PALACKY UNIVERSITY OLOMOUC

Faculty of Science
2012

Ph.D. Thesis
Physical Chemistry

STRUCTURE AND DYNAMICS OF HUMAN CYTOCHROME P450

Tereza Hendrychová

Department of Physical Chemistry
Supervisor: doc. RNDr. Michal Otyepka, Ph.D.

I, Tereza Hendrychová, declare that I have elaborated this thesis on my own with a kind support and supervision of Assoc. Prof. Michal Otyepka, Ph.D. with the use of literature listed at the end of this work.

In Opava 30th July 2012 _____

I would like to thank for great advice and support to my supervisor Assoc. Prof., Michal Otyepka, Ph.D. and to the members of the Physical Faculty in Olomouc, where I have found big help and friendly atmosphere.

My special thanks go to my colleagues at the Masaryk Secondary Agricultural School and College Opava, who helped me with many small unnecessary things, which were usually very important and they had to be done immediately. I am so happy that the headmaster and the deputy teacher of mentioned school abled me to finish my studies although it sometimes meant many compromises.

I would like to express the gratitude to my close family and friends, who supported me during all my studies and where I have found big support every time I needed.

TABLE OF CONTENTS

1. INTRODUCTION

"Only the educated are free."
Epictetus (55 AD – 135 AD), Discourses

Cytochrome P450 (CYP) enzymes play a prominent role in metabolism of eukaryotic and prokaryotic cell [1]. In humans, CYPs are involved in biosynthetic pathways of many endogenous compounds and metabolize the majority of xenobiotics, e.g. commercial drugs and pollutants [2-5]. They are responsible for many pharmacological and toxicological effects as well as drug-drug interactions [3, 6]. The mechanism, how CYPs accommodate and oxidize various substrates in highly stereo- and regio-specific manner, still remains the topic of great interest in the biochemistry of the CYPs with pharmacological implications.

Nowadays, there are many approaches, how to study the structure and other properties of CYPs, such as X-ray crystallography, NMR, cryo-EM and many other experimental methods. Results acquired by experimental techniques can be in some cases interpreted, complemented and enhanced by theoretical calculations. Among them, the molecular dynamics is often beneficial enabling simultaneous atomic and sub-picosecond resolution, hardly available for contemporary experimental methods. The molecular dynamics simulations provide unique view into the active site of the protein at an atomic resolution; it is able to give detailed information about structural changes in time. We can use it for description of flexibility, malleability and other properties of studied system in time (as shown later in this thesis).

The aim of this work is to study the structure and dynamics of human Cytochrome P450 enzymes mainly by molecular dynamics. We have employed molecular dynamic simulations at normal (0.1 MPa) and high (3 MPa) pressure conditions to study the flexibility and compressibility of CYPs in detail. Under both pressure conditions we described the flexibility of CYP's active sites together with their solvation. We have correlated results of molecular dynamics simulations with the spectroscopic data and with substrate preferences of studied CYPs [7, 8]. We have also shown how the flexibility of CYP2D6 active site was radically decreased by ligand (quinidine) binding [9]. From these works we concluded that there is a correlation between substrate preferences and flexibility

of CYPs. This implies that flexibility should be considered as significant contributor to properties, which determine substrate specificity of CYPs, and, therefore, should be taken into account in *in silico* predictions of metabolism of drugs by CYPs.

2. PROTEINS

Proteins are large molecules that are abundant in every living cell. They consist of basic amino acids, which are as beads sling on the file and make complicated structures. Every protein has its function which is conditioned by appropriate structure.

The protein structure can be hierarchically divided into four main levels [10]:

- Primary structure – represents the sequence of amino acids one after another, where each amino acid is connected to another one by a peptide bond.
- Secondary structure – is the geometrical organization caused by rotating of peptide bonds planes around C_α-atoms. Typical secondary structure elements are α-helices and β-sheets, which are stabilized mainly by H-bonds between –NH and –CO groups.
- Tertiary structure – express the arrangement of secondary structure elements in the 3D-space. The tertiary structure is stabilized by interaction of amino acids side chains.
- Quartery structure - is the superstructure where two or more tertiary structures are associated together (often also with cofactors) and create a functional protein.

Proteins are involved in many cellular processes, for example they act as building block, they transport many other chemicals, they play an important role in immunological response, they are hormones, receptors, provide biocatalysis as enzymes, and many other.

Enzymes are classified into six groups with corresponding EC numbers [10]:

EC 1: Oxidoreductases – catalyze oxidation/reduction reaction

EC 2: Transferases – catalyze transfer of a functional group

EC 3: Hydrolases – catalyze the hydrolytic reaction of various bonds

EC 4: Lyases – catalyze the cleavage reaction of various bonds by other mechanism than oxidation of hydrolytic reaction

EC 5: Isomerases – catalyze the isomeration changes in a molecule

EC 6: Ligases – catalyze the joining two molecules together through a covalent bond

2.1 Cytochromes P450

Cytochrome P450 (CYP) enzymes are heme-containing monooxygenases (EC 1 group) that play an important role in the metabolism of prokaryotic and eukaryotic cells. Mammalian CYPs are included in many biosynthetic pathways of many small regulatory molecules such as steroids, prostaglandins, fatty acid derivatives and many other [1]. They are also very important in metabolizing many xenobiotics, such as environmental pollutants and drugs and they mediate drug-drug interactions [2-5]. Thus interaction of drugs with CYPs can influence the efficacy of drug therapy [3].

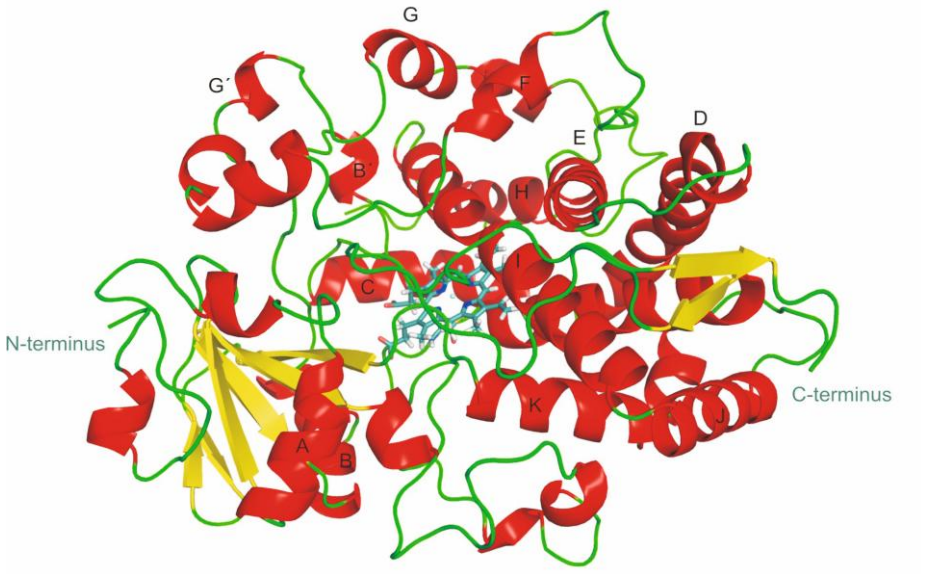
CYP are divided into families, where in one family the CYP sequences have higher than 40% sequence identity. The family is described by the Arabic number following the abbreviation CYP, e.g. CYP1 (the first family of the CYP). The letter following the Arabic number indicates a subfamily, where the sequence similarity of the subfamily members is higher than 55%, e.g. CYP1A (CYP belonging to the family 1 and subfamily A). The last Arabic number at the end of the abbreviation represents an individual gene, e.g. CYP1A2 (CYP from the family 1, subfamily A, gene name 2) [11].

2.2 Common CYP fold

Mammalian CYPs share common CYP-fold, which is predominantly α-helical. CYP active site is deeply buried in the structure and houses the heme molecule as a cofactor [1]. All α-helical segments, β-sheets, and loops are together with the overall CYP fold depicted on Figure **1**. The typical structural feature of CYP consists of the F/G segment (F-helix, F/G loop, G-helix) preceding highly structurally conserved I helix which is in the close vicinity to the heme cofactor [1]. Nevertheless the active site with the heme molecule is deeply buried in the protein; there are several access/egress pathways, which connect the hidden active site with the protein surface. Here, Wade et al. nomenclature of active site access channels is adopted [12]. The access to the CYP active site by the solvent channel is controlled by the G helix and the F/G loop and pathways 2 channels are immersed between G helix and the B/C-loop [12]. Another highly conserved parts of the CYP are A and C helices and parts of the K helix [13, 14]. Mammalian CYPs have quite a long insertion between F and G helices formed

by additional F´ and G´ helices [15]. Generally the more conserved element the closer to the heme molecule it is [1].

Figure 1 *The overall CYP fold (here shown for CYP1A2, PDBid: 2HI4). α-helical segments are depicted in red cartoons, β-sheets as yellow arrows and loops are shown as green tubes. Deeply inside CYP, the heme molecule is depicted in blue sticks (adopted from Ref. [16])***.**



Positioning and mutual orientation of α-helices and β-sheets, which make the CYP fold are well structurally conserved over the CYP families. In another words, the CYP fold is conserved throughout all CYPs.

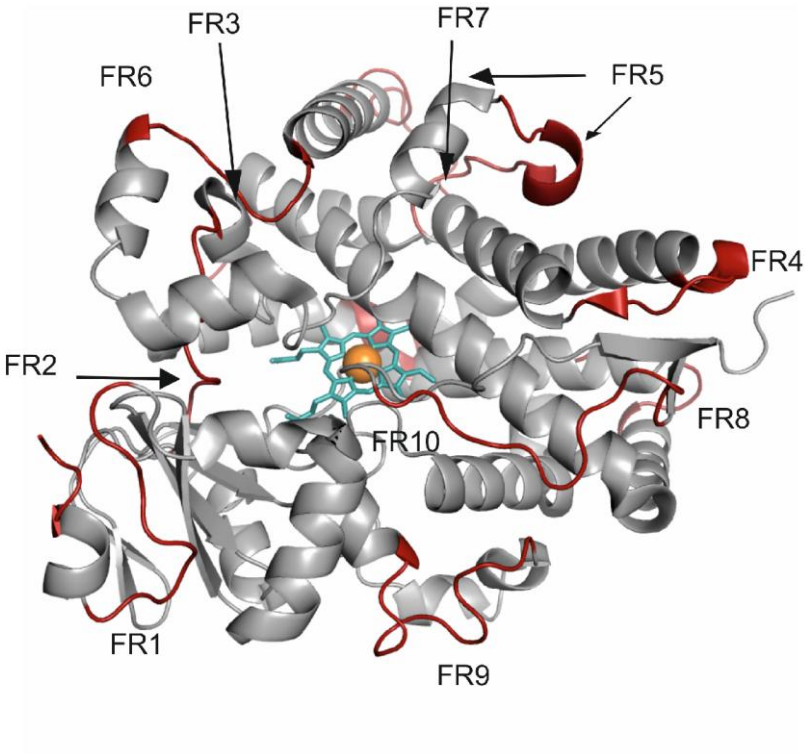
The CYP active site is deeply buried in the enzyme structure and therefore a substrate has to enter the active site by a channel connecting active site with enzyme exterior. Similarly, a substrate has to leave the active site by the same or another channel after reaction. Which channels are used to substrate/product passage to/from the active site, it is under an intensive scientific debate. However, analyses of available CYP crystal structures suggest that potential access/egress paths are not enough large to enable passage even of rather small molecules [12]. Enzymes are not rigid and therefore some breathing of their structures and internal channels can be anticipated. The fluctuation of access/egress channels has also been reported in many MD studies [8, 17, 18].

The flexibility of CYPs seems to be one of the key modulators of CYPs' enzyme specificity [19] and it plays an important role in the entry and accommodation of substrate in the active site [12, 20]. The α-helices and β-sheets making the core of CYP fold are rather rigid, that means that they do not

change their secondary structure or the position in the CYP's fold. There are also flexible regions located namely in the CYP distal part (on enzyme surface above the heme plane), which have enhanced fluctuations, i.e. higher B-factors (B-factor or the temperature factor reflects the flexibility of the system). The regions with the higher flexibility (plastic regions) were identified by Zhao et al. [21] from crystal structures. Later Skopalík et al. [18] introduced flexible regions (FRs), which were deduced from molecular dynamics simulations:

- FR1 – includes N-terminal domain loops (with the exception of the rigid β1 strand and residues 375-390)
- FR2 – B´/C loop
- FR3 – C/D loop
- FR4 – D/E loop
- FR5 – includes large segment beginning with E/F loop and G/H loop and joining all F/G segment
- FR6 – G´/G loop
- FR7 – H/I loop
- FR8 – J/J´ loop
- FR9 – includes the meander and residues 395-421
- FR10 – is formed by residues 460-475

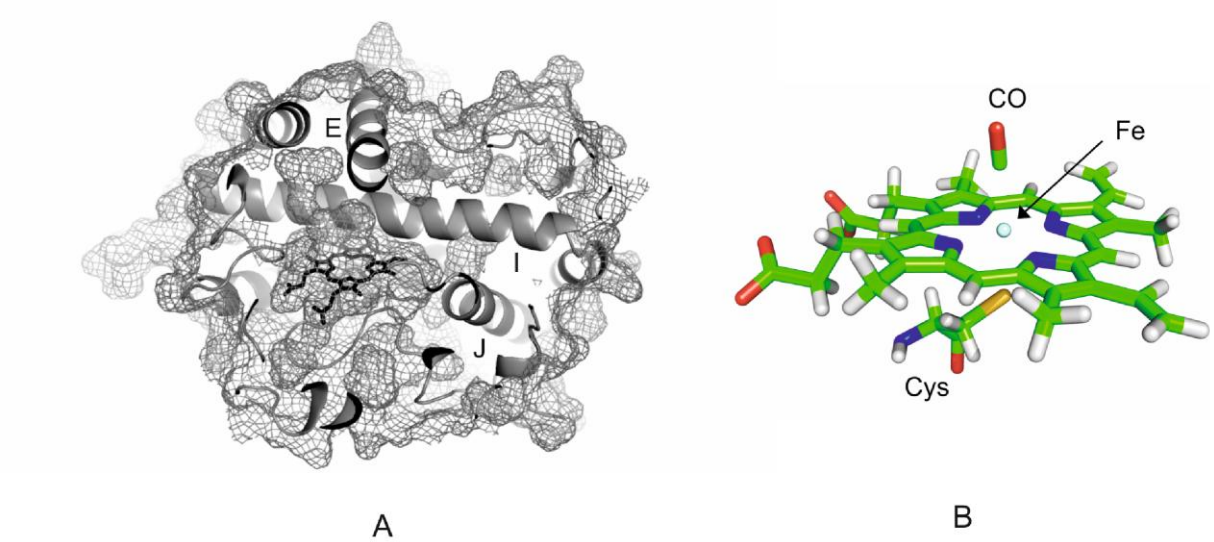
Figure 2 *Flexible regions (FR) are depicted in red color and they are labeled by respective labels (CYP3A4; PDBid: 1TQN). The heme molecule is depicted in blue sticks in the middle of the active site with the iron atom shown as an orange sphere (adopted from Ref. [16]).*



2.3 **The active site of the CYP**

The heme cofactor housed by the CYP active site (Figure 3) plays the dominant role in the catalytic cycle of CYPs (Figure 4). It consists of iron atom coordinated to a protoporphyrine ring and two additional axial ligands. At the proximal site there is a thiolate from a cystein side chain and at the distal site there are variable ligands that change during the catalytic cycle (e.g. water, oxygen, substrate etc.)[1]. The heme structure is described in Figure 3B.

Figure 3 *A – CYP2E1 with depicted heme molecule as sticks, the mesh represents a surface of CYP2E1, B – the structure of the heme molecule with the thiolate (Cys) and carbon monoxide (CO) molecule coordinated to the iron atom.*



Typical catalytic reaction of CYP is the monooxygenase reaction, where one atom of oxygen is inserted into an organic substrate (RH) while the other oxygen is reduced to water. This reaction can be summarized as follows:

$$RH + O_2 + NADPH + H^+ \rightarrow ROH + H_2O + NADP^+ \quad (1)$$

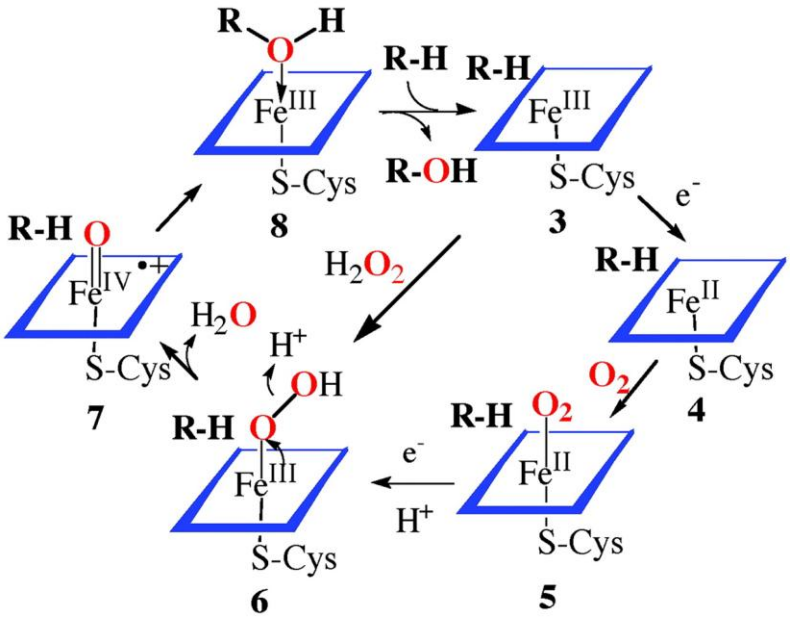
The CYP's cell cycle has several steps and it can be divided as follows (see Figure 4, [22]):

- the substrate binds to the enzyme; at the beginning there is a complex with the ferric state (Fe^{III}) of iron and water molecule; when the substrate enters the active site it splits the water molecule and induces the weaker interaction between the iron ion and water molecule; thus the complex

becomes good electron acceptor; and thus the enzyme-substrate adduct **3** is created

- the ferric state of cytochrome is reduced by an associated reductase with an NADPH-derived electron to the ferrous state (Fe^{II}) of CYP **4**
- the oxygen molecule binds to the ferrous heme to make a dioxygen complex **5**
- the protonation and other electron reduction leads to the Fe(III)-hydroperoxy complex **6**
- a reactive iron-oxo intermediate **7** is made by the protonation and heterolytic lceaveage of O-O bond; the water molecule is formed
- the final oxygen-atom transfer from the **7** complex to the substrate produce complex **8** and the dissociation reaction finishes the cycle

Figure 4 *Consensus catalytic cycle for oxygen activation and transfer by cytochrome P450 (adopted form* **[22]).**



The volume of the active site varies among CYP forms. We can roughly divide CYPs into three groups according to their active site volume. The first group has the smallest active site, with the active site volume around 300 Å. CYP2A6 and CYP2E1 can fit the first group of CYP with smallest active sites [8]. The second group of CYP with the moderate active site volume about 2,000 Å includes CYP2C9, CYP2D6 and CYP3A4 [8, 23]. CYP2B4 is the typical representative of the last group with the largest active site. It is worth to say that this structure

does not have a native conformation and the open structure is unnatural state caused by the cocrystalization of more CYP2B4 structures. The volume of its active site is approximately 12,500 Å [8]. This structure shows the large flexibility potential of CYPs.

The active sites of various human CYP P450 accommodate beside the heme molecule the number of water molecules. At the resting state of CYP P450 the active site is occupied by many water molecules, which have to enter the active site by various channels connecting the active site to surroundings (described later in chapter 2.4). The water molecules probably use the same channels as the substrate or product, but because waters are quite small molecules they can pass also various other ways [23]. It is believed that the number of active site water molecules reflects the relative size of the active site volume thereby the cytochrome except CYP2B4, the CYP3A4 should accommodate the largest number of water molecules [23]. The number of water molecules fluctuates in the time whilst the water molecule travels from the bulk solvent into the protein and from the active site to the bulk back [18, 23]. Some water molecules present in the active site have to be also displaced when the substrate binds into the active site.

2.4 **CYP active sites access/egress paths**

The catalytical reaction takes place in the deeply buried active site and the heme molecule can be considered as the engine for the reaction (Figure 4). So it is necessary for the substrate or other molecules to enter the active site and for the products to leave via various pathways. As described in the previous paragraph, it is known that the water molecules access the active site and surround the heme molecule [23, 24]. Active site access/egress pathways were identified and classified as ligand pathway channels [12, 25]. The numbering of channels corresponds to the rule that the pathway 1 has the highest average crystallographic thermal factor. Channel 2 is further divided into seven subclasses with the requirement of sharing at least one secondary structure element (see Figure 5, [25]). The most extensive analysis of access/egress channels was published by Cojocaru et al. [12], who applied the CAVER software [23] to available X-ray structures.

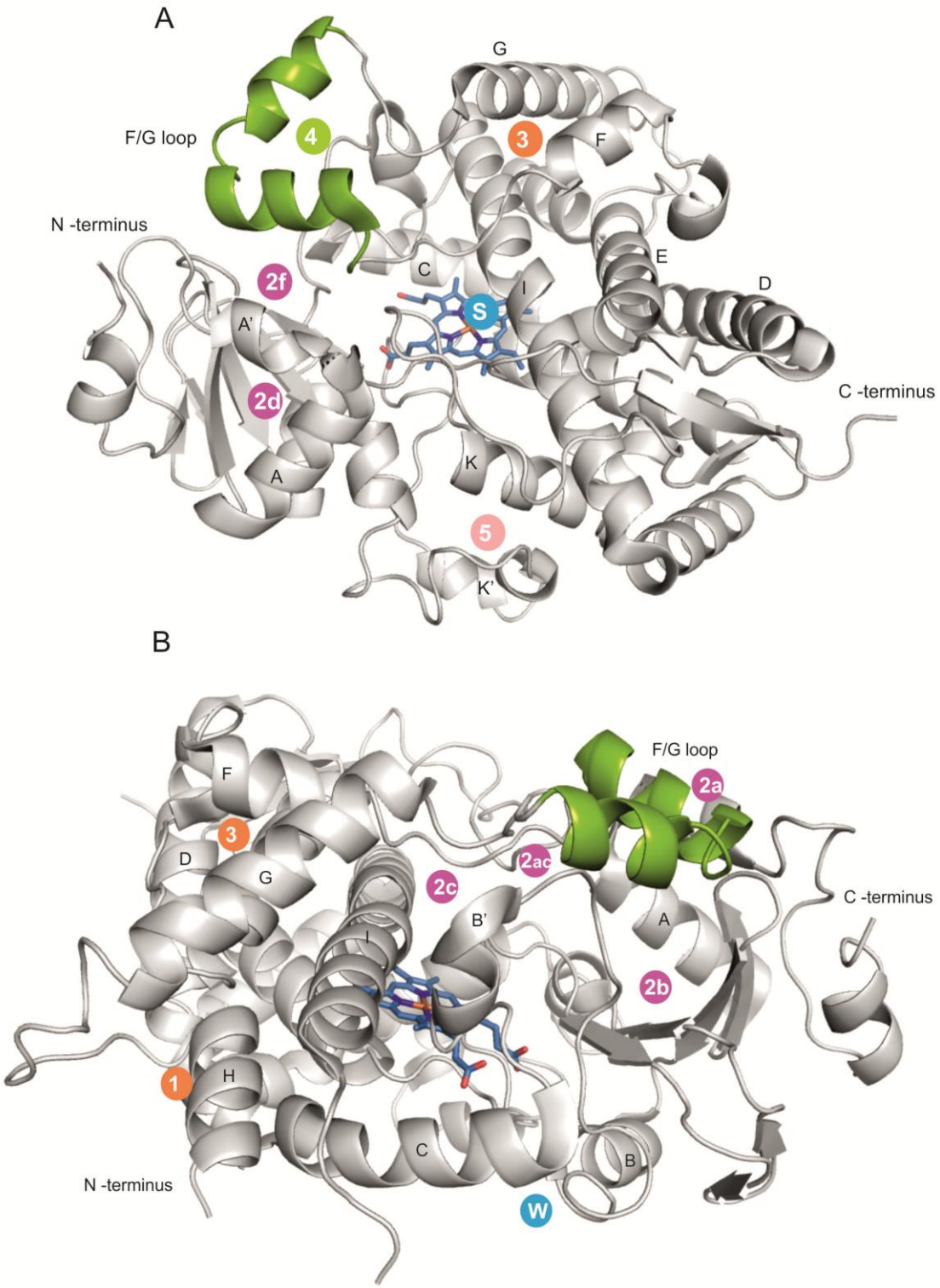
The functional role of mentioned channels attracts a large interest and has been subject of many studies, including e.g. random expulsion molecular dynamics (REMD) method [26-29] and steered molecular dynamics simulations [28, 30-32]. The REMD simulations showed channel 1 and 2c, 2e and Solvent channel favored by water molecules. Channels 2c, 2e and Solvent channel are also routes for egress of soluble products while 2a, described by steered MD, is an access way for hydrophobic products [25].

The following Table 1 and Figure 5 summarize the overview of CYPs pathways.

Table 1 *The overview of the access/aggress pathways in the human cytochrome (based on [12]*.

CHANNEL		THE LOCATION OF THE CHANNEL
1		egress between the C/C′ and H or L helices, close to the G-H loop and β2 sheet
2	2a	All 2 channels are lined by the B–C loop/B′ helix; egress between the F–G loop, B′ helix/B–B′ loop/B–C loop and the β1 sheet
	2b	egress between the B–B′ loop and the β1 and β3 sheets
	2c	egress between the G and I helices and the B′ helix/B–C look
	2ac	egress between the tip of the B–C loop (or B′ helix) and the G-helix, between channels 2a and 2c
	2d	egress between the N-terminus and helices a/A′ and A
	2e	egress through the B–C loop
	2f	egress between the F′ helix/F–G loop, and the β5 sheet
3		egress between the F and G helices or at the E–F loop
4		egress through the F–G loop
5		egress between K and K′ helices
Solvent		egress between E, F, and I helices and the β5 sheet
Water		egress at the base of the B–C loop near the C-terminus of the B helix

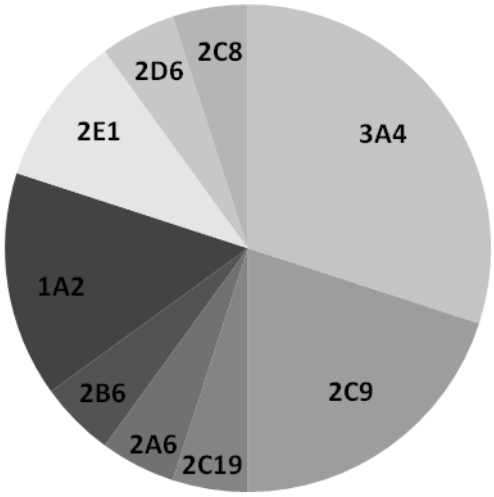
Figure 5 *The channels mapped to the structure of the CYP3A4 (PDBid: 1TQN; based on [25]).*



2.5 **Human Cytochrome P450**

Since the time we know the human genome of all genes we have identified 57 human CYPs which are divided into 18 families and 43 subfamilies [33]. The most of them are expressed primarily in the endoplasmic reticulum and only six of them are located in the mitochondria [1]. It is obvious that each CYP has different site of expression, relative abundance, polymorphism, substrates, reactions, etc. Many of human CYPs have solved structure by X-ray crystallographic method (Table 5).

Figure 6 *The contributions of particular CYP to the drug metabolism (based on ref.* [2]*).*



Here I would like to describe selected microsomal CYPs (the most significant forms for metabolism of drugs), namely CYP1A2, CYP2A6, CYP2C9, CYP2D6, CYP2E1, CYP3A4 and rabbit CYP2B4. These CYPs have been extensively studied in my work. I have chosen CYP3A4, 2C9, 1A2 and 2E1 because they metabolite the most of known drugs and we have supplied it by CYP2D6 and CYP2A6 because of their active sites, which are quite narrow (cf. Figure 6).

CYP1A2

Cytochrome P450 1A2 (CYP1A2) belongs to the CYP1 family together with CYP1A1 and CYP1B1. It is localized in the endoplasmic reticulum and in liver tissues only. The expression is substantial, it is 10-15% of the total P450s [1]. CYP1A2 has less than 40% amino acid sequence similarity when compared to other human microsomal cytochromes [34]. CYP1A2 is involved in the metabolism of many xenobiotics, such as caffeine, melatonine, acetaminophen (3') (known as a paracetamol) and many other marked drugs such as lidocaine, zolmitriptan, flutamide, clozapine, drugs for treating schizophrenia, etc. [34-37]. There are also several inhibitors known from clinical work, such as furafylline or fuvoxamine [1, 38] Also drug interactions with CYP1A2 have been reported and it was mentioned that the high activity of CYP1A2 is associated with the enhanced risk of colon cancer [39].

CYP2A6

Cytochrome P450 2A6 (CYP2A6) belongs to the family CYP2, which is divided into 13 other subfamilies and it includes 16 types of human cytochrome P450. It is expressed mainly in the endoplasmatic reticulum of livers, but we can find it also in other tissues such as nasopharyngeal region, trachea, lung, esophageal mucosa [1]. The overexpression of CYP2A6 in the liver tissue is associated with chronic inflammation and cirrhosis and causes liver cancer [40]. Persons with low CYP2A6 activity are believed to smoke less and they might have lower cancer risk [41]. It was also proved that many carcinogens from smoking are activated by CYP2A6, but not all experiments are in the coincidence. The typical CYP2A6 substrates are coumarine, cotinine (metabolite of nicotine), nicotine and 7-ethoxycoumarine. Typical inhibitors are grapefruit juice containing mercaptane [42].

CYP2C9

Cytochrome P450 2C9 (CYP2C9) is one of the major enzymes involved in the drug metabolism. It belongs to the CYP2 family. It is primarily hepatic CYP and the level of expression is the highest from all known cytochromes, except CYP3A4 [1]. It is typical for all CYP 2C enzymes that they are not expressed in fetal liver, but the level raises quickly in the first month after the birth [43]. Another place of occurrence is in the small intestines [1].

CYP2C9 is involved in oxidation of ~16% of current used drugs drugs [44], where it has a selectivity for oxidation of mainly small, lipophilic anions (e.g. non-steroidal anti-inflamantory drugs as fluorbiprofen, ibuprofen and diclofenac) and hypolipidemic agents (e.g. gemfibrozil) [45]. It also regulates blood pressure through the participation in the synthesis of the arachidonic acid epoxides [46]. To date two crystal structures with the substrate are known. The first one with S-warfarine [47] and the second one with flurbiprofen [45].

CYP2D6

Cytochrome CYP 2D6 (CYP2D6) belongs to the CYP2 family and it is expressed mainly in the liver tissue, but we can find it also in lung and brain (localized in the large principal neurons). It accounts 5% of total amount of P450 [1]. It is involved in the oxidative reactions of ~30% of marketed drugs (mainly CNS and cardiovascular drugs) [48], which includes aliphatic and aromatic hydroxylation, heteroatom dealkylation, etc [1]. The common substrate has at least one basic nitrogen with the distance of 5 Å or 7 Å to oxidation site and a negative molecular electrostatic potential above the planar part of the molecule. One of the strongest inhibitor is quinidine, which is often used as diagnostic involvement of CYP2D6 in the drug metabolism [49]. The drug-drug interaction and the influence of CYP2D6 were profoundly studied also thanks to the polymorphism. The polymorphic effect occurs in 7% of Caucasian population, where these people are poor metabolizers [50]. Comparing CYP2D6 to other cytochromes, CYP2D6 is not inducible [1].

CYP2E1

Cytochrome CYP 2E1 (CYP2E1) is the CYP2 family member and it is expressed mainly in the liver tissue and also in extrahepatic sites such as lung, esophagus, small intestines, brain, nasal mucosa and pancreas [1]. It is not presented in the fetal liver, but it appears within a few hours after the birth and the level rises during the first year of childhood [51]. The expression in moderately abundant and as in previous cytochrome CYP2D6 it is polymorphic.

CYP2E1 has a significant role in human health and drug metabolism. It is known that the level of CYP2E1 varies depending on alcohol consumption, diabetes, obesity and fasting [52]. This cytochrome together with alcohol dehydrogenase and aldehyde dehydrogenase is important for the decomposition

of alcohol into acetaldehyde and acetate. Furthermore it oxidates many low molecular weight cancer suspects such as nitrosamine, benzene, styrene, vinyl chloride, acetonitryle, vinyl carbamate and many others [1]. There are also many known inhibitors, where among the most well known belong disulfiram, known as Antabus (used by patients in alcohol aversion therapy). There are also many inhibitors from natural origin, which come from onions, garlic and cruciferous vegetable [1].

CYP3A4

Cytochrome CYP 3A4 (CYP3A4) belongs to the CYP3 family, which includes only one subfamily. It is the most abundant cytochrome in the human body and it plays very important role in drug metabolism. It is expressed in livers and small intestines and it makes ~25-30% of total P450 in small intestines and in liver the level is even higher. It was also found in some extrahepatic tissue such as lung, stomach, colon and adrenal [1]. Cytochrome P450 metabolizes ~50% of marketed drugs among those are mainly chemotherapeutic drugs. It catalyzes also many atypical reactions such as desaturation, oxidative carboxylic acid ester cleavage, etc. [1]. There is a quite large active site, which can accommodate many large substrates such as cyclosporine, bromocriptine, progesterone, erythromycin and many others [20].

CYP2B4

Cytochrome CYP 2B4 (CYP2B4) belong to the CYP2 superfamily, but it is mentioned here as the last one, because its X-ray structure was obtained after the separation not from the human cells but from the rabbit liver microsomes. It was firstly separated in 1973 by Coon et al. [53] under the name LM2. The ligand free CYP2B4 structure is considered to be special for its open conformation [54]. It has large open cleft between α-helical and β-sheet domains without the perturbing the overall P450 fold and this open formation is reversible depending on the substrate bounding into the active site, when it closes [54, 55].

As preceding CYPs it also converts lipophilic xenobiotics to more polar metabolites and it is believed that the wide conformational flexibility may be responsible for wide range of metabolizes xenobiotics [54].

Table 2 *Summary of solved structures of human cytochromes P450 (date back to 1/5/2012).*

CYP	TISSUE SITES	SUBCELLULAR LOCALIZATION	PDBid	LIGAND	REF.	CYP	TISSUE SITES	SUBCELLULAR LOCALIZATION	PDBid	LIGAND	REF.
1A2	liver	ER	2HI4	α-naphtoflavone	[34]	**2E1**	liver, lung, other tissue	ER	3KOH	8-(1H-imidazol-1-yl) octanoic acid	[57]
1B1	extrahepatic sites (e.g. lung, kidney)	ER	3PM0	α-naphtoflavone	[56]						
2A6	liver, lung, many extrahepatic sites	ER	2PG6		[58]				3GPH	10-(1H-imidazol-1-yl) decanoic acid	[57]
			2PG7		[58]						
			3EBS	phenacetin	[59]	**2R1**	liver	ER	3C6G	vitamin D3,	[60]
			2FDY	4,4´-dipyridyl disulfide	[61]					β-cyclodextrine	
			2PG5	1,2-ethandiol	[58]				3DL9	1-hydroxy-vitamin D2,	[62]
			1Z10	coumarin	[63]					β-cyclodextrine	
			1Z11	methoxalen	[63]				3CZH	vitamin D2,	
			2FDU	N,N-dimethyl(5-(pyridin-3yl)furan-2yl)methanamin	[61]					β-cyclodextrine	[64]
			2FDV	N-methyl(5-(pyridin-3-yl)furan-2-yl)methan amine	[61]	**3A4**	liver, small intestine	ER	1TQN		[65]
			2FDW	(5-(pyridin-3-yl)furan-2-yl)methanamin	[61]				1W0E		[66]

2A13	nasal tissue	ER	2P85	indole	[67]				1W0F	progesterone	[66]
2B6	liver, lung	ER	3IBD	4-(4-chlorophenyl) imidazole	[68]				1W0G	metyrapone	[66]
2C8	liver	ER	1PQ2	palmitic acid	[69]				2J0D	erythromycin A	[20]
			2NNH	retinoic acid,	[70]				2V0M	ketoconazole	[20]
				palmitic acid					3NXU	ritonavir	[71]
			2NNI	montelukast,	[70]	**7A1**	liver	ER	3DAX		[72]
				palmitic acid		**11A1**	adrenals, other steroidogenic tissues	Mit	3NA0	20,22-dihydroxy cholesterol	[73]
			2NNJ	felodipine,	[70]						
				palmitic acid		**19A1**	steroidogenic tissues, adipose, brain	ER	3EQM	androstenedione	[74]
			2VN0	troglitazone,							
				palmitic acid	[70]	**46A1**	brain	ER	2Q9G		[75]
2C9	liver	ER	1OG2		[47]				2Q9F	cholesterol-3-sulphate	[75]
			1OG5	S-warfarin	[47]				3MDV	clotrimazole	[76]
			1R9O	flurbiprofen	[77]				3MDT	voriconazole	[76]
2D6	liver	ER	2F9Q		[78]				3MDR	tranylcypromine	[76]
2E1	liver, lung, other tissue	ER	3E4E	4-methyl-1H-pyrazole	[79]				3MDM	thioperamide	[76]
			3E6I	indazole	[79]	**51A1**	liver, testes	ER	3JUV		[80]
			3LC4	12-(1H-imidazol-1-yl)	[57]				3JUS	econazole	[80]
				dodecanoic acid					3LD6	ketoconazole	[80]

3. METHODS

The next several pages deal with theoretical methods which were used in this work. The biggest attention is paid to the molecular dynamics simulations, which were used throughout all the work.

3.1 Molecular dynamics simulations

Molecular dynamics simulations are computer techniques based on theoretical chemistry methods and experimental data. The theoretical chemists use the basic data from experiment, such as structural and conformational data, many parameters describing various constants, lengths and forces, etc. and use them to calculate and predict the structure of studied system and its behaving in a time [81]. Molecular dynamics simulation is a powerful tool used for investigating the structure and dynamics under various conditions, predicting ligand binding and enzyme-reaction mechanism, denaturation of proteins or their folding on time scales up to microseconds.

Molecular dynamics simulations are based on the basic Newtonian equations of motion (2):

$$F_i = m_i \cdot a_i = m_i \cdot \frac{\partial v_i}{\partial t} = m_i \cdot \frac{\partial^2 r_i}{\partial t^2} = -\frac{\partial V}{\partial r_i}, (2)$$

where F_i is the force applied along Cartesian coordinates r_i on each atom *i*; m_i is the mass of atom *i*; a_i and v_i is the acceleration and velocity of the atom *i*, respectively, and *V* is the potential energy of the system. The potential energy is in classical molecular dynamics simulations calculated from empirical potential (molecular mechanics). A simulation first determines the force on each atom equaling to the negative gradient of the potential energy with respect to the position of atom *i*. The change in velocity is equal to the integral of acceleration over time. Then the motion equation is integrated for every explicit atom N in the system in time in small time steps, Δt (usually 0.5-2 fs) [82].

The potential energy of all systems in molecular mechanics can be calculated using various force fields (see below). The biggest utilization of the molecular mechanic is in energy minimization, which could be done by various methods, for e.g. simulated

annealing, the Metropolis algorithm, Monte Carlo methods, etc. The aim is to obtain the structure with the lowest energy.

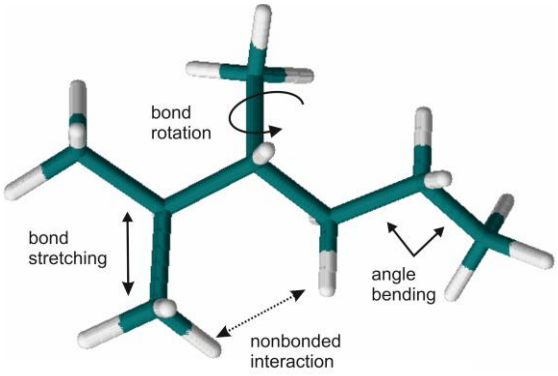
Molecular mechanic methods are based on several rules:

- Nuclei and electrons are lumped together and treated as unified atom-like particles.
- Atom-like particles are treated as spherical balls.
- Chemical bonds between particles are considered as springs.
- Interactions between particles are treated using potential functions derived from classical mechanics.
- Individual potential functions are used to describe different types of interactions.
- Potential energy functions rely on empirically derived parameters that describe

the interactions between sets of atoms [82].

The potential energy is described as a sum of bonded and non-bonded interactions. The bonded interactions span bond stretching, angle bending, dihedral term and the non-bonded interactions include long-range electrostatics and van der Waals forces. The potential energy is described by following equations (Figure 7).

Figure 7 *The picture of individual molecular mechanic terms and equations describing the potential energy.*



$$\mathbf{E_{TOT} = E_{BON} + E_{ANG} + E_{TOR} + E_{VDW} + E_{ELS}} \quad (3)$$

$$E_{BON} = \sum K_r (r - r_0)^2$$

$$E_{ANG} = \sum K_\theta (\theta - \theta_0)^2$$

$$E_{TOR} = \sum K_\phi \left[1 + \cos(n\phi - \phi_0)\right]$$

$$E_{VDW} = \sum_{i<j} \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}}$$

$$E_{ELS} = \frac{1}{4\pi\varepsilon_0} \frac{q_i q_j}{r_{ij}}$$

where K_r, K_θ and K_ϕ are force constants for bond, angle and dihedral terms respectively, r describes the distance and r_0 is an equilibrium distance, θ defines angle and θ_0 defines equilibrium angle, ϕ is a dihedral and ϕ_0 is an equilibrium dihedral; n is the periodicity of the Fourier term; i and j are indexes of interacting atoms with the distance r_{ij} and van der Waals parameters A_{ij} and B_{ij}; ε_0 is the dielectric constant of vacuum.

3.2 AMBER

Equation (3) reproduces the potential energy at an atomic level. The potential energy function together with the parameters accompanying it (r_0, θ_0, ϕ_0, etc.) makes the force field. Nowadays we can choose from many force field families such as CHARMM [83], OPLS [84], AMBER [85], GROMOS [86] and many others. In this thesis the AMBER family force field is used. AMBER stands for the abbreviation: Assisted Model Building and Energy Refinement, the molecular dynamics package, which was developed in 1980s by Kollman group [87]. It has been used for modelling proteins, small organic molecules, carbohydrates and nucleic acids. This approach adopts ideal bond lengths and angles from microwave spectroscopy and neutron diffraction, torsion constants are taken from microwave and NMR, nonbonded parameters are from crystal packaging calculations and charges were derived from quantum chemical calculations [88].

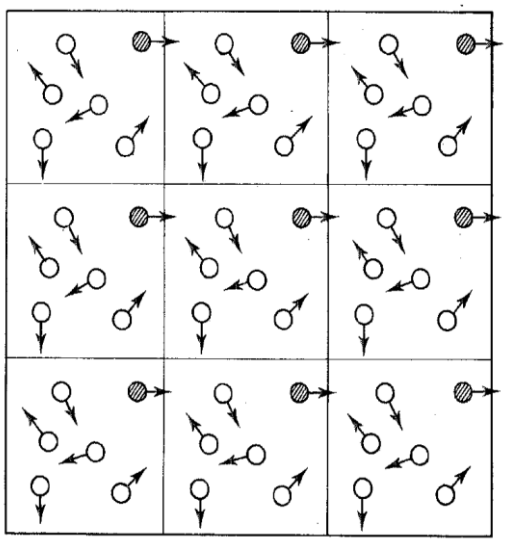
As the increasing computations were came out and the computer technology was developing the second generation force field ff94 [89]. Force field ff94 implements the set of new parameters for better describing for the pair-additive potential for condensed phase, but the partial charges derived by this way overestimate molecular dipole

moments and thus neglect explicit polarization [90]. Nevertheless some studies have shown that it overemphasizes the helical conformation and thus the ff96 was developed [91]. Force field ff96 improved the peptide backbone ψ/φ torsion parameters [91]. In 2000, ff99 was introduced as the revision of ff94 and ff96 based on the new quantum mechanical data for torsion parameters [92]. Later, Simmerling et al. proved that the ff99 does not accurately represent glycine behaviour and thus it causes the over-stabilization of the α-helical peptide conformations and destabilize β-bend propensity. That is why ff99SB was developed to improve the balance between α-helical structures and β structures [92]. The third-generation AMBER force field ff03 was introduced by Duan et al. where all ψ/φ parameters were revised and partial charged were recalculated using B3LYP/cc-pVTZ method, where the solvent polarization was solved with esp=4. Other corrections followed, ff99SB* [93] and ff03* [93] corrected ψ backbone dihedral correction of ff99SB and ff03 respectively. Recent studies has been showing new reparameterization of known force fields mainly based on fixing parameters of backbone and side chains mainly of small proteins and salvation free energies of small molecule [94, 95].

3.3 Boundaries

Setting up the boundaries for molecular dynamic simulation is crucial because it enables us to describe macroscopic properties of the system by using relatively small amount of particles. To represent a real system enough reliably and to avoid any unwanted behavior at the boundary, the periodic boundary conditions (PBC) are used. The principle of this method is simple. The system is placed into the box, which is replicated in all directions to give an array (Figure 8). If the particle leaves the box in one direction, it is replaced by the same particle that enters from the opposite site. That means that the number of particles in a box is constant. Nevertheless, this method has some limitations. It does not achieve fluctuations with the wavelength greater than the length of the cell [81].

Figure 8 *Periodic boundary conditions in two dimension (adopted from [81]).*



The protocol of molecular dynamics simulation

The setup of a molecular dynamics simulation can be from practical point of view divided into several steps, which are called the simulation protocol. The protocol should be well tested for the system under consideration. Here I present the molecular dynamics protocol used in our laboratory.

1. Obtaining initial coordinates (NMR or preferentially X-ray structures) from PDB database (http://www.pdb.org/pdb/home/home.do); building the missing structural elements (amino acid side chains or larger parts usually using a suitable template).
2. Developing force field parameters for non-standard residues according to Cornell et al. procedure [96].
3. Choosing the appropriate force field for studied systems, adding missing hydrogens, adding counter ions (usually Na^+, Cl^-, some other could be used), setting periodic boundary conditions, solvating whole system by explicit solvent model.
4. Energy minimization and equilibration of solvent, short MD simulation of solvent (to balance system density ~1 g/mL), energy minimization of all atoms in our system, warming the system to selected temperature (usually to 298 K).
5. Setting up simulation parameters, such as coupling with thermostat, barostat, etc.
6. Running MD production on the required time scale.

7. Evaluating convergence determination (RMSD, Rg, thermal B-factors, etc.).
8. Analyses of trajectories.

3.4 Solvation

Most of the biological processes take place in a water solution so the influence of the surrounding environment is obvious. Today molecular modeling approach provides two types of solvent, explicit and implicit model of a solvent [81].

Explicit solvent

In the explicit solvent, each water molecule is represented as a single individual. Every water molecule has its own geometry and molecular mechanic parameters, mainly atomic charges and Lenard-Jones parameters. Thus the detailed information about the water information is provided. The big disadvantage of using the explicit water model is the high demands to computers and increasing cost of the computing time. So far many explicit water models have been designed but there are some preferred. Among them the most popular explicit solvents for biomolecules are TIP*n*P [97-99] and SPC/E [100]. It is worth noting, that it is generally accepted that a choice of solvent model in molecular dynamics simulations has negligible impact on structure and dynamics of folded protein solutes [101].

4. RESULTS AND DISCUSSION

This chapter introduces and discusses obtained results and it is divided into three parts. In the first subchapter entitled "Flexibility of human cytochrome P450 at normal and high pressure simulation" – the flexibility of some human liver microsomal CYPs analyzed by all-atomistic molecular dynamics simulations under normal (0.1 MPa) and high (3.0 MPa) pressure is presented.

The second subchapter – "Dynamics and hydration of the active site of mammalian cytochrome P450" – is aimed at the long time scale (100+ ns) simulations of selected CYPs. The structural variation of individual CYPs is examined, mainly in terms of active site volume changes and solvation, and the opening of access channels during the simulation is mentioned.

The third subchapter – "Binding of quinidine into the cytochrome P450 2D6 active site" – is focused on the changes in structure, flexibility and compressibility induced by quinidine binding to the active site of CYP2D6.

4.1 Flexibility of human Cytochrome P450 at normal and high pressure simulation

The active site of CYP is deeply buried in the structure of protein. Each substrate enters the active site via various pathways, it is changed into the product in the appropriate place of the active site and the product is released out of the enzyme. The flexibility of enzyme is believed to play an important role in the entry of the substrate and release of the product in and out of the CYP actives site, respectively. The X-ray analysis can grasp the CYP's structural properties very well, but it provides only limited and indirect evidence regarding flexibility. Hence we enhanced information available from crystal structures by other techniques, such as molecular dynamics simulations and electronic and vibrational spectroscopies. UV/VIS spectroscopy at normal and high pressure provides information about the compressibility of the active site. Molecular dynamics simulations provide direct information about atomic motions on nanoseconds time scale (with sub ps resolution). We applied all-atomic molecular dynamics simulations under normal (NP, 0.1 MPa) and high pressure (HP; 300 MPa) together with UV/VIS and resonance Raman spectroscopy to five human liver cytochromes P450 (CYPs 1A2, 2A6, 2C9, 2D6 and 3A4) to gain the information about compressibility and flexibility of their active sites.

Common features of studied CYPs

We run 10-ns-long molecular dynamics simulations of five human liver cytochromes P450 (CYPs 1A2, 2A6, 2C9, 2D6 and 3A4. In all cases the overall topology of the enzyme did not significantly deviate from the X-ray structure, which can be documented by average RMSD of backbone atoms with values from 1.28 to 1.85 Å (Table 3).

Table 3 *Key data of the MD simulations (adopted from [7]).*

		NP				HP			
CYP	initial PDB	t[a] (ns)	R_g[b] (Å)	RMSD[c] (Å)	B_f[d]	t[e] (ns)	R_g[b] (Å)	RMSD[c] (Å)	B_f[d]
1A2	2HI4	10.0	22.73±0.05	1.35±0.08	9.2	10.0	22.57±0.05	1.09±0.07	8.7
2A6	1Z10	10.0	22.55±0.06	1.28±0.14	11.1	10.0	22.24±0.04	1.12±0.14	8.1
2A6	1Z10[f]	5.0	22.63±0.05	1.37±0.08	-	-	-	-	-
2D6	2F9Q	10.0	22.52±0.06	1.84±0.07	11.4	10.0	22.29±0.05	1.50±0.10	11.0
2C9	1OG2[g]	10.0	22.41±0.04	1.85±0.10	10.5	10.0	22.25±0.06	1.42±0.07	8.1
2C9	1OG2[f,g]	8.5	22.67±0.05	1.81±0.09	-	-	-	-	-
3A4	1TQN	10.0	22.66±0.05	1.38±0.06	9.7	10.0	22.54±0.05	1.12±0.06	8.7
3A4	1TQN[f]	5.0	22.73±0.06	1.33±0.06	-	-	-	-	-

[a] duration of MD production simulations after equilibration (5 ns long) under normal conditions (T = 298.15 K, p = 1 atm)
[b] R_g is the mean radius of gyration of C_α, C, N atoms ± SMD calculated from the last 3 ns of the respective simulations
[c] RMSD is the mean root-mean-square-deviation ± SMD of C_α, C, N atoms from the starting X-ray structure calculated from the last 3 ns of the respective simulations
[d] mean temperature factors averaged over backbone atoms from the last 1 ns of the respective simulations
[e] duration of MD simulations under high pressure conditions (T = 298.15 K, p = 3.000 atm), before the system was equilibrated for 5 ns under normal conditions (T = 298.15 K, p = 1 atm)
[f] data taken from our previous MD simulations [102]
[g] 1OG2 structure was modified according to the WT-CYP2C9 sequence

The active site in a close proximity of the heme molecule (the space around the heme in the distance not more than 5Å) is relatively rigid and the flexibility is localized mainly to the solvent exposed regions at both normal and high pressure conditions. MD simulations identified highly flexible regions, which coincided with flexible regions previously described by Skopalík et al. [18] (cf. chapter Methods, Flexible regions – FR). Under HP the average RMSD of the backbone atoms and temperature B-factors decrease, which speaks about the decreasing size of molecule and flexibility, respectively, under HP (Table 3). The high pressure causes increase in density, which is also reflected in the active site, where enzyme backbone atoms and water molecules get closer to the heme molecule. The mean coordination number of water molecules (CN^W) in the proximity of heme was determined by integration of radial distribution function (RDF) of water oxygen atoms found close to the carbon monoxide at the distance of 4.0 Å ($CN^W_{4.0}$) and 5 Å ($CN^W_{5.0}$) from the center of the C-O bond. To measure how close are protein atoms to heme, the coordination number of backbone atoms (CN^B) was introduced by integrating the respective RDF at the distances of 4.0 and 5.0 Å. The

decreasing size of the active site under HP is documented by increasing values of CN^B and CN^W.

Active site properties

The CN^W and CN^B are the coordination numbers of water molecules and backbone atoms to the heme molecule (CO molecule), respectively. To some extent, the CN^W reflects the water solvation of the active site (namely in the close proximity of heme cofactor). The number of the water molecules in the active site of all studied CYPs increase under HP with the exception of CYP3A4 (cf. Table 4). The different behavior of CYP3A4 can be explained in terms of its active site topology. As the only CYP considered in this study, CYP3A4 does not have a steric restriction above the iron atom and thus the high pressure may push out the water molecule from the cavity.

The active site compressibility can also be deduced from UV/VIS spectroscopy measuring Soret band shift with pressure. The Soret band is occurred after absorption of visible light by molecule in narrow range and after UV excitation the red fluorescence is occurred [103, 104]. The Soret band shift α value (Table 4) is traditionally considered to reflect an active site compressibility [105]. The α value was the lowest for CYP2A6 and CYP1A2 (α_{CYP2A6} = -0.233 and α_{CYP1A2} = -0.253, Table 4). The pressure perturbative UV/VIS spectroscopy documents also their high stability against denaturation. These CYPs are known to have relatively small active sites (the volume of the CYP2A6 active site is ~300 Å; the CYP1A2 active site volume is ~375 Å [106], as described by Otyepka et al. [15] earlier). More negative α values were observed at CYP2C9 and CYP2D6 (-0.273 and -0.301 cm^{-1}/MPa, respectively). We also observed the change of CYP2C9 and CYP2D6 to a denatured P420 state at medium hydrostatic pressure (100-150 MPa). The CYP3A4 showed similar denaturation but it appeared to be the most compressible one (α value of -0.449 cm^{-1}/MPa) with the steepest slope of the correlation between Soret band shift and the pressure obtained. This α value may indicate the high degree of compressibility of CYP3A4 active site. Our findings are in coincidence with latest research of CYP's behavior and their Soret's band shifts [107].

We have correlated the α factor (mentioned in the previous paragraph, [105]) with the data obtained from the MD simulations to gain the interpretation of the effect of high pressure on the UV/VIS spectra. Clear correlations were found between –α and CN^Ws at

4 and 5 Å under both NP and HP (see Table 4). The linear dependence of α on ΔCN^W and ΔCN^B at 5Å was found, according to

$$-\alpha = -0.14(0.06)\Delta CN^{W}_{5.0} + 0.45(0.27)\Delta CN^{B}_{5.0} + 0.29(0.05), \qquad (1)$$

where numbers in parentheses are the respective standard deviations. This correlation shows that the changes in the UV/VIS spectra under HP may reflect structural rearrangements in close proximity to the heme.

Table 4 *Coordination numbers of water molecules and CYP backbone atoms toward the CO heme ligand (based on [7])* .

	NP				HP				HP-NP				
CYP	CN^W 4.0 Å	CN^W 5.0 Å	CN^B 4.0 Å	CN^B 5.0 Å	CN^W 4.0 Å	CN^W 5.0 Å	CN^B 4.0 Å	CN^B 5.0 Å	ΔCN^W 4.0 Å	ΔCN^W 5.0 Å	ΔCN^B 4.0 Å	ΔCN^B 5.0 Å	-α cm^{-1}/MPa
1A2	0.84	1.95	2.61	4.23	1.02	2.27	2.15	4.29	0.18	0.32	-0.46	0.06	0.253 [108]
2A6	0.01	0.06	0.60	2.76	0.18	0.57	0.53	2.88	0.18	0.51	-0.07	0.12	0.233
2D6	1.00	1.19	0.68	3.39	1.11	2.17	0.72	3.58	0.11	0.98	0.04	0.19	0.301
2C9	1.09	2.54	0.43	2.51	1.78	3.74	0.50	2.91	0.69	1.19	0.07	0.40	0.273
3A4	3.00	4.47	0.33	2.48	2.80	4.16	0.36	2.69	-0.20	-0.31	0.04	0.21	0.449
r	-0.97	-0.86	0.41	0.42	-0.89	-0.72	0.41	0.42	0.64	0.65	-0.41	-0.18	

[a] CN, coordination number of the CO heme ligand; the changes of CN, i.e ΔCN (= CN(HP) – CN(NP)) are calculated from NP and HP pressure MD simulations. The CN^W and CN^B values are the coordination numbers of water molecules and CYP backbone atoms, respectively, calculated at two distances, 4.0 and 5.0 Å (see 2.1 for details). The experimental parameter -α is related to compressibility and was obtained from the UV/VIS high pressure experiments; pressure applied up to 250 MPa, with CYP1A2, up to 300 MPa . The Pearson's correlation coefficient *r* is shown in the last row.

Flexible regions

Ten flexible regions (FRs) were identified in human Cytochromes P450 from MD simulations by Skopalík et al. [18]. These flexible regions were also observed in CYPs considered here and they are summarized in Table 5.

All ten flexible regions were observed in all systems, but some FR regions are more intensive in some CYPs (described later). FR1 localized in the N-terminal domain loops was presented in all observed CYPs at normal and high pressure with the exception of the CYP2C9 at normal pressure. FR2, FR3 and FR4 (B'/C, C/D and D/E loops, respectively) are the least exposed regions and we observe lower flexibility in these regions under high pressure. FR5, FR6 and FR8 (E/F and G/H loop, G'/G loop, J/J'

loop, respectively) are pronounced in all CYPs with some exceptions and the lower flexibility is observed in some studied CYPs under the high pressure. FR7, FR9 and FR10 (H/I, meander and amino acids residues 395-421 and amino acids residues 460-475) are observed in all studied CYPs.

Some of the FRs can be important for substrate access and product release, because they line walls of access/egress paths. Regions FR2, FR3 and FR6 partially contribute to active site access channel 2 and thus it may regulate substrate access to the active site. This channel is localized in proximity of B'/C-loop and it was suggested to be the route for substrate passage to the active site [12]. On the other hand FR10 makes partially wall of the solvent channel [12], which is considered as a potential egress route for products.

Table 5 *The depiction of observed flexible regions in human CYPs. Observed flexible regions in respective MD simulations are shown in grey color. The dark grey color stands for observed flexible regions. The light grey color stands for flexible regions, which flexibility was lowered at HP with respect to NP. The white color is used at place, where no significant flexibility was observed.*

CYP		**FR1**	**FR2**	**FR3**	**FR4**	**FR5**	**FR6**	**FR7**	**FR8**	**FR9**	**FR10**
2A6	NP										
	HP										
1A2	NP										
	HP										
2C9	NP										
	HP										
2D6	NP										
	HP										
3A4	NP										
	HP										

Comparing the movements of secondary structure elements, the most flexible form of studied CYPs is CYP3A4, whereas CYP1A2 and CYP2A6 are the most rigid ones. The intermediate position between these two groups is occupied by CYP2D6 and CYP2C9. To summarize the findings concerning FR regions we can say that the flexibility is localized mainly to the solvent exposed loops and mostly in the CYP distal side, which is responsible for substrate binding. This may indicate that flexibility plays an important role in substrate channeling to the active site and product release from the active site.

4.2 Dynamics and hydration of the active site of mammalian Cytochrome P450

In this chapter, the 100+ ns long molecular dynamics simulations of Cytochrome P450 forms CYP2A6, 2B4, 2C9, 2D6, 2E1 and 3A4 are in detail described. The longer simulation time enables to observe larger conformational changes, e.g. domain motions, which cannot be observed at short time scales. Goals of this work were as follows; to identify the timescale on which selected structural features converge and to examine changes in the active site volume, solvation and the opening of access channels during the simulation.

Overall structural convergence

The overall convergence of all simulations is documented by main structural parameters listed in Table 6 (more detailed figures can be found in Ref. [8]). Most structures achieved RMSD equilibrium within 50 ns, but each timescale was a bit different. We observed that there is a group of CYPs, namely CYP2A6 and CYP2E1, which equilibrated "very quickly", within 20 ns. On the other hand, CYP2C9 did not reach the equilibrium till the end of the 100 ns long simulation, because a non-negligible trend in RMSD vs. time curves was apparent.

We can divide all structures into three groups according to behavior of radius of gyration (Table 6), which is a measure of the effective size of a protein. The first group contains CYP3A4, CYP2A6 and CYP2E1, which retain their closed structures. The two members (CYP2C9 and CYP2D6) of the second group display a limited degree of opening. The last group with the only member CYP2B4, which shows a large conformational change from an open to closed conformation during the first 50 ns of the MD simulation, which is in details described later.

Table *6* *The key data for all MD simulations (adopted from [8]).*

CYP	PDB	Rg [Å]	RMSD [Å]	V_{init} [Å^3]	V_{sim} [Å^3]	V_{ligand} [Å^3][c]	N_{WAT}
2A6	1Z10	22.44 ± 0.05	1.69 ± 0.08	330	480 ± 430	140 ± 65	3 ± 1.1
2B4[a]	1PO5	22.75 ± 0.07	4.13 ± 0.10	12600	3900 ± 2000		14 ± 1.9
2C9	1OG2	22.79 ± 0.05	2.68 ± 0.11	1960	480 ± 440	250 ± 30	2.5 ± 1.5
2D6	2F9Q	22.70 ± 0.06	2.44 ± 0.09	1800	1500 ± 1100	265 ± 40	4 ± 1.4
2E1	3E6I	22.55 ± 0.04	1.95 ± 0.05	290 (800)[b]	440 ± 310	95 ± 30 (226)[d]	1 ± 1.2
3A4	1TQN	22.66 ± 0.05	2.33 ± 0.05	1950	2000 ± 1300	550 ± 275	14 ± 2.2

V_{init} is the volume of the active site cavity above heme plane, as calculated with the CASTp server, from the crystal structure [109]. V_{sim} is the volume of the same cavity averaged from 10 snapshots acquired over a time frame of 10 ns. N_{WAT} – number of waters identified by visual inspection inside the active site; quoted value represents an average from 10 snapshots acquired over 10 ns. The radius of gyration (Rg) and RMSD were calculated from the last 5 ns of the 100ns trajectory. a – values for CYP2B4 were calculated from a longer 150 ns simulation. b – value in parentheses is the sum including the cavity present in the CYP2E1 crystal structure bound with lipid molecules. c – averaged over at least 8 most typical substrates for each human CYP (listed in Table ST2). d – van der Waals volume for lauric acid.

Specific findings

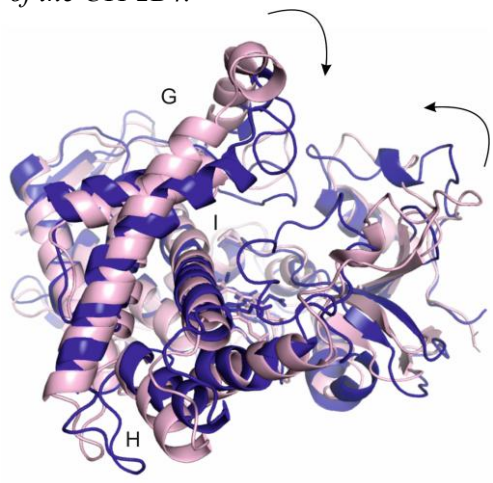
We have found also some specific findings concerning individual members of these human CYPs, which are discussed below.

CYP2A6 is known to metabolize small molecules, what seems to be reflected by its rather small active site. From MD simulation we can say that CYP2A6 is rather rigid with no remarkable structural changes. It also has smaller B-factors than other examined CYPs, with the exception of CYP2E1, which was similarly rigid (see Table 6). The paths to the CYP2A6 active site seem to be closed (the channel bottlenecks radii are smaller than 1.5 Å; measured by MOLE software [110]). Nevertheless, water molecules can enter and leave active site via alternative pathways, such as 2b channel.

We have analyzed by MD simulations also the **CYP2B4** open structure to see the structural changes on 100+ ns long time scale. The used crystal structure of the unliganded rabbit CYP2B4 is known to be an open homodimer [54]. This widely open structure was a subject of an intensive debate in CYP community because some researchers argued that this structure represented a crystallization artifact, while the other pointed out that it was a paradigm of CYP malleability. This open structure was closing during the course of MD simulation, where the movement and closing was very slow (Figure 9). The F/G-loop moved toward the N-terminal β-sheet, where this movement caused also the shift of the H/helix. The B'/C-loop also contributed to the

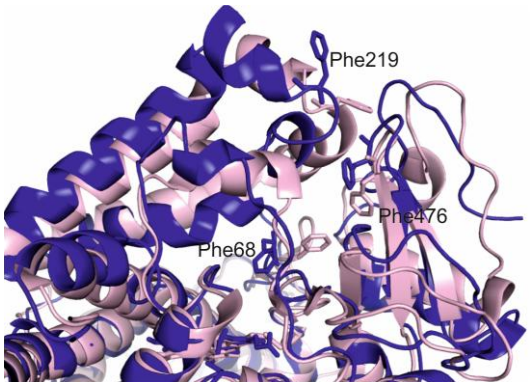
closing of the whole segment. The largest conformational changes took place during the first 30 ns, while some subsequent smaller conformational changes affecting A-helix resulting to slow side-chain rearrangements were observed on even longer time scale. Generally we can say that the open form of the CYP2B4 shows a tendency to close under ambient conditions. The open form of CYP2B4 and its movements were previously studied by Halpert et al. [55], who also described the closing movement of open CYP2B4 structure.

Figure 9 *The distal side of the CYP2B4 is depicted in cartoons. The X-ray structure is colored by the light pink color, the 100ns structure is colored by the blue color. The black arrows show the "closing" movement of the CYP2B4.*



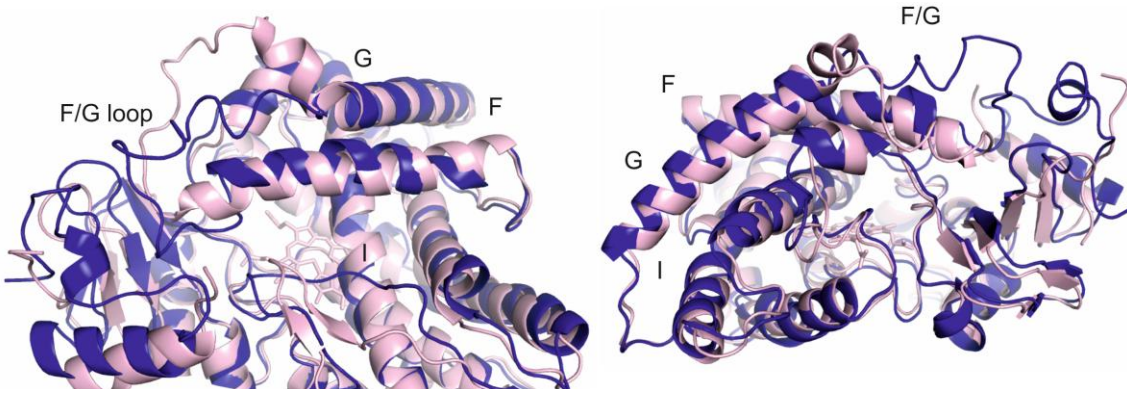
After the visual inspection, the F/G loop of **CYP2C9** is considered to be the most flexible part of this CYP. The shift of this region is reflected into a movement of the A-helix together with the conformational changes of some phenylalanines (F69, F219 and F476; Figure 10). After the conformational changes at A and B' helices between 30 and 50 ns the solvent channel has opened and the solvation of the active site has increased.

Figure 10 *Aligned structure of the 1OG2 (light pink) and 100ns structure (blue). Here the A segment shift is visible. Phe69, Phe219, Phe476 and HEME molecule are highlighted by sticks.*



CYP2D6 has solved structure with quite low resolution (3 Å), so some parts (mainly A'-helix) are missing in the structure because of lack of well resolved electron density. The CYP2D6 is also the structure of chimeric enzyme with seven mutations in the F/G loop (K206E, I215V, C216Y, S220P, P221A, I222L and I223L), which were introduced into primary sequence to enhance the crystallization. We have reversed these mutations to simulate behavior of wild type enzyme. The position of backbone missing residues (between P41 to Q52) were modeled according to CYP2C8 template (PDB: 1PQ2, [69]) and the side chains were built using the program Leap of the AMBER package. The core fold of CYP2D6 was stable during the simulation with the only exception regarding the movement of the F/G-loop. It is worth noting that F/G-loop has contacts with the A'- helix, which was built as mentioned above (Figure 11).

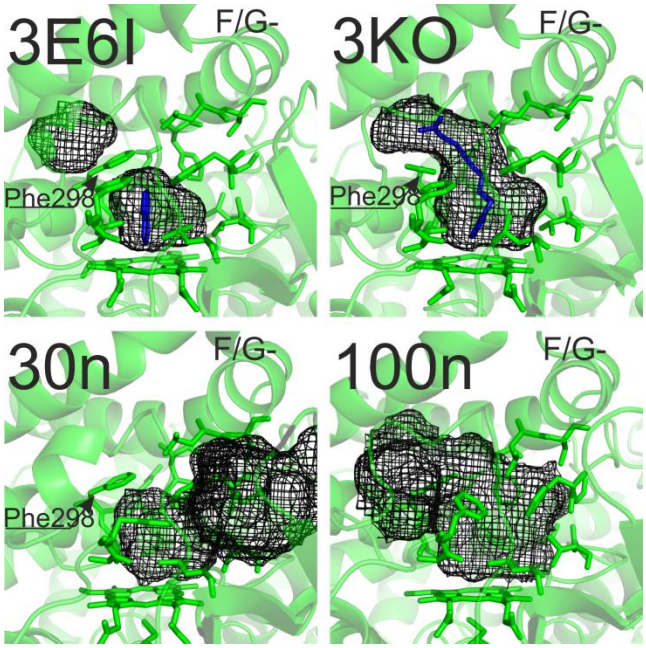
Figure 11 *The CYP2D6 structure. X-ray (PDB ID: 2F9Q) is depicted in light pink, 100ns snap is in blue color. The shift of the F/G segment is obvious. Left picture shows the front view and the right picture shows rotated view of CYP2D6.*



As previous CYP2A6, the **CYP2E1** showed a high rigidity. Inside the structure, we observe two additional cavities, which are near to the active site. The second cavity is located between I, F, G and B' helices and it was described earlier [52]. This cavity is not present in any other CYP structure. The third cavity is placed among β1 sheet and F' helix and B/C loop. Our results showed that the second cavity opens separately and joined the active site cavity. The third cavity was opening during the slide motion of the phenylalanine F478 together with rotation of F106 and F116 (Figure 12). Merging of these cavities with the active site void lead to significant increase in active site volume.

Figure 12 *The opening and closing of CYP2E1 cavities. The upper pictures depict observed cavities in crystal structures, the lower pictures show observed cavities during the simulation (adopted from [8]).*

.



The last CYP, **CYP3A4,** displayed higher flexibility in F'-helix, the F/G loop and the G/H-loop. Also some phenylalanines from the phenylalanine roof, which comprises F108, F213, F215, F241, F304 and F220, showed the sliding and rotational motion, namely F215 and F220. The residues from the phenylalanine roof are believed to form the gate for the substrate to the CYP3A4 active site [111].

Active sites of studied CYPs

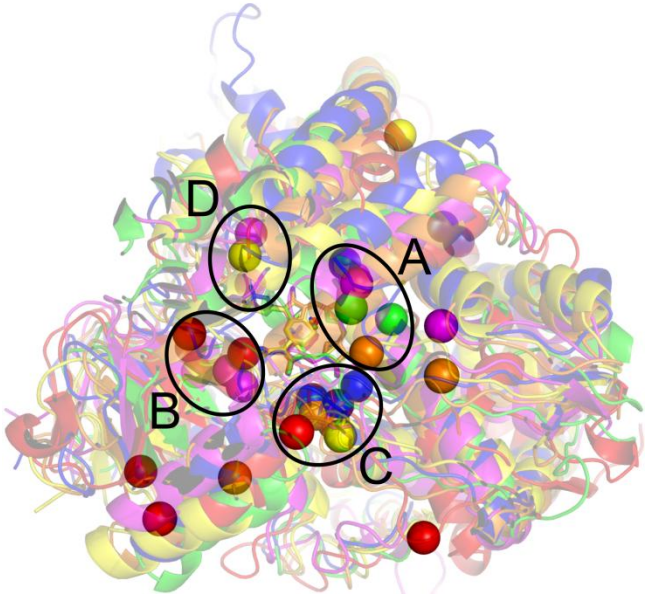
Generally the active site of considered CYP contains structurally conserved residues, where most of these residues are the part of the substrate recognition sites [112] (abbreviated as SRS; which are places typical for substrate interaction with the protein). We have analyzed the motions and changes in the active side and we can conclude that the biggest motions were caused by the main chain movements, which were caused by the secondary structure movements, rather than by the movement of the amino acids side chains lining the active site. The biggest active site movement was observed at the CYP2B4, the less flexible were CYP3A4, CYP2D6 and CYP2C9 and the smallest active site flexibility was at the CYP2A6.

We have also measured the volume of the active site and studied CYPs can be divided into three groups according to their active site volumes. The first group concerned CYPs

with the smallest active site volume (around 300 Å^3), namely CYP2A6 and CYP2E1. The second group contains CYP2C9, CYP2D6 and CYP3A4 with the active site around 1,900 Å^3. The largest active site was observed at CYP2B4 with the volume around 12,500 Å^3. All measured active site volumes are summarized in Table 6. The volume of the active size reflects also the size of the ligand, which is usually accommodated by the active site. For example, CYP2A6 and CYP2E1, which have the smallest active sites, are known to metabolize smaller substrates, whereas CYP3A4 with the larger active site is known to metabolize larger substrates.

Also the solvation of the active site was analyzed from MD simulations. The solvation of the active site was analyzed by radial distribution function (RDF) for the positions of water molecules relative to the heme's CO ligand and also by visual inspection. In studied CYP four common regions with high water solvation occurred: (A) the region between the heme and the I-helix, (B) the entrance to the water channel, (C) the inner cavity between the heme and the end of the K-helix and (D) the B'/C-loop region (see Figure 13). Nevertheless each CYP has unique water solvation regions and not all of mentioned regions are occupied in all systems together.

Figure 13 *Four mentioned places with the highest active site hydration are depicted. The water molecules are shown as spheres (adopted from [8]).*



Water molecules in CYP2A6 mostly preferred region D, where they interact with the residues from B'/C-loop. The second most occupied place was the region A, where one molecule binds in the 'kink' of the I-helix. Region B is also strongly solvated. The last

region is the region C with smaller occupancy, which is probably caused by the presence of some hydrophobic amino acids residues (F432, V365 and I366). Other strong-solvated site occurred; it is placed between I-helix and the beginning of the E-helix.

The most of water molecules in CYP2B4 were occupied in the region A and the second most occupied region is the region C. There were no water molecules in regions B and D, which can be caused by the motion of the B/C-loop during the closing of the open structure. However there is another strongly solvated region located at the end of E-helix and near R145 and S181.

CYP2C9 has the most of water molecules in the region A, namely in the water groove of the I-helix, but some of them occurred also between I- helix and E-helix. The second most occupied region is the region B followed by region C. Region D contained no waters.

The active site of CYP2D6 contained fewer water molecules than in other systems. The kink of the I-helix did not contain any water, but some water molecules were localized in regions B and C.

As in previous case, there is no water molecule in the kink of the I-helix in the CYP2E1. Generally, the active site was not hydrated for more than half of the MD simulation. The most occupied region was the region C and the solvent channel near the E, F and I-helices. Regions B and D were occupied sporadically by water molecules.

The last cytochrome CYP3A4 had water molecules in all four regions, but the occupancy was different in each region. The most occupied region was region B. Waters in region A and C seemed to be more mobile than in other cytochromes. The water molecules also occupied the place above the heme, near the carbon monoxide molecule, where no other water molecules in other cytochromes showed this behavior.

As described above, the active site of various cytochromes showed different solvation pattern, although some places are more occupied by water molecule and seemed to be preferred. The least solvated active site has CYP2D6 and especially CYP2E1, which are known for metabolizing non-polar substrates.

Opening channels

The residence time of water molecules in the active site is affected by the opening and closing of access channels into the active site. Water molecules usually leaked in and out the active site via these channels. The access path bottlenecks decreased over the

course of simulation for all cytochromes. The structures of studies CYPs contain several different access/egress paths:

CYP2A6 – No channel was fully opened during the simulation, but 2b and 2e channels exhibited widest openings.

CYP2B4 – The widest channel was 2a channel, but during the closing of the structure 2e channel opened. Occasionally 2ac opened.

CYP2C9 – The 2b and solvent channel were opened, also 2a and 2e were partially opened.

CYP2D6 – The only opened channel was the solvent channel.

CYP2E1 – We have not observed any fully opened channel. The most widely-opened channels were 2ac and 2e.

CYP3A4 – Two channels were opened, 2e and solvent channel. 2b and 2a channels were also partially opened.

Generally the widest channels were usually 2b, 2e and solvent channel. Nevertheless, the water is able to penetrate into the active site by gradual adjustment of the channel walls.

4.3 Binding of quinidine into the Cytochrome P450 2D6 active site

In our previous work we have analyzed flexibility of selected human Cytochromes P450 and changes in their structure and active site properties at normal and high pressures. Here we were interested to what extent the ligand binding can affect the flexibility of CYP2D6. We have chosen CYP2D6 and its inhibitor quinidine as a ligand. The role of quinidine molecule and its inhibition to CYP2D6 has been extensively studied [49, 112] and also modelling studies of CYP2D6 and CYP2D6-drug interaction were described [113]. Quinidine is a marketed drug used as an antiarrhytmic agent and it is known that it strongly binds into CYP2D6 without being metabolized. It has been established as a competitive inhibitor due to its features typical to CYP2D6 substrates (having planar aromatic ring and polar nitrogen).

Here we have analyzed changes in structure, flexibility and compressibility caused by the quinidine binding by means of high-pressure UV/VIS spectroscopy and MD.

Common molecular dynamic features

We have run molecular dynamics simulations at normal (0.1 MPa) and high pressure (300 MPa) with and without quinidine present in the CYP2D6 active site. We again focused on some measures of structural features related to enzyme flexibility. We have observed differences in Rg values under normal and high pressures simulations between ligand free and ligand bound forms. The decrease in the Rg under HP is significantly higher in ligand free enzymes whereas CYP2D6 with quinidine bound did not show significant change under HP in Rg values (Table 7). This indicates that CYP2D6 without quinidine ligand is more compressible while the same enzyme with the quinidine ligand bound is more rigid. This fact corresponds to measured UV/VIS-pressure perturbation spectroscopy data.

Table 7 *The key data for all MD simulations.*

			NP				HP			
CYP	initial PDB	ligand	*t* (ns)	R_g[b] (Å)	RMSD[c] (Å)	B_f[d]	*t*(ns)	R_g[b] (Å)	RMSD[c] (Å)	B_f[d]
2D6	1OG2	-	100	22.70±0.06	2.46±0.08	11.37	100	22.24±0.04	1.51±0.04	11.06
2D6	prof. Gordon	quinidine	50	23.07±0.05	3.32±0.08	13.36	50	23.05±0.05	2.52±0.07	11.49

b - last 3 ns

c - last 3ns

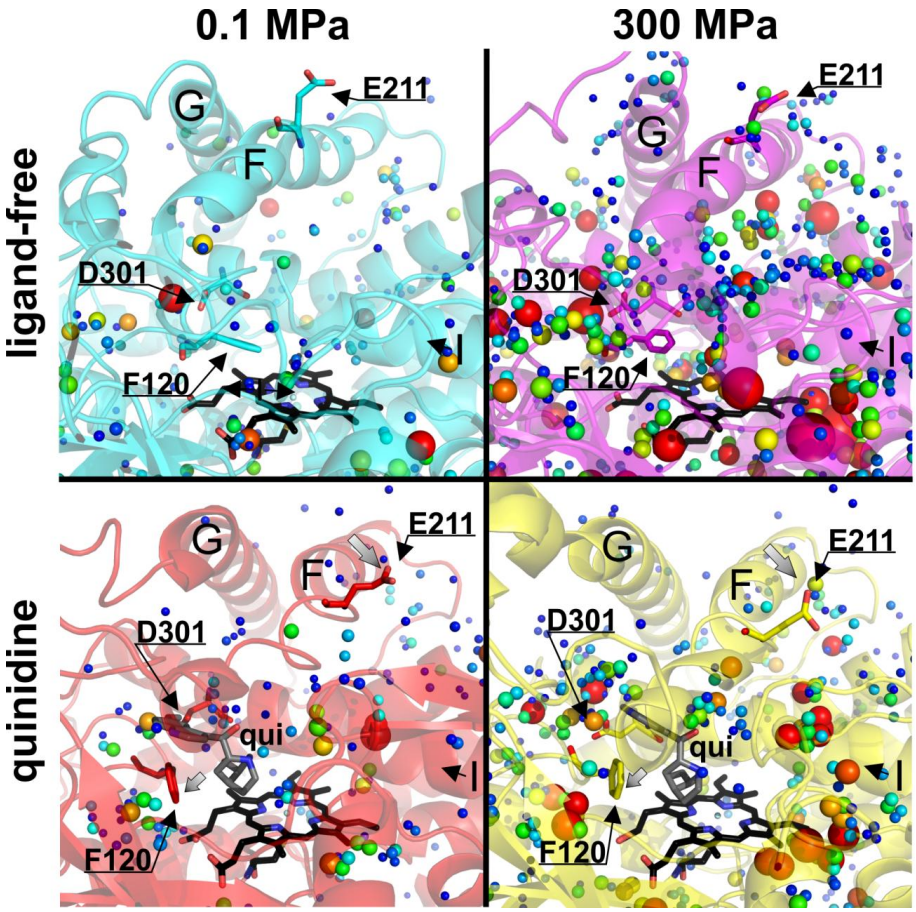
d - last 1ns

The differences in flexibility of CYP2D6 with or without ligand bound in the active site are inferred also from temperature B-factors. At both structures with or without quinidine, the flexibility of enzyme is lowered by HP. The detailed analyses showed that the core of the protein is more rigid in the structures with quinidine bound, where the B-factors are considerably lower after ligand binding. Also under high pressure the behavior pattern is the same, the active site of CYP2D6 without quinidine is more flexible than CYP2D6 with the quinidine bound. Generally we can conclude that the binding of the quinidine molecule in the active site lowered the active site flexibility.

Active site properties

We have also analyzed the active sites, because our previous work had shown that the active site solvation is pressure depended (described earlier). Generally more water molecules remained in the active site close to the heme molecule in systems without quinidine. In all HP systems we have observed water molecules. The place near D301 in the B/C loop is highly occupies by water molecule in ligand-free systems and after quinidine binding this place is occupied by the ligand, so no more water molecule can take place there (see Figure 14). It seems that the presence of the quinidine in the active site causes the blocking of the entry for water molecules into the actives site. This could contribute to high denaturation resistance of ligand bound system, which is also documented from high pressure experiments.

Figure 14 *Positions occupied by water molecules inside the CYP2D6 structures. Water molecules are shown as spheres, where the different size of spheres means different remaining time of particular water on its position during the simulation. Those molecules, which remained on its position 1/3 of the simulation are depicted as larger red spheres, molecules which were remaining at their positions at least 1/10 of the simulation are depicted as smaller blue spheres. Black arrows show positions of important residues (adopted from [9]).*

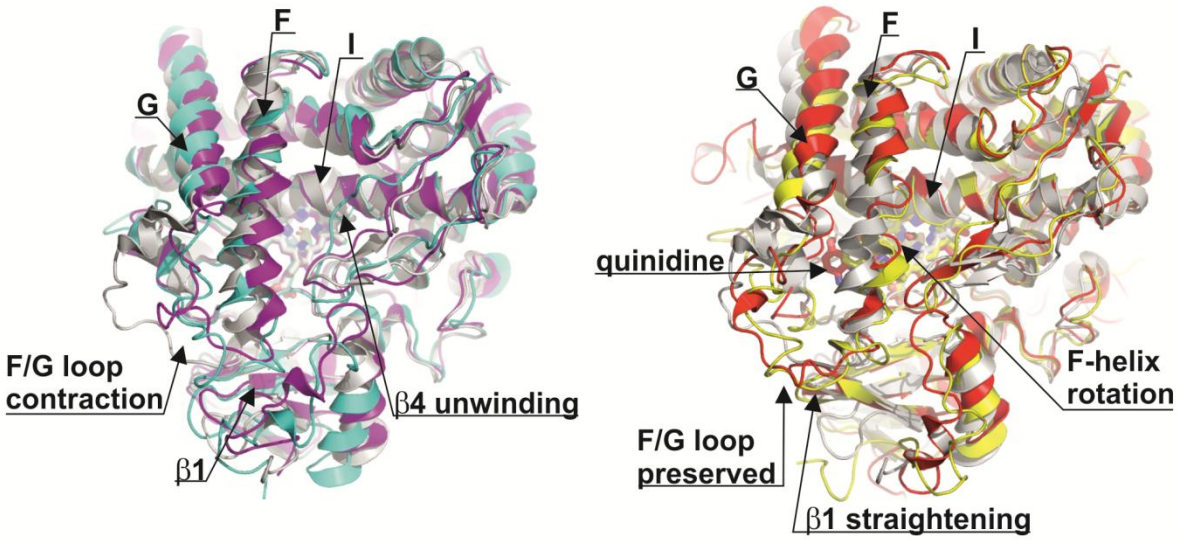


Structural changes

We have observed the structural changes at the CYP2D6 upon quinidine binding. Overall CYP fold remained the same with some small changes mainly in the distal part of the enzyme. The F/G loop of unliganded CYP2D6 at NP showed the contraction of this segment accompanied by unwinding of β4 sheet, whereas the rest of the structure remained similar (see Figure 15). Nevertheless the presence of quinidine conserves both these segments unchanged. Also the disruption of the F helix caused by the changed position E211 led to the stabilizing the expanded F/G loop. The contraction of the F/G loop and unwinding of β4 sheet also led to the contraction of the active site cavity under

pressure. The molecule of quinidine itself is placed in the vicinity of key residues (F120, D301) and its position is stable during all MD.

Figure 15 *Structural changes observed in CYP2D6 structures. The crystal structure is depicted in grey; at the left picture there are structures at NP and HP depicted in cyan and violet respectively. On the right side the structure at NP is depicted in red and at HP in yellow (adopted from [9]).*



5. SUMMARY

The aim of this work was to study the flexibility and dynamics of most important human drug metabolizing Cytochromes P450. Molecular dynamics simulations were used as the main technique but all observations were cross-checked against available experimental data. The data acquired from molecular dynamics simulations and presented in this work showed together with the experimental data various aspects of flexibility of human microsomal CYPs. In addition, the stability of considered enzymes was described, the active site properties and its behavior was documented and the solvation of active site was analyzed.

The high pressure simulations allowed us the deeper insight into the problem of CYP pressure compressibility and it provided detailed structural information about the whole enzyme as well as its active site cavity [103, 105, 108]. The active-site-related attributes have strong relevance to the enzyme function, so there are differences between various CYPs in respect to their physiological and pharmaceutical relevant qualities [19].

In our work, CYP3A4 was identified as the most flexible CYP form, the intermediate position is occupied by CYP2D6 and CYP2C9 and the most rigid ones are CYP1A2 and CYP2A6. These observations agree with previous CYP's knowledge about their flexibilities [1]. The places with the higher flexibility (FR) were localized to the solvent exposed regions. These FR regions are in coincidence with previously described flexible regions by Skopalík et al. [18] and with plastic regions defined by Zhao et al. [113].

Generally the high pressure decreases the radius of gyration and decrease temperature B-factors that means that the flexibility of CYP is somewhat restricted under high-pressure conditions. This is in line with earlier observations by McCarthy and Grigera [114]. We have observed these cut-downs in majority of flexible regions, but some exceptions occurred (namely FR3, FR5).

The high pressure studies also yield the information about active site solvation. Comparing NP and HP simulations we can get important information about changes in hydration of the active site under HP. At the high pressure we observed the increasing number of water molecules in a close proximity of heme in all studied systems except CYP3A4. The hydration of the nonpolar protein cavity under pressure has also been observed experimentally [115] and described also in independent MD simulations [116]. The specific finding at CYP3A4 can be explained by different active site

topology, where there are not any steric restrictions of the active site above the heme iron atom. So the increasing pressure may push out the water molecule from the active site cavity.

We have also found the correlation between $-\alpha$ value and CN^W at 4 and 5 Å under NP and HP, which indicates that CYPs containing more water molecule in the active site close to the heme molecule exhibit larger $-\alpha$ value and can be considered as more compressible. This finding is also supported by Jung et al. [105].

The long time scale (100+ ns) simulations showed that on average the CYPs' structures are equilibrated on 50+ ns, where more rigid structures (CYP2A6 and CYP2E1) relaxed on shorter timescales (<30 ns) opposite to flexible ones that needed longer times. The widely opened structure of CYP2B4 closes slowly during the MD simulation and it was approaching the classical CYP's fold.

We have measured the volume of the active sites and it fluctuated by more than 50% during the MD simulation. The biggest fluctuations of active site were observed in simulations of CYP2E1, where opening and closing of two cavities adjacent to the active site occurred. We believe that this finding could explain why CYP2E1 is able to metabolize small non-polar substrates as well as longer fatty acids. CYP2E1 had also the least solvated active site, which is in coincidence with its preferences for non-polar substrates. The most solvated actives sites were observed at CYP2B4 and CYP3A4.

The active sites of CYP2A6 and CYP2E1 are deeply buried in the protein, with the access paths that are not big enough for passing of water. In spite of this fact, water molecules are still able to penetrate the active site by local adaptations of the channel width to accommodate the passage of water molecules. This may imply that various substrates does not need a fully open channel to get to the active site of CYPs, but the local adaptations and changes of a channel can enable the substrate passage.

In the last part of this work the changes in compressibility of CYP2D6 upon quinidine binding are shown. The ligand-free structures are more compressible and denaturation prone under high pressure, which was documented in our previous work [7]. The structure with the quinidine molecule in the active site is rigid and stable to denaturation. The active site flexibility (determined by B_f values) decrease upon binding of quinidine. The same behavior was described earlier at other proteins [117]. The active site solvation also decreases upon binding of quinidine.

The presented study shows the predictive power of molecular dynamics simulations to explain the behavior of mammalian cytochromes P450 and to complement available

experimental data. It is worth to say that the computational results are not self-explanatory and it is better to join the simulation results together with the experiment and thus strengthen obtained and explained results.

6. SHRNUTÍ

Předložená disertační práce se zabývá popisem flexibility a dynamiky savčích forem cytochromů P450. V první části práce jsou u vybraných systémů sledovány základní veličiny popisující vlastnosti cytochromů sledované během molekulové dynamiky, např. B-faktory, RMSD, gyrační poloměr, flexibilní regiony, dále pak chování aktivního místa a jeho solvataci a další. Tyto údaje jsou korelovány s experimentálními výsledky (zejména z UV/VIS spektroskopie) a jsou popsány matematickou rovnicí. Je zde sledován vliv vysokého tlaku (300 MPa) na chování jednotlivých systémů.

Použití vysokotlakých simulací nám umožňuje lepší pohled na flexibilitu enzymů, jejich kompresibilitu a chování aktivního místa. Celkově lze říci, že flexibilita jednotlivých systémů je za použití vysokého tlaku značně omezena a dochází k poklesu kompresibility. Jednotlivé použité cytochromy mohou být rozděleny do tří skupin podle toho, jaké flexibilní chování vykazují: nejvíce flexibilní je CYP3A4, který je následován skupinou cytochromů se střední flexibilitou – CYP2D6 a CYP2C9 a nejvíce rigidní jsou CYP2A6 a CYP1A2. V jednotlivých systémech byly popsány flexibilní regiony společné pro všechny systémy (tzv. FR) a bylo pozorováno, že tyto oblasti jsou ve většině případů vystaveny směrem do obklopujícího solventu. Celkově ale pohyby těchto flexibilních regionů nijak zvlášť neovlivňují celkovou strukturu typickou pro cytochromy P450.

V druhé části práce jsou studovány dlouhé molekulově-dynamické simulace (100+ ns) a opět je posuzován vliv vysokého tlaku na jednotlivé systémy. Základní veličiny popisující flexibilitu systému, jako jsou teplotní B-faktor a gyrační poloměr, jsou nižší za vysokého tlaku oproti naměřeným hodnotám za normálního tlaku, což vysvětluje omezenou flexibilitu jednotlivých systému.

Druhým cílem předkládaného článku je specifikovat konvergenční kriteria a čas, který jednotlivé systémy potřebují pro svou stabilizaci. Většina systémů zkonvergovala na časové škále 50+ns a systémy, které všeobecně vykazují větší rigiditu, měly tento čas ještě kratší (<30ns). Jedinou výjimku tvoří systém CYP2B4, který po celou dobu simulace vykazoval vysokou flexibilitu, která byla způsobena zavíráním otevřené struktury a snahou cytochromu dosáhnout klasického uspořádání typického pro cytochromy P450.

Objem aktivního místa studovaných cytochromů fluktuuje 50% celkového času simulace. Největší rozdíly v objemu aktivního místa jsou sledovány u systému CYP2E1. Tyto rozdíly jsou tvořeny otevíráním a zavíráním dvou dalších přídatných kavit, které jsou umístěny nedaleko aktivního místa. Tato pozorování mohou také vysvětlit, proč je cytochrom CYP2E1 schopen metabolizovat ve svém aktivním místě jak malé nepolární substráty, tak i objemnější molekuly mastných kyselin. Současně má CYP2E1 nejméně solvatované aktivní místo ze všech pozorovaných enzymů, což je ve shodě s dříve publikovaným faktem, že jeho aktivní místo je silně nepolární. Nejvíce molekul vod bylo nalezeno v aktivním místě CYP2B4 a CYP3A4.

Ačkoliv aktivní místa CYP2A6 a CYP2E1 jsou uložena hluboko ve struktuře proteinu a jejich přístupové kanály se zdají být nevhodné pro vstup vod do aktivního místa, i přesto jsou jednotlivé molekuly schopné se do aktivního místa dostat. Tento jev je způsobený postupným přizpůsobováním tvaru jednotlivých kanálů molekulám vody, když těmito kanály prostupují. Celkově je tento model vhodný i pro vysvětlení vstupu substrátu do aktivního místa i přes to, že jednotlivé přístupové kanály nemusejí mít patřičný průměr vhodný pro vstupující molekulu.

Poslední část práce je věnovaná studiu CYP2D6 a navázáním jednoho z jeho substrátu – quinidinu do aktivního místa. V práci je popsán vliv quinidinu na celkovou flexibilitu systému, která je tímto navázáním snížena, tak i vliv na chování aktivního místa. Vysokotlaké simulace ukázaly, že u systému obsahujícím ligand je snížena i kompresibilita systému a cytochrom CYP2D6 s quinidinem je méně náchylný k denaturaci.

Závěrem bych chtěla zmínit, že předložená práce ukazuje sílu molekulové dynamiky jako metody, která slouží k popisu molekul na atomární úrovni. Nesmíme ale zapomenout, že i tato metoda má svá omezení a proto je vodné ji doplnit i experimentálními výsledky a tak je lépe interpretovat a využít.

7. LIST OF PH.D. CANDIDATE`S PUBLICATIONS

Hendrychova T., Anzenbacherova E., Hudecek J., Skopalik J., Lange R., Hildebrandt P., Otyepka M., Anzenbacher P.: *Flexibility of human cytochrome P450 enzymes: molecular dynamics and spectroscopy reveal important function-related variations.* Biochim. Biophys. Acta, 2011, **1814**(1): p. 58-68.
Attached as Appendix I

Berka K., **Hendrychova T**., Anzenbacher P., Otyepka M.: *Membrane position of ibuprofen agrees with suggested access path entrance to cytochrome P450 2C9 active site.* J. Phys. Chem. A, 2011, **115**(41): p. 11248-55.
Attached as Appendix II

Berka, K., **Hendrychova T**., Anzenbacherova E., Lange R., Masek V., Anzenbacher P., Otyepka M.: *Binding of quinidine radically increases the stability and decreases the flexibility of the cytochrome P450 2D6 active site.* J. Inorg. Biochem., 2012. **110C**: p. 46-50.
Attached as Appendix III

Hendrychova T., Berka K., Navratilova V., Anzenbacher P., Otyepka M.: *Dynamics and hydration of the active sites of mammalian cytochromes P450 probed by molecular dynamics simulations.* Current Drug Metab., 2012, **13**(2): p. 177-189.
Attached as Appendix IV

Child E.S., **Hendrychova T**., McCague K., Futreal A., Otyepka M., Mann DJ: *A cancer-derived mutation in the PSTAIRE helix of cyclin-dependent kinase 2 alters the stability of cyclin binding.* Biochim. Biophys. Acta, 2010, **1803**(7): p. 858-64
Attached as Appendix V

8. LIST OF ABBREVIATIONS

Bf	temperature B-factor
CN^B	the coordination number for backbone atoms
CN^W	the coordination number for water molecules
CYP	Cytochrome P450
Cys	cysteine
EC	enzyme commission number
ff	force field
FR	flexible region
GB	Generalized Born
HP	high pressure
MD	molecular dynamics
NMR	nuclear magnetic resonance
NP	normal pressure
NpT	the ensemble with constant number of particles, pressure and temperature
PB	Poisson-Boltzman
PDBid	the identification code for protein database
QUI	quinidine
RDF	radial distribution function
Rg	radius of gyration
RMSD	Root-mean-square deviation
t	time
UV/VIS	Ultra Violet/Visible Spectroscopy
X-ray	Roentgen rays

9. REFERENCES

1. Montellano, P.R.O.d., *Cytochrome P450: structure, mechanism, and biochemicsty*. 3rd ed. 2005, New York: Kluwer Academic/Plenum Publisher.
2. Evans, W.E. and M.V. Relling, *Pharmacogenomics: translating functional genomics into rational therapeutics.* Science, 1999. **286**(5439): p. 487-91.
3. Anzenbacher, P. and E. Anzenbacherova, *Cytochromes P450 and metabolism of xenobiotics.* Cellular and molecular life sciences, 2001. **58**(5-6): p. 737-47.
4. Guengerich, F.P., *Cytochrome P450 and chemical toxicology.* Chemical Research in Toxicology, 2008. **21**(1): p. 70-83.
5. Rendic, S., *Summary of information on human CYP enzymes: Human P450 metabolism data.* Drug Metabolism Reviews, 2002. **34**(1-2): p. 83-448.
6. Bernhardt, R., *Cytochromes P450 as versatile biocatalysts.* Joural of Biotechnology, 2006. **124**(1): p. 128-45.
7. Hendrychova, T., et al., *Flexibility of human cytochrome P450 enzymes: molecular dynamics and spectroscopy reveal important function-related variations.* Biochimica et biophysica acta, 2011. **1814**(1): p. 58-68.
8. Hendrychova, T., et al., *Dynamics and hydration of the active sites of mammalian cytochromes P450 probed by molecular dynamics simulations.* Current Drug Metabolism, 2012. **13**(2): p. 177-189.
9. Berka, K., et al., *Binding of quinidine radically increases the stability and decreases the flexibility of the cytochrome P450 2D6 active site.* Jouranl of Inorganic Biochemistry, 2012. **110C**: p. 46-50.
10. Alberts B., B.D., Johnson A., Lewis J., Raff M., Roberts K., Walter P., *Essential Cell Biology: An Introduction to the Molecular Biology of the Cell*. 2nd ed. 1998, New York: Garland Publishing, Inc. 630.
11. Nelson, D.R., et al., *The P450 superfamily: update on new sequences, gene mapping, accession numbers, early trivial names of enzymes, and nomenclature.* DNA Cell Biol, 1993. **12**(1): p. 1-51.
12. Cojocaru, V., P.J. Winn, and R.C. Wade, *The ins and outs of cytochrome P450s.* Biochimica Et Biophysica Acta, 2007. **1770**(3): p. 390-401.
13. Mestres, J., *Structure conservation in cytochromes P450.* Proteins, 2005. **58**(3): p. 596-609.
14. Lisitsa, A., et al., *Bioinformatic insight into the unity and diversity of cytochromes P450.* Methods and Findings in Experimental Clinical Pharmacology, 2003. **25**(9): p. 733-45.
15. Otyepka, M., et al., *What common structural features and variations of mammalian P450s are known to date?* Biochimica Et Biophysica Acta, 2007. **1770**(3): p. 376-389.
16. Hendrychova, T., et al., *Flexibility of human cytochrome P450 enzymes: molecular dynamics and spectroscopy reveal important function-related variations.* Biochimica Et Biophysica Acta, 2011. **1814**(1): p. 58-68.
17. Winn, P.J., V. Cojocaru, and R. Wade, *The ins and outs of the cytochromes P450: understanding the mechanics of a key drug metabolising enzyme family.* Febs Journal, 2007. **274**: p. 261-261.
18. Skopalik, J., P. Anzenbacher, and M. Otyepka, *Flexibility of human cytochromes P450: molecular dynamics reveals differences between CYPs 3A4, 2C9, and 2A6, which correlate with their substrate preferences.* The journal of physical chemistry B, 2008. **112**(27): p. 8165-73.
19. Tokuriki, N. and D.S. Tawfik, *Protein Dynamism and Evolvability.* Science, 2009. **324**(5924): p. 203-207.

20. Ekroos, M. and T. Sjogren, *Structural basis for ligand promiscuity in cytochrome P450 3A4.* Proceedings of National Academy of Science U S A, 2006. **103**(37): p. 13682-7.
21. Zhao, Y., et al., *Structure of microsomal cytochrome P450 2B4 complexed with the antifungal drug bifonazole: insight into P450 conformational plasticity and membrane interaction.* Journal of biological chemistry, 2006. **281**(9): p. 5973-81.
22. Groves, J.T., *The bioinorganic chemistry of iron in oxygenases and supramolecular assemblies.* Proceedings of the National Academy of Science U S A, 2003. **100**(7): p. 3569-74.
23. Rydberg, P., et al., *Dynamics of water molecules in the active-site cavity of human cytochromes P450.* Journal of Physical Chemistry B, 2007. **111**(19): p. 5445-5457.
24. Skopalik, J., P. Anzenbacher, and M. Otyepka, *Flexibility of human cytochromes P450: molecular dynamics reveals differences between CYPs 3A4, 2C9, and 2A6, which correlate with their substrate preferences.* Jouranl of Physical Chemistry B, 2008. **112**(27): p. 8165-73.
25. Wade, R.C., et al., *A survey of active site access channels in cytochromes P450.* Journal of inorganic biochemistry, 2004. **98**(7): p. 1175-1182.
26. Winn, P.J., et al., *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.* Proceedings of the National Academy of Sciences of the United States of America, 2002. **99**(8): p. 5361-5366.
27. Schleinkofer, K., et al., *Do mammalian cytochrome P450s show multiple ligand access pathways and ligand channelling?* Embo Reports, 2005. **6**(6): p. 584-589.
28. Ludemann, S.K., V. Lounnas, and R.C. Wade, *How do substrates enter and products exit the buried active site of cytochrome P450cam? 2. Steered molecular dynamics and adiabatic mapping of substrate pathways.* Journal of molecular biology, 2000. **303**(5): p. 813-30.
29. Li, W.H., et al., *Reduced Catalytic Activity of P450 2A6 Mutants with Coumarin: A Computational Investigation.* Journal of Chemical Theory and Computation, 2009. **5**(5): p. 1411-1420.
30. Ludemann, S.K., V. Lounnas, and R.C. Wade, *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.* Journal of molecular biology, 2000. **303**(5): p. 797-811.
31. Ludemann, S.K., V. Lounnas, and R.C. Wade, *How do substrates enter and products exit the buried active site of cytochrome P450cam? 2. Steered molecular dynamics and adiabatic mapping of substrate pathways.* Jouranl of Molecular Biology, 2000. **303**(5): p. 813-30.
32. Anzenbacher, P. and M. Otyepka, *Editorial: cytochromes P450: flexibility and plasticity - properties determining substrate preferences.* Curr Drug Metab, 2012. **13**(2): p. 129.
33. Nelson, D., *The Cytochrome P450 Homepage.* Human Genomics 4, 2009: p. 59-65.
34. Sansen, S., et al., *Adaptations for the oxidation of polycyclic aromatic hydrocarbons exhibited by the structure of human P450 1A2.* Journal of Biological Chemistry, 2007. **282**(19): p. 14348-14355.
35. Tassaneeyakul, W., et al., *Specificity of substrate and inhibitor probes for human cytochromes P450 1A1 and 1A2.* Jouranl of Pharmacol and Experimental Therapeutics, 1993. **265**(1): p. 401-7.
36. Butler, M.A., et al., *Human cytochrome P-450PA (P-450IA2), the phenacetin O-deethylase, is primarily responsible for the hepatic 3-demethylation of caffeine and N-oxidation of carcinogenic arylamines.* Proceeding of the National Academy of Science U S A, 1989. **86**(20): p. 7696-700.
37. Kalow, W. and B.K. Tang, *The use of caffeine for enzyme assays: a critical appraisal.* Clinical Pharmacology and Therapeutics, 1993. **53**(5): p. 503-14.
38. Racha, J.K., A.E. Rettie, and K.L. Kunze, *Mechanism-based inactivation of human cytochrome P450 1A2 by furafylline: detection of a 1:1 adduct to protein and evidence for the formation of a novel imidazomethide intermediate.* Biochemistry, 1998. **37**(20): p. 7407-19.

39. Lang, N.P., et al., *Rapid metabolic phenotypes for acetyltransferase and cytochrome P4501A2 and putative exposure to food-borne heterocyclic amines increase the risk for colorectal cancer or polyps.* Cancer Epidemiol Biomarkers Prevention, 1994. **3**(8): p. 675-82.
40. Raunio, H., et al., *Cytochrome P4502A6 (CYP2A6) expression in human hepatocellular carcinoma.* Hepatology, 1998. **27**(2): p. 427-32.
41. Pianezza, M.L., E.M. Sellers, and R.F. Tyndale, *Nicotine metabolism defect reduces smoking.* Nature, 1998. **393**(6687): p. 750.
42. Seden, K., et al., *Grapefruit-drug interactions.* Drugs, 2010. **70**(18): p. 2373-407.
43. Yasumori, T., et al., *Expression of a human P-450IIC gene in yeast cells using galactose-inducible expression system.* Molecular Pharmacology, 1989. **35**(4): p. 443-9.
44. Rendic, S. and F.J. Di Carlo, *Human cytochrome P450 enzymes: a status report summarizing their reactions, substrates, inducers, and inhibitors.* Drug Metababolism Review, 1997. **29**(1-2): p. 413-580.
45. Wester, M.R., et al., *The structure of human cytochrome P450 2C9 complexed with flurbiprofen at 2.0-A resolution.* Journal of Biological Chemistry, 2004. **279**(34): p. 35630-7.
46. Fleming, I., *Cytochrome P450 epoxygenases as EDHF synthase(s).* Pharmacological Research, 2004. **49**(6): p. 525-33.
47. Williams, P.A., et al., *Crystal structure of human cytochrome P450 2C9 with bound warfarin.* Nature, 2003. **424**(6947): p. 464-8.
48. Smith, D.A., et al., *Human cytochrome P450s: selectivity and measurement in vivo.* Xenobiotica, 1998. **28**(12): p. 1095-128.
49. McLaughlin, L.A., et al., *Why is quinidine an inhibitor of cytochrome P450 2D6? The role of key active-site residues in quinidine binding.* Journal of biological chemistry, 2005. **280**(46): p. 38617-24.
50. Saraceno, M., A. Coi, and A.M. Bianucci, *Molecular modelling of human CYP2D6 and molecular docking of a series of ajmalicine- and quinidine-like inhibitors.* Interantional Journal of Biological Macromolecules, 2008. **42**(4): p. 362-71.
51. Vieira, I., M. Sonnier, and T. Cresteil, *Developmental expression of CYP2E1 in the human liver. Hypermethylation control of gene expression during the neonatal period.* European Journal of Biochememistry, 1996. **238**(2): p. 476-83.
52. Porubsky, P.R., K.M. Meneely, and E.E. Scott, *Structures of human cytochrome P-450 2E1. Insights into the binding of inhibitors and both small molecular weight and fatty acid substrates.* Jouranl of Biological Chemistry, 2008. **283**(48): p. 33698-707.
53. Coon, M.J., et al., *Biochemical studies on cytochrome P-450 solubilized from liver microsomes: partial purification and mechanism of catalysis.* Annals of the New York Academy of Science, 1973. **212**: p. 449-57.
54. Scott, E.E., et al., *An open conformation of mammalian cytochrome P450 2B4 at 1.6-A resolution.* Proceedings of National Academy of Science U S A, 2003. **100**(23): p. 13196-201.
55. Wilderman, P.R. and J.R. Halpert, *Plasticity of CYP2B enzymes: structural and solution biophysical methods.* Current Drug Metabolism, 2012. **13**(2): p. 167-76.
56. Wang, A., et al., *Structural Characterization of the Complex between alpha-Naphthoflavone and Human Cytochrome P450 1B1.* Journal of Biological Chemistry, 2011. **286**(7): p. 5736-5743.
57. Porubsky, P.R., K.P. Battaile, and E.E. Scott, *Human cytochrome P450 2E1 structures with fatty acid analogs reveal a previously unobserved binding mode.* Journal of biological chemistry, 2010. **285**(29): p. 22282-90.
58. Sansen, S., et al., *Structural insight into the altered substrate specificity of human cytochrome P450 2A6 mutants.* Archives of biochemistry and biophysics, 2007. **464**(2): p. 197-206.
59. DeVore, N.M., et al., *Key Residues Controlling Phenacetin Metabolism by Human Cytochrome P450 2A Enzymes.* Drug Metabolism and Disposition: The Biological Fate of Chemicals, 2008. **36**(12): p. 2582-2590.

60. Strushkevich, N., et al., *Structural analysis of CYP2R1 in complex with vitamin D-3.* Journal of Molecular Biology, 2008. **380**(1): p. 95-106.
61. Yano, J.K., et al., *Synthetic inhibitors of cytochrome P-450 2A6: Inhibitory activity, difference spectra, mechanism of inhibition, and protein cocrystallization.* Journal of Medicinal Chemistry, 2006. **49**(24): p. 6987-7001.
62. Strushkevich, N.V., et al., *Crystal structure of CYP2R1 in complex with 1-alpha-hydroxy-vitamin D2.* To be published.
63. Yano, J.K., et al., *Structures of human microsomal cytochrome P450 2A6 complexed with coumarin and methoxsalen.* Nature structural & molecular biology, 2005. **12**(9): p. 822-823.
64. Strushkevich, N.V., et al., *Crystal structure of CYP2R1 in complex with vitamin D2.* To be published.
65. Yano, J.K., et al., *The structure of human microsomal cytochrome P450 3A4 determined by X-ray crystallography to 2.05-angstrom resolution.* Journal of Biological Chemistry, 2004. **279**(37): p. 38091-38094.
66. Williams, P.A., et al., *Crystal structures of human cytochrome P450 3A4 bound to metyrapone and progesterone.* Science, 2004. **305**(5684): p. 683-6.
67. Smith, B.D., et al., *Structure of the human lung cytochrome P450 2A13.* Journal of Biological Chemistry, 2007. **282**(23): p. 17306-17313.
68. Gay, S.C., et al., *Crystal structure of a cytochrome P450 2B6 genetic variant in complex with the inhibitor 4-(4-chlorophenyl)imidazole at 2.0-A resolution.* Molecular pharmacology, 2010. **77**(4): p. 529-38.
69. Schoch, G.A., et al., *Structure of human microsomal cytochrome P4502C8 - Evidence for a peripheral fatty acid binding site.* Journal of Biological Chemistry, 2004. **279**(10): p. 9497-9503.
70. Schoch, G.A., et al., *Determinants of cytochrome P4502C8 substrate binding - Structures of complexes with montelukast, troglitazone, felodipine, and 9-cis-retinoic acid.* Journal of Biological Chemistry, 2008. **283**(25): p. 17227-17237.
71. Sevrioukova, I.F. and T.L. Poulos, *Structure and mechanism of the complex between cytochrome P4503A4 and ritonavir.* Proceedings of the National Academy of Sciences of the United States of America, 2010. **107**(43): p. 18422-18427.
72. Strushkevich, N.V., et al., *Crystal structure of human CYP7A1.* To be published.
73. Strushkevich, N.V., S.A. Usanov, and H. Park, *Crystal structure of human CYP11A1 in complex with 20,22-dihydroxycholesterol.* Proceedings of the National Academy of Sciences of the United States of America, 2011. **published before print**(June 2).
74. Ghosh, D., et al., *Structural basis for androgen specificity and oestrogen synthesis in human aromatase.* Nature, 2009. **457**(7226): p. 219-U119.
75. Mast, N., et al., *Crystal structures of substrate-bound and substrate-free cytochrome P450 46A1, the principal cholesterol hydroxylase in the brain.* Proceedings of the National Academy of Sciences of the United States of America, 2008. **105**(28): p. 9546-51.
76. Mast, N., et al., *Structural basis of drug binding to CYP46A1, an enzyme that controls cholesterol turnover in the brain.* Journal of biological chemistry, 2010. **285**(41): p. 31783-95.
77. Wester, M.R., et al., *The structure of human cytochrome P4502C9 complexed with flurbiprofen at 2.0-angstrom resolution.* Journal of Biological Chemistry, 2004. **279**(34): p. 35630-35637.
78. Rowland, P., et al., *Crystal structure of human cytochrome P450 2D6.* Journal of Biological Chemistry, 2006. **281**(11): p. 7614-7622.
79. Porubsky, P.R., K.M. Meneely, and E.E. Scott, *Structures of Human Cytochrome P-450 2E1.* Journal of Biological Chemistry, 2008. **283**(48): p. 33698-33707.
80. Strushkevich, N., S.A. Usanov, and H.W. Park, *Structural Basis of Human CYP51 Inhibition by Antifungal Azoles.* Journal of Molecular Biology, 2010. **397**(4): p. 1067-1078.
81. Leach, A.R., *Molecular Modelling; Principle and Applications*. 2nd ed. 2001: Pearson Education

82. Tsai, C.S., *An introduction to computational biochemistry*. first edition ed. 2002, New York: Wiley-Liss.
83. Patel, S., A.D. Mackerell, Jr., and C.L. Brooks, 3rd, *CHARMM fluctuating charge force field for proteins: II protein/solvent properties from molecular dynamics simulations using a nonadditive electrostatic model.* Journal of Computer Chemistry, 2004. **25**(12): p. 1504-14.
84. Jorgensen WL, M.D., TiradoRives J, *Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids.* Journal of the American Chemical Society, 1996. **118**(45): p. 11225-11236.
85. Duan, Y., et al., *A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations.* Jouranl of Computational Chemistry, 2003. **24**(16): p. 1999-2012.
86. Schuler LD, D.X., Van Gunsteren WF, *An improved GROMOS96 force field for aliphatic hydrocarbons in the condensed phase.* Journal of Computational Chemistry, 2001. **22**(11): p. 11225-11236.
87. Weiner, S.J., Kollman, P. A., Case, D. A., Singh, U. C., Ghio, C., Alagona, G., Profeta, S., Weiner, P., *A new force field for molecular mechanical simulation of nucleic acids and proteins.* Jouranl of American Chemical Society 1984. **106**: p. 765–784.
88. Ponder, J.W. and D.A. Case, *Force fields for protein simulations.* Advances in Protein Chemistry, 2003. **66**: p. 27-85.
89. Cornell, W.D., et al., *A 2nd Generation Force-Field for the Simulation of Proteins, Nucleic-Acids, and Organic-Molecules.* Journal of the American Chemical Society, 1995. **117**(19): p. 5179-5197.
90. Bayly, C.I., et al., *A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges - the Resp Model.* Journal of Physical Chemistry, 1993. **97**(40): p. 10269-10280.
91. Kollman, P.A., *Advances and continuing challenges in achieving realistic and predictive simulations of the properties of organic and biological molecules.* Accounts of Chemical Research, 1996. **29**: p. 461-469.
92. Hornak, V., Abel, R., Okur, A., Strockbine, B., Roitberg, A. and Simmerling, C., *Comparison of multiple amber force-fields and development of improved protein backbone parameters.* Proteins, 2006. **65**: p. 712–725.
93. Best, R.B. and G. Hummer, *Optimized molecular dynamics force fields applied to the helix-coil transition of polypeptides.* Journal of Physical Chemistry B, 2009. **113**(26): p. 9004-15.
94. Garcia, A.E. and K.Y. Sanbonmatsu, *Alpha-helical stabilization by side chain shielding of backbone hydrogen bonds.* Proceeding of the National Academy of Science U S A, 2002. **99**(5): p. 2782-7.
95. Lindorff-Larsen, K., Piana, S., Palmo, K., Maragakis, P., Klepeis, J. L., Dror, R. O. and Shaw, D. E., *Improved side-chain torsion potentials for the Amber ff99SB protein force field.* Proteins, 2010. **78**: p. 1950–1958.
96. Cornell, W.D., et al., *Application of Resp Charges to Calculate Conformational Energies, Hydrogen-Bond Energies, and Free-Energies of Solvation.* Journal of the American Chemical Society, 1993. **115**(21): p. 9620-9631.
97. Michael W. Mahoney, W.L.J., *A five-site model for liquid water and the reproduction of the density anomaly by rigid, nonpolarizable potential functions* Journal of Chemical Physics, 2000. **112**(20): p. 8910-8922.
98. Mahoney, M.W.J., W. L., *A five-site model for liquid water and the reproduction of the density anomaly by rigid, nonpolarizable potential functions.* Journal of Chemical Physics, 2001. **114**(1): p. 363-366
99. Jorgensen, W.L.C., J.; Madura, J. D.; Impey, R. W.; Klein, M. L., *Comparison of simple potential functions for simulating liquid water.* Journal of Chemical Physics, 1983. **79**(2): p. 926-935.

100. Berendsen, H.J.C.G., J. R.; Straatsma, T. P., *The missing term in effective pair potentials.* Journal of Physical Chemistry, 1987. **91**(24): p. 6269-6271.
101. Florova, P., et al., *Explicit Water Models Affect the Specific Solvation and Dynamics of Unfolded Peptides While the Conformational Behavior and Flexibility of Folded Peptides Remain Intact.* Journal of Chemical Theory and Computation, 2010. **6**(11): p. 3569-3579.
102. Skopalik, J., P. Anzenbacher, and M. Otyepka, *Flexibility of human cytochromes P450: Molecular dynamics reveals differences between CYPs 3A4, 2C9, and 2A6, which correlate with their substrate preferences.* Journal of Physical Chemistry B, 2008. **112**(27): p. 8165-8173.
103. Anzenbacher, P. and J. Hudecek, *Differences in flexibility of active sites of cytochromes P450 probed by resonance Raman and UV-Vis absorption spectroscopy.* Journal of inorganic biochemistry, 2001. **87**(4): p. 209-13.
104. Bancel, F., et al., *High pressure: a new tool to study P450 structure and function.* Methods Enzymol, 2002. **357**: p. 145-57.
105. Jung, C., et al., *Compressibility of the heme pocket of substrate analogue complexes of cytochrome P-450cam-CO. The effect of hydrostatic pressure on the Soret band.* European journal of biochemistry, 1995. **233**(2): p. 600-6.
106. Wang, B. and S.F. Zhou, *Synthetic and natural compounds that interact with human cytochrome P450 1A2 and implications in drug development.* Current Medicinal Chemistry, 2009. **16**(31): p. 4066-218.
107. Otyepka, M., K. Berka, and P. Anzenbacher, *Is there a relationship between the substrate preferences and structural flexibility of cytochromes P450?* Current Drug Metabolism, 2012. **13**(2): p. 130-142.
108. Anzenbacherova, E., et al., *Flexibility and stability of the structure of cytochromes P450 3A4 and BM-3.* European journal of biochemistry, 2000. **267**(10): p. 2916-20.
109. Dundas, J., et al., *CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues.* Nucleic acids research, 2006. **34**(Web Server issue): p. W116-8.
110. Petrek, M., et al., *MOLE: a Voronoi diagram-based explorer of molecular channels, pores, and tunnels.* Structure, 2007. **15**(11): p. 1357-63.
111. Fishelovitch, D., et al., *Theoretical characterization of substrate access/exit channels in the human cytochrome P450 3A4 enzyme: involvement of phenylalanine residues in the gating mechanism.* The journal of physical chemistry B, 2009. **113**(39): p. 13018-25.
112. Gotoh, O., *Substrate recognition sites in cytochrome P450 family 2 (CYP2) proteins inferred from comparative analyses of amino acid and coding nucleotide sequences.* Journal of biological chemistry, 1992. **267**(1): p. 83-90.
113. Zhao, Y. and J.R. Halpert, *Structure-function analysis of cytochromes P450 2B.* Biochimica et biophysica acta, 2007. **1770**(3): p. 402-12.
114. McCarthy, A.N. and J.R. Grigera, *Effect of pressure on the conformation of proteins. A molecular dynamics simulation of lysozyme.* Journal of Molecular Graphics and Modelling, 2006. **24**(4): p. 254-61.
115. Collins, M.D., et al., *Cooperative water filling of a nonpolar protein cavity observed by high-pressure crystallography and simulation.* Proceedings of the National Academy of Science U S A, 2005. **102**(46): p. 16668-71.
116. Day, R. and A.E. Garcia, *Water penetration in the low and high pressure native states of ubiquitin.* Proteins, 2008. **70**(4): p. 1175-84.
117. MacRaild, C.A., et al., *Global changes in local protein dynamics reduce the entropic cost of carbohydrate binding in the arabinose-binding protein.* Journal of Molecular Biology, 2007. **368**(3): p. 822-32.

APPENDIX I

Hendrychova, T., et al., *Flexibility of human cytochrome P450 enzymes: molecular dynamics and spectroscopy reveal important function-related variations.* Biochimica et biophysica acta, 2011. **1814**(1): p. 58-68.

http://www.sciencedirect.com/science/article/pii/S1570963910002116

APPENDIX II

Berka K., **Hendrychova T**., Anzenbacher P., Otyepka M., *Membrane position of ibuprofen agrees with suggested access path entrance to cytochrome P450 2C9 active site.* J Phys Chem A, 2011. **115**(41): p. 11248-55.

http://pubs.acs.org/doi/abs/10.1021/jp204488j

APPENDIX III

Berka, K., Anzenbacherova E., **Hendrychova** T., Lange R., Masek V., Anzenbacher P., Otyepka M., *Binding of quinidine radically increases the stability and decreases the flexibility of the cytochrome P450 2D6 active site.* J Inorg Biochem, 2012. **110C**: p. 46-50.

http://www.sciencedirect.com/science/article/pii/S0162013412000542

APPENDIX IV

Hendrychova, T., et al., *Dynamics and hydration of the active sites of mammalian cytochromes P450 probed by molecular dynamics simulations.* Current Drug Metabolism, 2012. **13**(2): p. 177-189.

http://www.benthamscience.com/contents-JCode-CDM-Vol-00000013-Iss-00000002.htm

APPENDIX V

Child E.S., **Hendrychova T**., McCague K., Futreal A., Otyepka M., Mann DJ, *A cancer-derived mutation in the PSTAIRE helix of cyclin-dependent kinase 2 alters the stability of cyclin binding.* Biochim Biophys Acta, 2010. **1803**(7): p. 858-64

http://www.ncbi.nlm.nih.gov/pubmed/20399812

Biochimica et Biophysica Acta 1814 (2011) 58–68

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbapap

Flexibility of human cytochrome P450 enzymes: Molecular dynamics and spectroscopy reveal important function-related variations

Tereza Hendrychová [a], Eva Anzenbacherová [b], Jiří Hudeček [c], Josef Skopalík [a], Reinhard Lange [d], Peter Hildebrandt [e], Michal Otyepka [a,*], Pavel Anzenbacher [f,*]

[a] Department of Physical Chemistry, Faculty of Science, Palacky University, Olomouc, Czech Republic
[b] Department of Medicinal Chemistry, Faculty of Medicine and Dentistry, Palacky University, Olomouc, Czech Republic
[c] Department of Biochemistry, Faculty of Science, Charles University, Prague, Czech Republic
[d] INSERM, University of Montpellier, France
[e] Technische Universität Berlin, Institut für Chemie, Sekr. PC14, Straße des 17. Juni 135. D-10623 Berlin, Germany
[f] Department of Pharmacology, Faculty of Medicine and Dentistry, Palacky University, Olomouc, Czech Republic

ARTICLE INFO

Article history:
Received 13 March 2010
Received in revised form 11 July 2010
Accepted 14 July 2010
Available online 23 July 2010

Keywords:
Molecular dynamics
Flexibility
Malleability
Cytochrome P450
UV/VIS
Raman spectroscopy

ABSTRACT

To gain more complete insight into flexibility and malleability of five forms of human liver cytochrome P450 enzymes, which play major roles in drug metabolism (CYPs 1A2, 2A6, 2C9, 2D6 and 3A4), we employed UV/VIS and resonance Raman spectroscopy in combination with all-atomic molecular dynamics simulations under normal and high pressure conditions (300 MPa). In general, the high pressure reduces the flexibility of CYPs, which become more dense and compact as their radii of gyration and temperature B-factors diminish. The flexibility of CYPs spans the regions, which are localized in solvent exposed loops. A considerable degree of flexibility is also observed at amino-acids making the pw2 and solvent channels, which are suggested to serve for substrate access and/or product release. The number of water molecules as well as the number of protein backbone atoms of the active site in close proximity of heme cofactor generally increases under high pressure. This finding provides new insights regarding the interpretation of pressure-related Soret band red shifts. Presented results also point towards considerable differences between the CYP forms studied: CYP2A6 and CYP1A2 have the least malleable active sites while those of CYP2D6, CYP2C9 and CYP3A4 have considerably greater degrees of flexibility or malleability. In addition, the number of water molecules in the active site cavity of CYP3A4 anomalously decreases under high pressure due to opening of the active site. These results correlate with the known substrate promiscuity of the respective CYP forms, with CYP3A4 displaying the highest substrate promiscuity, corresponding to the most open and malleable active site, whereas CYP1A2 and CYP2A6 show a high substrate-specificity and have a small and rigid active sites.

© 2010 Elsevier B.V. All rights reserved.

1. Introduction

Cytochromes P450 (CYP) are heme-containing monooxygenases constituting a superfamily of enzymes present in both eukaryotic and prokaryotic organisms [1]. Human liver microsomal CYP enzymes, discussed here, play major roles in the first phase of metabolism of xenobiotics [2–5]. The mechanisms whereby the microsomal CYPs accommodate various substrates and oxidize them (in a stereo-specific and regio-specific manner) remain intriguing aspects of the biochemistry of CYPs, with significant biological, pharmacological, toxicological and biotechnological implications [3,6]. Knowledge recently acquired from microsomal CYP X-ray structures has enabled the identification of their common structural features (e.g. topology, Fig. 1) and variations. The most important structural variations among CYPs described to date include: (i) the preferred access/egress paths to their buried active sites [7], (ii) the degree of their flexibility [8,9], (iii) the size of the active site [8,10] and (iv) the water network inside the active site [11,12].

Flexibility, one of the key modulators of enzyme specificity [13], appears to play an important role in the entry and accommodation of substrates in the active site, and in the release of the respective products of reactions catalyzed by CYPs [7,9,14–17]. X-ray analysis has provided highly valuable insights into the structural properties of the enzymes, but it provides only indirect evidence regarding their flexibility [9]. Therefore, conclusions regarding flexibility based on X-ray crystal structures should be crosschecked using other techniques, such as electronic (UV/VIS absorption) and vibrational (IR, Raman) spectroscopy [18–20] and molecular

Abbreviations: HP, high pressure (300 MPa); NP, normal pressure/normal conditions (0.1 MPa); CYP, cytochrome P450; RR, resonance Raman; RDF, radial distribution function; CN, coordination number; Rg, radius of gyration; B_f, B-factor

* Corresponding authors. M. Otyepka is to be contacted at tr. 17. listopadu 12, 771 46, Olomouc, Czech Republic. Fax: +420 585634761. P. Anzenbacher, Hnevotinska 3, 775 15 Olomouc, Czech Republic. Tel.: +420 585632569.
E-mail addresses: michal.otyepka@upol.cz (M. Otyepka), anzen@tunw.upol.cz (P. Anzenbacher).

1570-9639/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.bbapap.2010.07.017

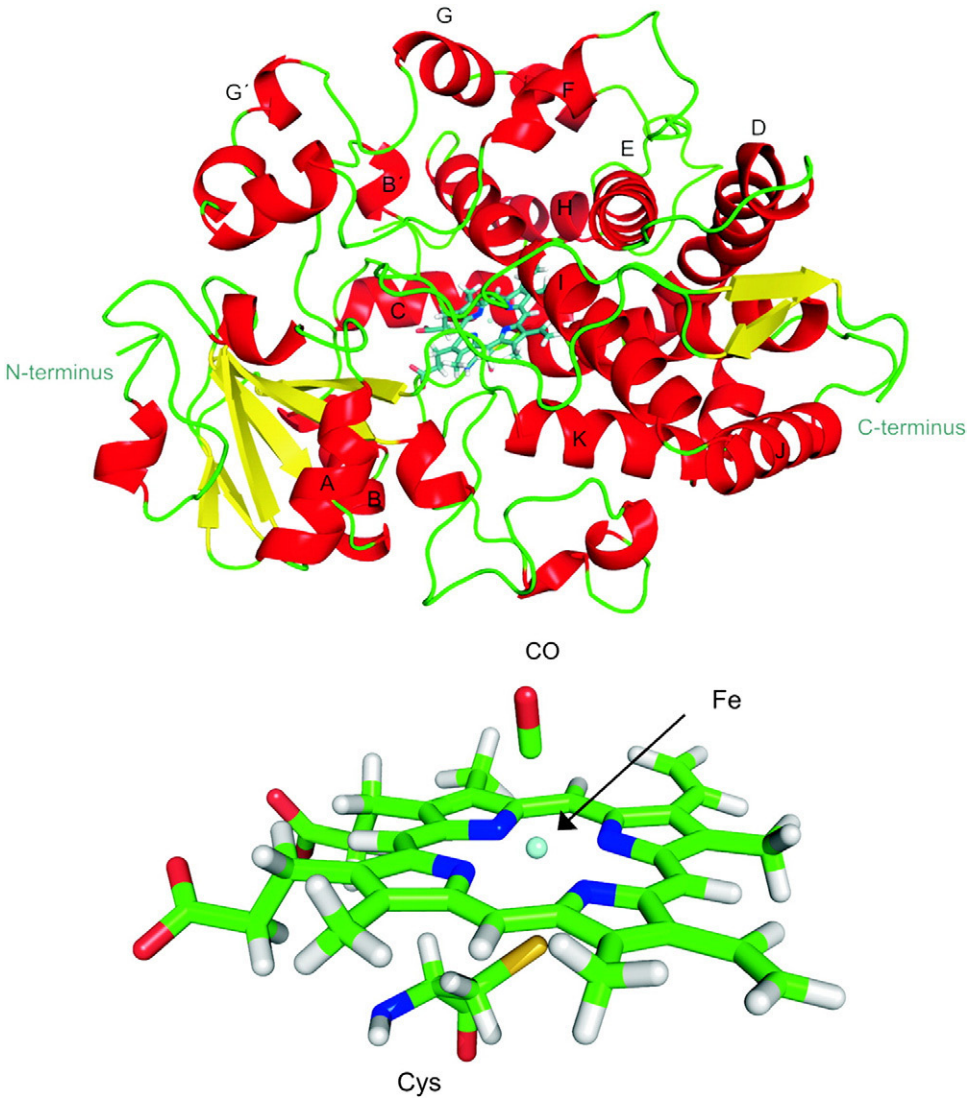


Fig. 1. *Top*: topology of cytochrome P450 (here CYP3A4) with labeled secondary structure elements (α-helices in red curls labeled by capitals, β-strands in yellow stripes). The heme (represented by sticks) is located in the buried active site near the I helix. *Bottom*: structure of the heme as used in the present MD simulations. The central iron atom (cyan ball) is bound to the protoporphyrin ring, coordinated by a CO molecule (on the distal side) and cysteinate (on the proximal side).

dynamics (MD) simulations [9,11,21–23]. Spectroscopic analyses at normal and high pressures provide information on the compressibility of the active site [18,24,25] and resonance Raman (RR) spectra allow assessing, inter alia, the flexibility of the heme vinyl side chains [18,26,27]. Analysis of the heme vinyl vibrations has shown that two stretching modes at ca. 1620 and 1631 cm^{-1} are present in the CYP spectra, which have been assigned to in-plane and out-of-plane conformations of these substituents. The presence of both bands in the spectrum suggests a relatively freely placed heme chromophore, with little or no constraints of the heme vinyls [26]. MD simulation can provide valuable additional information since it offers all-atomic resolution, but its utility is limited by the quality of the force field used and conformational space sampled in the nanoseconds timescale.

In our previous study [9], we have analyzed the flexibility of three human CYPs and its changes caused by high temperature. Here we have extended this approach to study the changes accompanying an increased pressure and to correlate the theoretical results with spectroscopic data. Also, we have included two additional human CYP forms (1A2 and 2D6) in the comparison.

The CYP enzymes examined here thus cover the metabolically most important systems. CYP3A4 is the most abundant P450 enzyme in human liver and intestines, involved in the biotransformation of more than half of the marketed drugs with known metabolism [2,3]. CYP2C9 is the major CYP2C form in human liver, responsible for the biotransformation of both weakly acidic drugs and many endogenous compounds [1,3]. CYP2A6 metabolizes about 1% of the drugs transformed by CYPs, typical substrates being coumarin and steroids [1,3]. It also contributes to the metabolism of nicotine and tobacco-specific procarcinogens [28,29]. CYP2D6 is responsible for the metabolism of at least 20% of known drugs that have a basic nitrogen and a planar aromatic ring [30]. As in CYP2A6, the CYP2D6 active site is quite small corresponding to the low-molecular weight substrates. The last CYP studied in this work is CYP1A2, which is mainly present in human liver [3]. This CYP exhibits less than 80% amino acid sequence identity to other mammalian cytochromes P450 [31]. Hence, the five CYPs selected for this study include two enzymes (CYP2A6 and CYP1A2) with relatively small active sites [29,32] and relatively narrow substrate specificities [1,3,32], the most promiscuous CYP known (CYP3A4) [1,3] with a large active site [14,33,34], and two enzymes known to have active sites with intermediate volumes (CYP2D6 and CYP2C9) [30,35,36].

2. Methods

2.1. Molecular dynamics

The MD simulations presented here model the behavior of the enzymes under conditions corresponding to those applied in the experimental studies. We have developed force field parameters for the non-standard heme residue to model the heme coordinated by the cysteinate sulfur as the fifth axial heme ligand (typical of all CYP enzymes) following the procedure of Cornell et al. [37]. The heme cofactor is, hence, represented in all simulations with a Cys residue bound to the heme iron and the carbon monoxide in the second axial position (Fig. 1). A similar approach has been successfully applied in our previous work [9]. The starting geometry of the heme residue heavy atoms for QM calculation was taken from the X-ray CYP2C9 structure (PDB ID: 1OG2 [35]) and missing CO and hydrogen atoms were added (see Supplementary materials for parameters). The system described was optimized and the RESP charges were produced in the same way as described in our previous publication [9]. The non-standard force field parameters were taken from the literature [38–40]. A 10 ns long *NpT* ($T = 298.15$ K, $p = 0.1$ MPa) simulation in TIP3P water box of the nonstandard residue capped by acetyl and N-methyl residues showed that the heme geometry is well maintained.

MD simulations of all CYP enzymes studied (Table 1) were carried out using the AMBER 9.0 package [41] with the *parm99* force field [42]. Starting structures of CYP1A2 (PDB ID: 2HI4 [31]), CYP2A6 (PDB ID: 1Z10 [29]), CYP2D6 (PDB ID: 2F9Q [30]), and CYP3A4 (PDB ID: 1TQN [34]) were taken from the Protein Data Bank database. The structure of wild-type (wt) CYP2C9 was taken from our previous MD simulation [9] because the X-ray structure (PDB ID: 1OG2) of CYP2C9 [35] represents a chimera since seven additional mutations in the F/G-loop (K206E, I215V, C216Y, S220P, P221A, I222L and I223L) have been inserted in the primary sequence to enhance crystallization. The inserted mutations do not significantly affect the catalytic activity of CYP2C9 as reported, and it is believed that they do not significantly alter its structure [21,35]. However, our previous classical MD simulations indicated that they cause small but significant differences in access path preferences [9]. The structures of ligand-free CYP1A2 and CYP2A6 forms were derived from the X-ray structures of CYP1A2-α-naphthoflavone (PDB ID: 2HI4) and CYP2A6-coumarin complexes (PDB ID: 1Z10), respectively. The missing residues (between P41 and Q52) in the CYP2D6 structure were inserted from the CYP2C8 (PDB ID: 1PQ2 [43]) structure after necessary modifications. Backbone atoms of residues 37-52 were taken from the structure of CYP2C8 and the respective side chains (according to the CYP2D6 sequence) were built in the Leap program from the AMBER package. All structures for MD simulations represent soluble CYP forms with truncated N-terminal transmembrane helices. Well-established MD simulation protocols [9,44] were used as follows. Initially, the protonation states of all histidines were checked by WHATIF [45] to create an optimal H-bond network. Then, all hydrogens were added using the Leap program from the AMBER package. The structures were neutralized by adding seven Cl^- (CYP1A2), six Cl^- (CYP2A6), two Na^+ (CYP2C9), six Na^+ (CYP2D6) and six Cl^- (CYP3A4) counterions. Each system was inserted in a rectangular TIP3P water box in which each protein atom was at least 10 Å away from the nearest edge of the box. Then, each system was minimized prior to the production part of the MD run in the way described in [9]. Subsequent to relaxation, the system was heated (0.5 ns) and after equilibration (5 ns long) production MD simulations (10 ns long) were carried out under periodic boundary conditions in the *NpT* ensemble at ambient (NP, 0.1 MPa) or high (HP, 300 MPa) pressures and at a temperature of 298.15 K using 2-fs time integration steps (see ref. for further details [9]). Each of the five systems comprises ~50,000 atoms. Table 1 summarizes details of the production phases of the systems. All MD trajectories were analyzed by the ptraj module of the AMBER package. The mean structures were produced from the last 5-ns of the respective MD simulations.

To assess the effects of pressure on the properties of the active sites, the radial distribution functions (RDF) of water molecules and enzyme backbone atoms around the heme were calculated. The mean coordination number of water molecules (CN^W) in the proximity of heme was determined by integrating the RDF of water oxygen atoms found close to the carbon monoxide (acting as the heme axial ligand), at distances of 4.0 and 5.0 Å from the center of the C-O bond. In analogy, the mean coordination number of enzyme backbone atoms (CN^B) was determined by integrating the respective RDF at distances of 4.0 and 5.0 Å.

2.2. UV/VIS spectroscopy at high hydrostatic pressure

Absorption spectra were acquired with a Cary 3E (Varian, Palo Alto, CA) spectrophotometer equipped with a high-pressure cell with sapphire windows placed in the sample compartment. Details of the instrument and the applications of the method in protein and CYP studies have been presented previously [19]. The CYP enzymes were generally diluted to a 2 μM concentration by a 20 mM K/PO_4 buffer (pH 7.4) containing 20% (v/v) glycerol, reduced by solid dithionite and complexed with carbon monoxide. The Soret absorption band position was determined by spectral analysis and its shift with applied pressure was plotted according to the equation $S = S_0 + \alpha P$,

Table 1
Key data of the MD simulations.

NP						HP			
CYP	Initial PDB	t[a] (ns)	R_g[b] (Å)	RMSD[c] (Å)	B_f[d]	t[e] (ns)	R_g[b] (Å)	RMSD[c] (Å)	B_f[d]
1A2	2HI4	10.0	22.73 ± 0.05	1.35 ± 0.08	9.2	10.0	22.57 ± 0.05	1.09 ± 0.07	8.7
2A6	1Z10	10.0	22.55 ± 0.06	1.28 ± 0.14	11.1	10.0	22.24 ± 0.04	1.12 ± 0.14	8.1
2A6	1Z10[f]	5.0	22.63 ± 0.05	1.37 ± 0.08	-	-	-	-	-
2D6	2F9Q	10.0	22.52 ± 0.06	1.84 ± 0.07	11.4	10.0	22.29 ± 0.05	1.50 ± 0.10	11.0
2C9	1OG2[g]	10.0	22.41 ± 0.04	1.85 ± 0.10	10.5	10.0	22.25 ± 0.06	1.42 ± 0.07	8.1
2C9	1OG2[f,g]	8.5	22.67 ± 0.05	1.81 ± 0.09	-	-	-	-	-
3A4	1TQN	10.0	22.66 ± 0.05	1.38 ± 0.06	9.7	10.0	22.54 ± 0.05	1.12 ± 0.06	8.7
3A4	1TQN[f]	5.0	22.73 ± 0.06	1.33 ± 0.06	-	-	-	-	-

[a] Duration of MD production simulations after equilibration (5 ns long) under normal conditions ($T = 298.15$ K, $p = 0.1$ MPa).
[b] R_g is the mean radius of gyration of C_α, C, N atoms ± SMD calculated from the last 3 ns of the respective simulations.
[c] RMSD is the mean root-mean-square-deviation ± SMD of C_α, C, N atoms from the starting X-ray structure calculated from the last 3 ns of the respective simulations.
[d] Mean temperature factors averaged over backbone atoms from the last 1 ns of the respective simulations.
[e] Duration of MD simulations under high pressure conditions ($T = 298.15$ K, $p = 300$ MPa), before the system was equilibrated for 5 ns under normal conditions ($T = 298.15$ K, $p = 1$ atm).
[f] Data taken from our previous MD simulations [9].
[g] 1OG2 structure was modified according to the WT-CYP2C9 sequence, see subsection 2.1.

where S is the Soret band position at given pressure P, S_0 is its position at normal pressure and slope α is related to compressibility [24]. The results were processed using Sigma Plot ver. 8.0 scientific graphing software (SPSS, Chicago, IL).

2.3. Resonance Raman spectroscopy

Resonance Raman (RR) spectral data were acquired using equipment consisting of a Kr^+ ion laser (Coherent 302, Santa Clara, CA) operating at the 413 nm line with laser power<8 mW at the sample. For the RR measurements, the protein concentration was adjusted with the phosphate buffer used for high pressure experiments to 15 μM. The samples were cooled to -140 °C using a Linkam THMS cryostat (Waterfield, UK) adapted to a confocal LabRam HR800 spectrometer (Horiba Jobin Yvon, Villeneuve d´Ascq, France) equipped with a liquid nitrogen-cooled CCD camera. All the acquired spectra were processed using Origin ver. 6.0 software (OriginLab, Northampton, MA).

3. Results

3.1. Common features of the CYPs studied

There are several general trends in the data obtained by MD simulations, common to all CYP forms studied. In all cases, the overall topology of the enzyme does not significantly deviate from the X-ray structure during the 10 ns long MD simulation, as shown by the mean RMSD of backbone atoms (values between 1.28 and 1.85, see Table 1). The ultimate heme core (defined by the structural elements not more than 5 Å away from the heme iron atom, thus including the cysteinate but not necessarily the complete substrate-binding active site), is relatively rigid and the flexibility is localized in the solvent-exposed regions and the substrate-binding site under both normal pressure (NP) and high pressure (HP) conditions (*cf.* also Figs. S3 and S4). These segments coincide with the ten previously identified "flexible regions": FR1-10, according to nomenclature used in our previous work [9] (Fig. 2, see figure caption for FR definitions). The HP MD simulation gave very similar results, i.e. the regions identified as flexible under NP conditions were also flexible under HP conditions. However, there are subtle differences in the behavior of individual FRs, differing for the various CYP forms, which will be discussed for individual CYP forms later. Under HP the mean RMSD of backbone atoms decreases (Table 1, Fig. S2). In addition, the positional fluctuations of atoms are on average smaller under HP, i.e. the mean temperature B-factor (B_f) of the system is lower under HP (Table 1, Fig. 3). A decrease in the radius of gyration (Rg, reflecting the size of the molecule) is observed under high pressure (Table 1) indicating that the size of the molecule decreases under HP. In most cases, the increase in values of CN^B and CN^W indicates that enzyme backbone atoms and water molecules are closer to the heme (and CO ligand) under HP. However, in individual cases there are significant exceptions to this rule (Table 2), to be discussed later for the respective CYP forms. The average structures resulting from the MD of the last 5 ns have been analyzed under both the HP and NP conditions. Comparisons of the structures showed that the secondary structure elements, i.e., α-helixes and β-sheets, are in almost all cases identically in terms of both positions and lengths under both conditions (Fig. S1).

The UV/VIS experimental data obtained for the five CYP enzymes (reduced complex with carbon monoxide) show red shifts of the Soret absorption maximum of the heme under high pressure (displayed in Fig. 5). This behavior of the Fe(II)-CO complex was reported previously [24]. The calculated values of the slope of the linear relationship between the red shift of the Soret band and pressure, related to compressibility of the heme active site (as indicated by $-\alpha$ values, see subsection 2.2), are summarized in Table 2. For the individual CYPs, different tendencies to form the inactive P420 state (exhibiting Soret absorption maximum of the carbon monoxide complex at about 420 nm) were observed as discussed below.

The RR data in the frequency region between 1600 and 1650 cm^{-1} including the vinyl stretching modes, show considerably different patterns for the various forms (*vide infra*, Fig. 6).

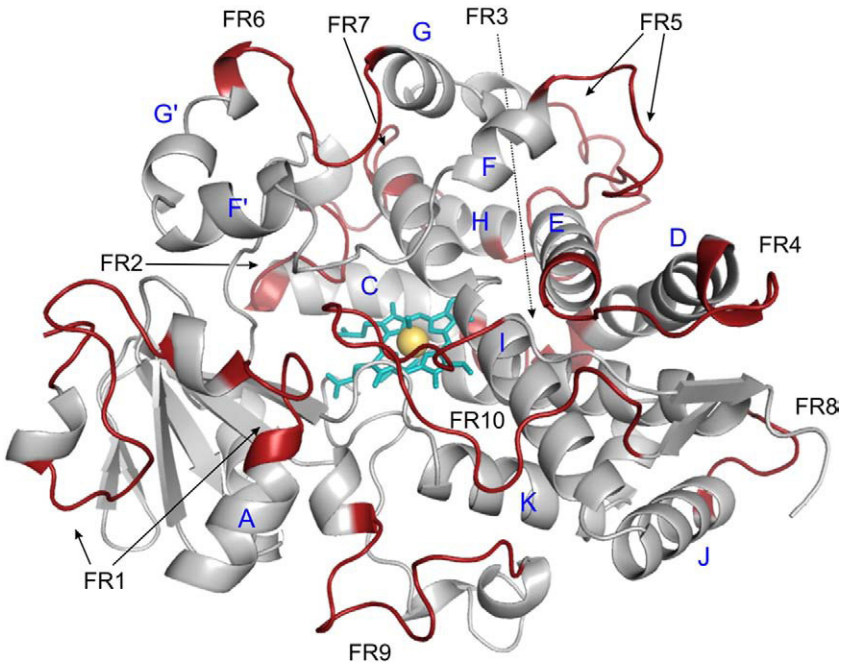


Fig. 2. Ten flexible regions (FRs, labeled from FR1 to FR10) were identified in the MD simulations: FR1 comprises of N-terminal domain loops (excluding rigid β1 core strands and residues contributing to the β1 sheet), FR2, FR3, and FR4 correspond to B′/C, C/D and D/E loops, respectively, FR5 is spread over residues of E/F and G/H loops, FR6, FR7 and FR8 coincide with the G′/G loop, H/I and J/J′ loop, respectively. FR9 corresponds to the meander and adjacent residues. FR10 corresponds to the loop following the $\beta3_1$ strand. The FRs are colored in red in the secondary structure of cytochrome P450 (here CYP3A4), helices are labeled by blue characters. The heme is shown in cyan sticks and the orange ball represents the iron atom of the heme cofactor.

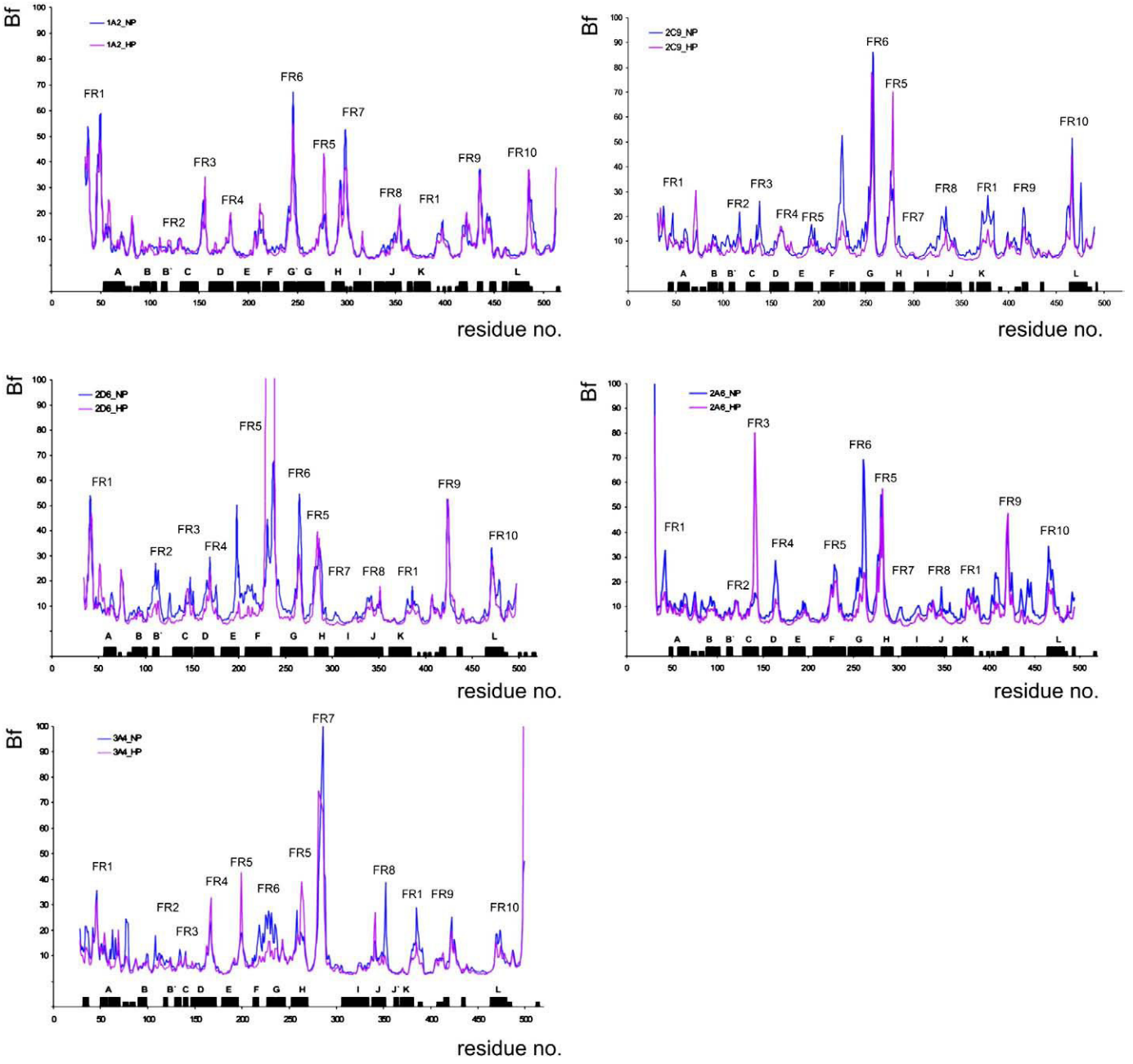


Fig. 3. Temperature B-factors calculated for all CYPs under NP and HP from last 1 ns of corresponding MD simulations. Normal pressure simulations (NP) are depicted in blue line, high pressure simulations (HP) are depicted in magenta line.

Table 2
Various parameters[a] derived from the MD simulations under normal and high pressure.

NP					HP				HP-NP				
CYP	CN^W 4.0 Å	CN^W 5.0 Å	CN^B 4.0 Å	CN^B 5.0 Å	CN^W 4.0 Å	CN^W 5.0 Å	CN^B 4.0 Å	CN^B 5.0 Å	ΔCN^W 4.0 Å	ΔCN^W 5.0 Å	ΔCN^B 4.0 Å	ΔCN^B 5.0 Å	$-\alpha$ cm^{-1}/MPa
1A2	0.84	1.95	2.61	4.23	1.02	2.27	2.15	4.29	0.18	0.32	-0.46	0.06	0.253 [25]
2A6	0.01	0.06	0.60	2.76	0.18	0.57	0.53	2.88	0.18	0.51	-0.07	0.12	0.233
2D6	1.00	1.19	0.68	3.39	1.11	2.17	0.72	3.58	0.11	0.98	0.04	0.19	0.301
2C9	1.09	2.54	0.43	2.51	1.78	3.74	0.50	2.91	0.69	1.19	0.07	0.40	0.273
3A4	3.00	4.47	0.33	2.48	2.80	4.16	0.36	2.69	-0.20	-0.31	0.04	0.21	0.449
r	-0.97	-0.86	0.41	0.42	-0.89	-0.72	0.41	0.42	0.64	0.65	-0.41	-0.18	

[a] CN, coordination number of the CO heme ligand; the changes of CN, i.e ΔCN (= CN(HP) – CN(NP)) are calculated from NP and HP pressure MD simulations. The CN^W and CN^B values are the coordination numbers of water molecules and CYP backbone atoms, respectively, calculated at two distances, 4.0 and 5.0 Å (see subsection 2.1 for details). The experimental parameter -α is related to compressibility and was obtained from the UV/VIS high pressure experiments (see subsection 2.2); pressure applied up to 250 MPa, with CYP1A2, up to 300 MPa . The Pearson's correlation coefficient *r* is shown in the last row.

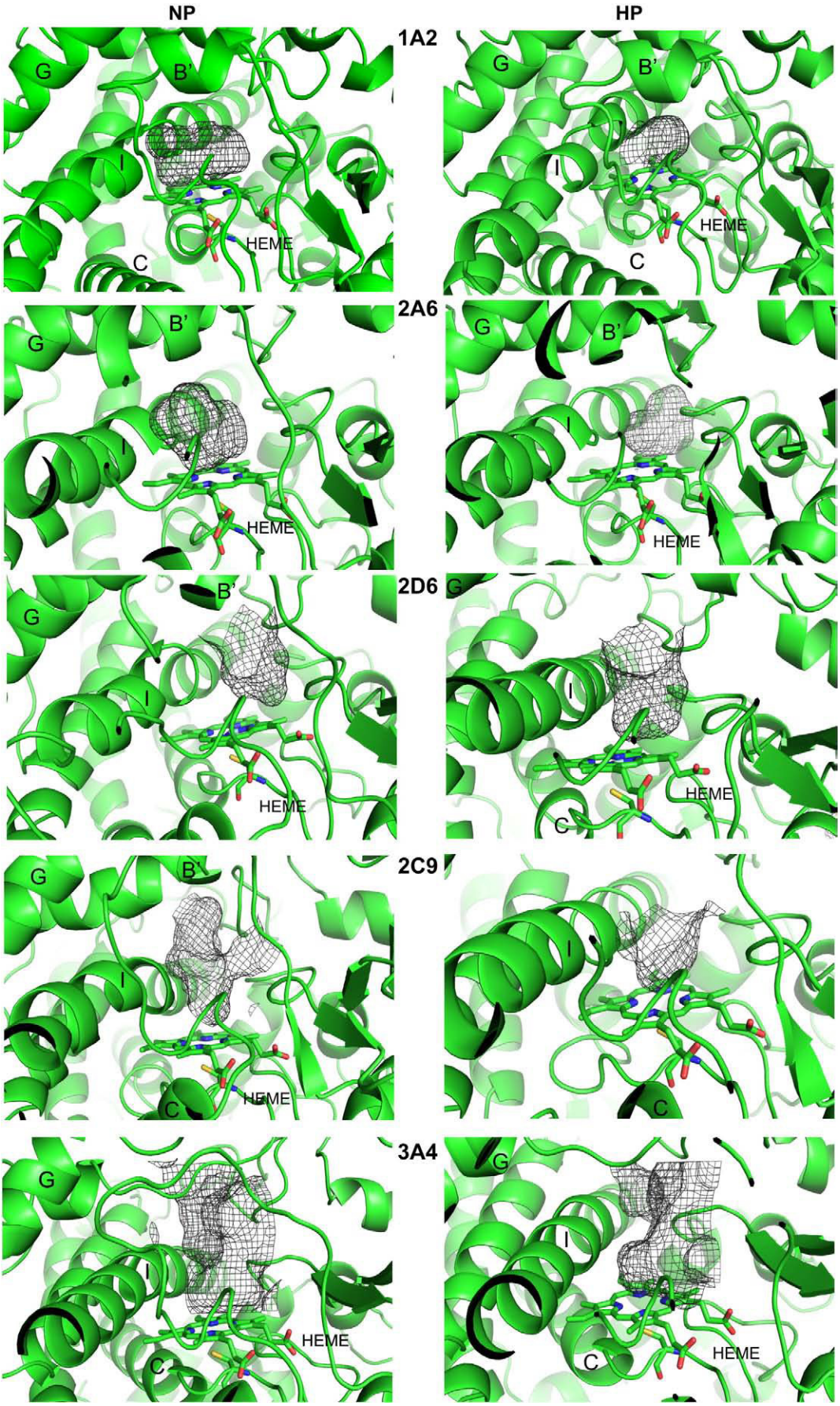


Fig. 4. Cavities (in mesh representaion) above the heme cofactor (in sticks) of Cytochromes P450 analyzed in this study (CYPs 1A2, 2A6, 2D6, 2C9 and 3A4, labeled in the middle). On the right and left side (of each row) the cavities are shown under normal pressure conditions (NP, 0.1 MPa) and under high pressure (HP, 300 MPa) conditions, respectively.

3.2. CYP2A6, CYP1A2

Difference spectra of the reduced complex of CYP2A6 with carbon monoxide (Fig. 5) (as well as the CYP1A2 [46]) does not indicate a significant degree of denaturation up to 300 MPa. Also, the $-\alpha$ value, reflecting compressibility of the heme active site, was the lowest for CYP2A6 (0.233 cm^{-1}/MPa, close to the previously reported value of 0.253 cm^{-1}/MPa for CYP1A2 [25]) (Table 2). Hence, both the CYP1A2 and CYP2A6 exhibit the least compressible heme active site. The heme active site of CYP1A2 was found to have considerable stability and low flexibility [25,46]. Also, the CYP2A6 form is known to have a relatively small active site [8,29] (Fig. 4), and the results of UV/VIS spectroscopy under pressure (Fig. 5) document its stability against denaturation to a P420 state and low compressibility.

In the 1620-1630 cm^{-1} region, the RR spectrum of CYP2A6 (Fig. 6) shows only one clearly distinguishable band at ca. 1630 cm^{-1}, corresponding to one vinyl conformer which, according to the high frequency, is attributed to an out-of-plane orientation. Hence, this finding points to an energetically less favorable conformer, suggesting that there is a certain strain in the active site which fixes the vinyl group in this position. The RR spectra of CYP1A2 protein, published previously [47], are also included in Fig. 6. In the frequency region between 1600 and 1650 cm^{-1}, the signal is dominated by a broad envelope peaking at 1623 cm^{-1} band, which includes the individual vinyl stretching modes as well as the porphyrin mode ν_{10} at ca. 1635 cm^{-1}. However, if both vinyl stretching modes of the in-plane and out-of-plane vinyl conformers are present, the envelope is expected to be more structured, specifically including a distinguishable peak at ca.1630 cm^{-1} corresponding to the out-of-plane conformer [18,47]. Thus, the CYP1A2 spectrum seems to be indicative of a single vinyl conformer.

The CYP2A6 exhibits the common flexibility pattern previously described [9]. The most significant fluctuations are localized mainly in FR1, FR4, FR5, FR7, FR9, FR10 under NP. The ultimate enzyme core is rigid and the flexibility is localized mainly in the solvent-exposed regions of the enzyme's distal side.

Under HP the common flexibility pattern remains intact, but the absolute flexibility of the C helix, FR4 and FR5 significantly decreases. However, some C helix residues (139-143) exhibit higher B_f values at HP than at NP (Fig. 3). The overall B_f values also decrease (Table 1). Beside the decrease in flexibility under HP, we also observed a decrease of the Rg by 0.31 Å (Table 1). The number of surrounding backbone atoms and the water molecules in the active site, CN^B and CN^W, were found to decrease under HP (Table 2).

Most of the flexibility of CYP1A2 under NP conditions refers to the previously identified "flexible regions": FR1, FR3, FR6, FR7, FR8, FR9 and FR10 (Fig. 2). Comparison of the average structures resulting from the MD of the last 5 ns runs under both the HP and NP conditions revealed the influence of HP on the FR5 and FR7 loops; the flexibility of the FR5 loop increases under HP while the flexibility of the FR7 loop is significantly lower than under NP. A decrease of Rg by 0.16 Å is observed under HP (Table 1). In addition, the CN^W value increased under HP, indicating that the water molecules are closer to the CO ligand (Table 2). The CN^B (at 5.0 Å) also increases, indicating that

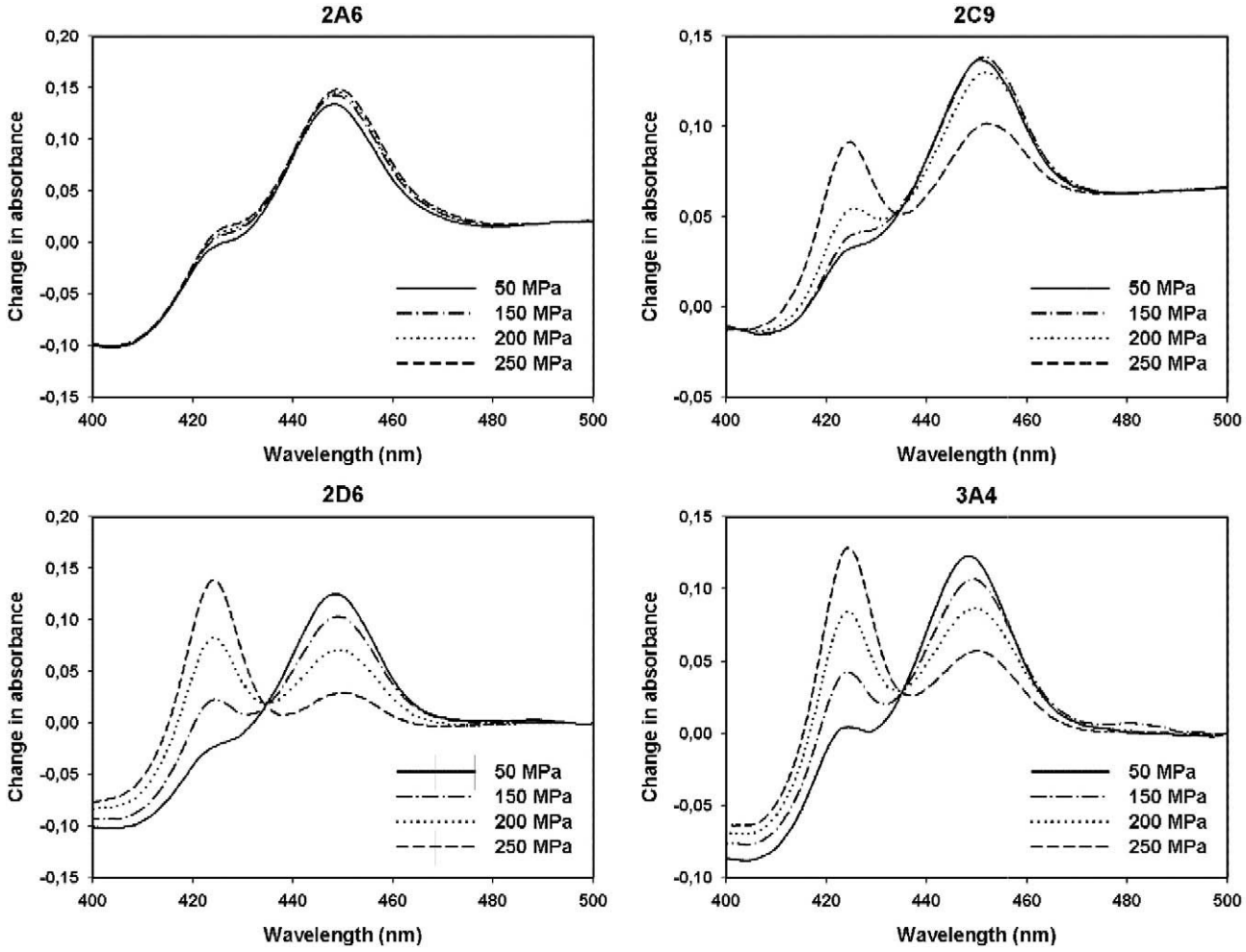


Fig. 5. Difference absorption spectra of the reduced complexes of CYP3A4, CYP2A6, CYP2D6 and CYP2C9 with carbon monoxide at hydrostatic pressures of 50 (solid line), 150 (dashed-dotted line), 200 (dotted line) and 250 MPa (dashed line). CYP concentrations were 2 μM, in 100 mM potassium phosphate buffer (pH 7.4) with 20% (v/v) glycerol.

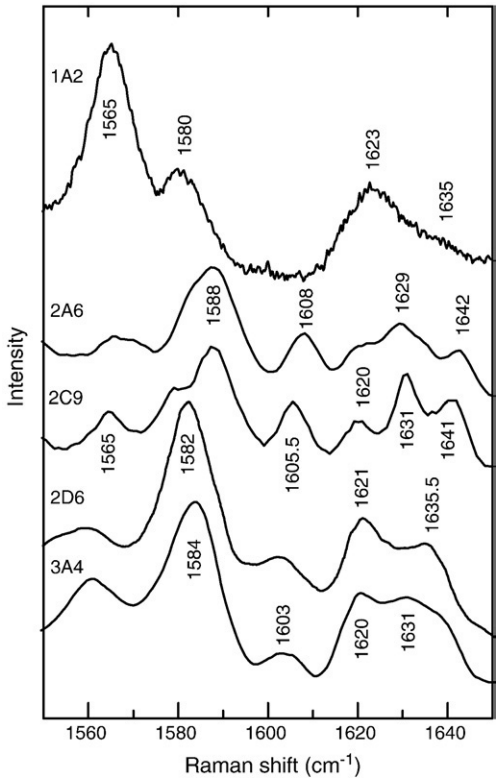


Fig. 6. Resonance Raman spectra of cytochromes P450 in the region between 1550 to 1650 cm^{-1} including the vinyl stretching modes. From top to bottom: 1A2 (data from ref. 48), 2A6, 2C9, 2D6 (data taken from ref. 49), and 3A4 (spectral trace based on data in ref. 48).

enzyme backbone atoms are closer to the heme under these conditions. It is also worth noting that the absolute CN^W values are significantly smaller than the CN^B values (Table 2).

3.3. CYP2C9, CYP2D6

The next two CYP enzymes analyzed spectroscopically under high hydrostatic pressure were CYP2C9 and CYP2D6. These two forms changed to the inactive P420 state at rather low hydrostatic pressures (CYP2C9 at 200 MPa, CYP2D6 at 100-150 MPa). This behavior corresponds to the results of the Soret band shift analysis, which gave $-\alpha$ values of 0.273 and 0.301 cm^{-1}/MPa for CYP2C9 and CYP2D6, respectively (Table 2).

The RR spectrum of the CYP2C9 protein shows the presence of both vinyl stretching modes at ca. 1621 and 1631 cm^{-1}, corresponding to of both an in-plane and an out-of-plane vinyl conformer. This finding indicates that the heme vinyls are less rigidliy fixed in the protein matrix than in CYP1A2 and CYP2A6 (Fig. 6). The RR spectrum of CYP2D6 in the absence of substrates was published by Bonifacio et al. [48]. The data show a more complex structure in the 1620-1635 cm^{-1} region; in this work [48], several band components were resolved including the 1620 and 1630 cm^{-1} bands of the vinyl stretching modes and the 1635 cm^{-1} band assigned to the ν_{10} mode of the heme.

The MD structure of CYP2C9, obtained after the 10 ns long MD simulation, does not significantly deviate from the X-ray starting structure with a mean RMSD of backbone atoms of 1.85 ± 0.10 Å (Table 1, Fig. S2). This value agrees well with the RMSD value of 1.81 ± 0.09 Å calculated from the 8.5 ns simulation run of CYP2C9 with a water molecule as the sixth heme axial ligand under ambient conditions (T = 298.15 K, p = 1 atm) [9]. The regions exhibiting high flexibility include the regions identified in our earlier work [9], namely FR2, FR3, FR5, FR6, FR8, FR9 and FR10 under NP. The flexible regions fit to the general flexibility pattern of mammalian CYPs, but with minor differences. Under HP, the flexibility of most of the regions (FR2, FR3, FR5, FR8 and FR9) significantly decreases, but the flexibility of the FR1 and FR7 loops increases under HP. Again, both the overall B_f and Rg values decrease under HP while CN^W and CN^B values increase (Tables 1 and 2).

The overall structure of CYP2D6 does not significantly change during the 10 ns long MD simulation from the X-ray starting structure, with a mean RMSD of backbone atoms of 1.84 ± 0.07 Å (Table 1). A RMSD variation from 6 to 8 ns under NP (Fig. S2) was caused by changes in two regions, residues 196-204 (in the E/F loop) and 250-256 (in the G helix). Although these residues are more flexible than the others, they do not influence the secondary structure of the protein (Fig. S1). The region that has to be modeled for the purposes of this study; for the residues 37-52 in the N-terminus (see subsection 2.1) an increase of the B_f values is noted (Fig. 3), which, however, does not cause a change in the conformation during the entire simulation and, furthermore, does not affect the overall CYP fold. The flexibility pattern of CYP2D6 also fits to the general picture of mammalian CYP [9]. The most significant fluctuations are localized mainly in FR1, FR2, FR5, FR7, FR9, FR10 under NP. The common features, i.e., a rigid enzyme core and flexible solvent-exposed regions on the enzyme distal side, are also conserved in the structure of this enzyme. Under HP the flexibility of FR2 and FR5 significantly decreases; there is no increase in flexibility observed for any loop or FR under HP in this protein. As expected, the overall B_f and Rg values decrease under HP (Table 1). In addition, both the CN^W and CN^B coordination numbers increase under HP (Table 2).

3.4. CYP3A4

This form of CYP showed similar denaturation behavior, i.e, formation of the P420 state, as a response to pressure as in the two previous cases, with significant denaturation at 100 - 150 MPa. The portion of P420 remained at approximately the same level when the pressure was subsequently relaxed to the NP level. The active site of this enzyme appeared to be the most compressible one, with an $-\alpha$ value of 0.449 cm^{-1}/MPa (Table 2), the steepest slope of the correlation between the Soret band shift of and the pressure obtained for any of the CYP enzymes studied so far. In earlier experiments [25] with CYP3A4 isolated from bacterial membranes, somewhat lower values of this coefficient were obtained (0.341 cm^{-1}/MPa in the presence of substrate). Thus, the results indicate a high degree of compressibility of the active site of this enzyme which is also relatively prone to denaturation to the P420 form.

We have shown previously [47], that CYP3A4 exhibits a complex band structure in the RR spectrum between 1600 and 1650 cm^{-1}, including two vinyl stretching modes at 1620 and 1631 cm^{-1} which suggest a less constrained heme moiety in this enzyme (Fig. 6).

As in the preceding cases, the overall structure of CYP3A4 does not significantly differ from the X-ray starting structure during the 10 ns long MD simulation, with a mean RMSD of backbone atoms of 1.38 ± 0.06 Å (Table 1). CYP3A4 displays the same common flexibility pattern as the other CYP forms. Flexible parts of the structure are localized mainly in the FR1, FR4, FR5, FR6, FR7, FR8, FR9 and FR10 regions under NP.

Comparison of the average structures of the CYP3A4 enzyme under NP and HP reveals changes in FR1, FR5 and FR6. These regions are more flexible under NP than HP. The overall B_f and Rg values decrease under HP, in analogy to all other CYP forms studied in this work (Table 1). However, among them, CYP3A4 is the only protein in which the water molecules are expelled from the active site as the CN^W number decreases (ΔCN^W equals –0.20 at 4.0 Å distance and –0.31 at 5.0 Å, Table 2). An increase of CN^B is observed at 4.0 and 5.0 Å, indicating that more backbone atoms are present in the heme active site after pressure is applied (Table 2).

4. Discussion

The experimental data and MD simulations presented in this work depict various aspects of the flexibility of human microsomal CYP enzymes. This multi-faceted property is reflecting the behavior of the peptide chain as a whole, i.e. the overall stability and compactness of the folded globule, the malleability and/or flexibility of the individual segments of the molecule, the active site properties, i.e. the substrate binding cavity, the heme side chains, and the heme ligand, and the enzyme`s tendency to convert to the inactive P420 state. In particular, the active-site-related attributes have a strong relevance for the enzyme function and its modulation [13]. Differences between various P450 forms in this respect may account for some of their physiologically and pharmaceutically relevant qualities, e.g. preferences for specific substrates or inhibitors.

In addition to the data for ambient conditions, the influence of high pressure (up to 300 MPa) was studied. This comparison allows for deeper insight into the complex problem of P450 flexibility, providing information about the active site cavity and its compressibility. Results of the MD simulations served as a basis for explaining the nature of the changes observed in UV/VIS spectra.

Experimental data, both from the UV/VIS high-pressure experiments and the RR spectra in the region of vinyl side-chain modes, afford a largely consistent picture. The most flexible form among the CYPs studied in this work (and to our knowledge, also among all forms studied so far) is CYP3A4, whereas the 1A2 and 2A6 forms are the most stable ones. The other two forms investigated in this study, 2D6 and 2C9, adopt an intermediate position in this respect. These findings agree well with the known enzymology of human CYPs, in particular with the functional data on denaturation, substrate accommodation, specificity and promiscuity of individual P450 forms (F.P. Guengerich in ref. [1]).

A new quantitative basis for interpretation of the effect of high pressure on the UV/VIS spectral properties is provided by the correlation of the α factor, related to compressibility [24] (see 2.2), with the data obtained from MD simulations. The α factor was proposed to correlate with the number of water molecules in the heme surrounding [24]. Therefore, the MD simulations were used to model the number or water molecules as well as protein backbone atoms within the spheres of 4.0 and 5.0 Å radius around the center of the C-O bond of the coordinated carbon monoxide, expressed in terms of CN^W and CN^B, see 2.1. Clear correlations were found between $-\alpha$ and CN^ws at 4 and 5 Å under both NP and HP, with R^2 values of up to 0.94 (4 Å, NP), indicating that CYPs containing more water molecules in close proximity to heme are more compressible than those in which water molecules are further away. This result supports the interpretations of Soret band shifts as proposed by Jung et al. [24].

Application of high pressure changes the numbers of both water and backbone atoms (CN^W and CN^B). Thus, the possibility of relating the changes of these parameters (ΔCN^B, and ΔCN^B) with the compressibility factor α was explored. Indeed, a linear dependence of α on ΔCN^W and ΔCN^B at 5 Å was found, according to

$$-\alpha = -0.14(0.06)\Delta CN^{W}_{5.0} + 0.45(0.27)\Delta CN^{B}_{5.0} + 0.29(0.05) \quad (1)$$

where numbers in parentheses are the respective standard deviations. The fact that the correlation coefficient for this relationship ($R^2 = 0.76$) is lower than for the individual CN^w values (R^2 up to 0.94, *vide supra*, Table 2), may reflect the higher inaccuracy of the delta values, which are obtained by subtraction. In other words, the correlation shows that the changes in the UV/VIS spectra under HP, from which the factor α is derived, reflect structural rearrangements in close proximity (< ~5 Å) to the heme.

Another experimental set of data reflecting the heme active site properties is related to the conformations of the vinyl side chains, as revealed by the RR spectra. The conclusions derived from the RR spectra of the five CYP enzymes reported in this work are consistent with the X-ray structures [8,10] showing differences in the flexibility and rigidity of the active sites. Specifically, the spectra indicate a structural constraints of the active sites for CYP2A6 and CYP1A2 whereas RR data for CYP2C9, CYP2D6 and CYP3A4 show the presence of both vinyl conformers, corresponding to a higher degree of flexibility of these substituents and thus of the heme *in toto* within the active site. These findings are in full agreement with the results of the HP UV/VIS experiments as discussed above. The positions of RR peaks correspond to dihedral angle between heme plane and vinyl group double bond [26,49] and we therefore analyzed the dihedrals of both vinyl groups in MD simulations. The dihedrals fluctuated close to 180° and any differences between CYP enzymes studied were not identified. This fact indicates that the heme vinyl dihedrals parameters (taken from literature, see Methods) overestimate barriers of vinyl group out-of-plane vibrations and should be improved. This prevents a comparison of RR data with those from MD simulations, on the other hand, the global behavior and main conclusions of this article can be hardy affected by this force field limitation.

On the other hand, the experimentally determined parameters for the compressibility do not directly reflect global structural changes of the CYP molecules under HP since they are derived solely from spectral properties of the heme site. However, one cannot rule out that local structural changes in close proximity of the heme might, at least to a certain extent, also depend on global structural changes. Here, the MD simulations provide essential complementary information.

The results of MD simulations confirm our previous conclusions [9], specifically since the present analysis was more thorough and robust due to longer production runs (10 ns vs. 5 ns in the cited study). The flexibility is localized in the ten previously defined regions (FR1 to FR10), mainly in solvent-exposed loops, i.e. in a protein shell, and mostly in the CYP distal side responsible for substrate binding. Regions FR2, FR3 and FR6 partially contribute to pw2 active site access channels, which are localized in proximity of B'/C-loop [7] and thus may regulate substrate access to the active site, because the pw2 channels were suggested to be possible routes for substrate passage to the active site. In contrast, FR10 is involved in the solvent channel [7] and may affect product release [50].

Common high pressure effects of mammalian CYPs as derived from MD simulations are a decrease of the radius of gyration and a decrease in the fluctuation flexibility as reflected by the decrease of the B_f values (Table 2). In general, these findings are consistent with results obtained for lysozyme by McCarthy and Grigera [51], who found a restricted flexibility of the enzyme under HP. However, the decrease in fluctuations of the CYPs is not isotropic (homogeneous) as the B_f values (*cf.* Fig. 3) decrease substantially only in some regions (FR2, FR4, FR6, FR7, FR8, FR9, FR10), marginally in others and even increases in two regions (FR3, FR5). The regions in which large conformational changes occur in all CYPs studied under HP are residues of the end of FR4 (terminal residues of the D helix), FR5, FR7, FR9 and FR10.

Comparison of the MD simulations under NP and HP also yields important information about hydration of the active sites. The number of water molecules in the active site (CN^w values, see Table 2) of all of the studied CYPs except for CYP3A4 increases under HP. The filling of typically nonpolar protein cavities by water molecules under hydrostatic pressure has been observed experimentally [52] and in MD study on ubiquitin [53] and expected in some CYP enzymes [24,54]. In the work presented here, both the experimental data as well as data from the molecular dynamics simulations may be interpreted in favor of this effect. The anomalous behavior of CYP3A4 might be explained from the topology of its active site. In contrast to all other CYPs examined in this work, CYP3A4 does not have a steric restriction of the active site above the iron atom (Fig. 4). Thus, the pressure increase may "push out" the water from the cavity. This interpretation is supported by the decrease of number of water

molecules (CN^W) and increase of backbone atoms (CN^B) in the active site (Table 2). Interestingly, Davydov et al. indicated a decreased water accessibility of the active site of CYP3A4 under pressure [55]. The remarkable flexibility of the CYP3A4 active site, documented here by spectroscopic methods as well as by theoretical approaches, is in line with the crystallographic data obtained by Ekroos and Sjögren [14] and explains the broad substrate specificity of this enzyme known to bind large variety of substrates of different structure and size.

5. Conclusions

Differences in the flexibility of five human liver drug-metabolizing CYP proteins show that most of their flexibility is in solvent-exposed loops. The flexibility of most of these regions is diminished under high pressure and the proteins adopt a more dense and compact structure as reflected by the decrease of the radius of gyration and the temperature B-factors. Active sites of CYPs differ significantly in their properties; all MD and UV/VIS and RR spectroscopic data indicate that CYP2A6 and CYP1A2 have the least malleable active sites while CYP2D6, CYP2C9 and CYP3A4 exhibit considerably greater degrees of flexibility or malleability. CYP3A4 has the most accessible active site and is exceptional since its active site opens under pressure, expelling water molecules in contrast to other CYPs for which the number of water molecules in the heme pocket was found to increase. These properties of the active sites can be related to the different substrate specificities of the CYP forms, because CYP3A4 exhibits the lowest and CYP1A2 and CYP2A6 the highest substrate specificities.

Acknowledgements

Support through the student project PrF_2010_025 of the Palacky University, GACR grants 303/09/1001 and 203/09/H046, and the Ministry of Youth, Sports and Education of the Czech Republic grants LC512, MSM6198959216, CZ.1.05/2.1.00/01.0030 and MSM0021620808, is gratefully acknowledged. P.H. acknowledges support by the DFG (Cluster of Excellence, UniCat).

This paper is dedicated to prof. Klaus Ruckpaul, who introduced us into the P450 community in 1979 and has been a “fixed star” in the P450 heaven for all the time.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at 10.1016/j.bbapap.2010.07.017.

References

[1] P.R. Ortiz de Montellano (Ed.), Cytochrome P450: Structure, Mechanism, and Biochemistry, 3rd. ed, Kluwer Academic, New York, 2005.
[2] W.E. Evans, M.V. Relling, Pharmacogenomics: Translating functional genomics into rational therapeutics, Science 286 (1999) 487–491.
[3] P. Anzenbacher, E. Anzenbacherova, Cytochromes P450 and metabolism of xenobiotics, Cell. Mol. Life Sci. 58 (2001) 737–747.
[4] F.P. Guengerich, Cytochrome P450 and chemical toxicology, Chem. Res. Toxicol. 21 (2008) 70–83.
[5] S. Rendic, Summary of information on human CYP enzymes: Human P450 metabolism data, Drug Metab. Rev. 34 (2002) 83–448.
[6] R. Bernhardt, Cytochromes P450 as versatile biocatalysts, J. Biotechnol. 124 (2006) 128–145.
[7] V. Cojocaru, P.J. Winn, R.C. Wade, The ins and outs of cytochrome P450s, Biochim. Biophys. Acta 1770 (2007) 390–401.
[8] M. Otyepka, J. Skopalik, E. Anzenbacherova, P. Anzenbacher, What common structural features and variations of mammalian P450s are known to date? Biochim. Biophys. Acta 1770 (2007) 376–389.
[9] J. Skopalik, P. Anzenbacher, M. Otyepka, Flexibility of human cytochromes P450: Molecular dynamics reveals differences between CYPs 3A4, 2C9, and 2A6, which correlate with their substrate preferences, J. Phys. Chem. B 112 (2008) 8165–8173.
[10] P. Anzenbacher, E. Anzenbacherova, R. Lange, J. Skopalik, M. Otyepka, Active sites of cytochromes P450: What are they like? Acta Chim. Slov. 55 (2008) 63–66.
[11] P. Rydberg, T.H. Rod, L. Olsen, U. Ryde, Dynamics of water molecules in the active-site cavity of human cytochromes P450, J. Phys. Chem. B 111 (2007) 5445–5457.
[12] C.W. Locuson, T.S. Tracy, Identification of binding sites of non-l-helix water molecules in mammalian cytochromes P450, Drug Metab. Dispos. 34 (2006) 1954–1957.
[13] N. Tokuriki, D.S. Tawfik, Protein Dynamism and Evolvability, Science 324 (2009) 203–207.
[14] M. Ekroos, T. Sjogren, Structural basis for ligand promiscuity in cytochrome P450 3A4, Proc. Natl Acad. Sci. U. S. A. 103 (2006) 13682–13687.
[15] Y.H. Zhao, M.A. White, B.K. Muralidhara, L. Sun, J.R. Halpert, C.D. Stout, Structure of microsomal cytochrome P4502B4 complexed with the antifungal drug bifonazole - Insight into p450 conformational plasticity and membrane interaction, J. Biol. Chem. 281 (2006) 5973–5981.
[16] E. Stjernschantz, N.P.E. Vermeulen, C. Oostenbrink, Computational prediction of drug binding and rationalisation of selectivity towards cytochromes P450, Expert Opin. Drug Metab. Toxicol. 4 (2008) 513–527.
[17] B.K. Muralidhara, S. Negi, C.C. Chin, W. Braun, J.R. Halpert, Conformational flexibility of mammalian cytochrome P4502B4 in binding imidazole inhibitors with different ring chemistry and side chains - Solution thermodynamics and molecular modeling, J. Biol. Chem. 281 (2006) 8051–8061.
[18] P. Anzenbacher, J. Hudecek, Differences in flexibility of active sites of cytochromes P450 probed by resonance Raman and UV-Vis absorption spectroscopy, J. Inorg. Biochem. 87 (2001) 209–213.
[19] F.R. Bancel, G. Hui Bon Hoa, P. Anzenbacher, C. Balny, R. Lange, High pressure: A new tool to study P450 structure and function, Meth. Enzymol. 357 (2002) 145–157.
[20] C. Jung, Fourier transform infrared spectroscopy as a tool to study structural properties of cytochromes P450 (CYPs), Anal. Bioanal. Chem. 392 (2008) 1031–1058.
[21] A. Seifert, S. Tatzel, R.D. Schmid, J. Pleiss, Multiple molecular dynamics simulations of human p450 monooxygenase CYP2C9: The molecular basis of substrate binding and regioselectivity toward warfarin, Proteins: Struct, Funct. Bioinform. 64 (2006) 147–155.
[22] W.H. Li, H. Liu, X.M. Luo, W.L. Zhu, Y. Tang, J.R. Halpert, H.L. Jiang, Possible pathway (s) of metyrapone egress from the active site of cytochrome P450 3A4: A molecular dynamics simulation, Drug Metab. Dispos. 35 (2007) 689–696.
[23] W.H. Li, H. Liu, E.E. Scott, F. Grater, J.R. Halpert, X.M. Luo, J.H. Shen, H.L. Jiang, Possible pathway(s) of testosterone egress from the active site of cytochrome P4502B1: A steered molecular dynamics simulation, Drug Metab. Dispos. 33 (2005) 910–919.
[24] C. Jung, G. Hui Bon Hoa, D. Davydov, E. Gill, K. Heremans, Compressibility of the heme pocket of substrate-analog complexes of cytochrome P450cam-CO the effect of hydrostatic pressure on the Soret band, Eur. J. Biochem. 233 (1995) 600–606.
[25] E. Anzenbacherova, N. Bec, P. Anzenbacher, J. Hudecek, P. Soucek, C. Jung, A.W. Munro, R. Lange, Flexibility and stability of the structure of cytochromes P450 3A4 and BM-3, Eur. J. Biochem. 267 (2000) 2916–2920.
[26] W.A. Kalsbeck, A. Ghosh, R.K. Pandey, K.M. Smith, D.F. Bocian, Determinants of the Vinyl Stretching Frequency in Protoporphyrins - Implications for Cofactor-Protein Interactions in Heme-Proteins, J. Am. Chem. Soc. 117 (1995) 10959–10968.
[27] E. Anzenbacherova, J. Hudecek, D. Murgida, P. Hildebrandt, S. Marchal, R. Lange, P. Anzenbacher, Active sites of two orthologous cytochromes P450 2E1: Differences revealed by spectroscopic methods, Biochem. Biophys. Res. Commun. 338 (2005) 477–482.
[28] V. Malaiyandi, E.M. Sellers, R.F. Tyndale, Implications of CYP2A6 genetic variation for smoking behaviors and nicotine dependence, Clin. Pharmacol. Ther. 77 (2005) 145–158.
[29] J.K. Yano, M.H. Hsu, K.J. Griffin, C.D. Stout, E.F. Johnson, Structures of human microsomal cytochrome P450 2A6 complexed with coumarin and methoxsalen, Nat. Struct. Mol. Biol. 12 (2005) 822–823.
[30] P. Rowland, F.E. Blaney, M.G. Smyth, J.J. Jones, V.R. Leydon, A.K. Oxbrow, C.J. Lewis, M.G. Tennant, S. Modi, D.S. Eggleston, R.J. Chenery, A.M. Bridges, Crystal structure of human cytochrome P450 2D6, J. Biol. Chem. 281 (2006) 7614–7622.
[31] S. Sansen, J.K. Yano, R.L. Reynald, G.A. Schoch, K.J. Griffin, C.D. Stout, E.F. Johnson, Adaptations for the oxidation of polycyclic aromatic hydrocarbons exhibited by the structure of human P450 1A2, J. Biol. Chem. 282 (2007) 14348–14355.
[32] S. Sansen, M.-H. Hsu, C.D. Stout, E.F. Johnson, Structural insight into the altered substrate specificity of human cytochrome P450 2A6 mutants, Arch. Biochem. Biophys. 464 (2007) 197–206.
[33] P.A. Williams, J. Cosme, D.M. Vinkovic, A. Ward, H.C. Angove, P.J. Day, C. Vonrhein, I.J. Tickle, H. Jhoti, Crystal structures of human cytochrome P450 3A4 bound to metyrapone and progesterone, Science 305 (2004) 683–686.
[34] J.K. Yano, M.R. Wester, G.A. Schoch, K.J. Griffin, C.D. Stout, E.F. Johnson, The structure of human microsomal cytochrome P450 3A4 determined by X-ray crystallography to 2.05-angstrom resolution, J. Biol. Chem. 279 (2004) 38091–38094.
[35] P.A. Williams, J. Cosme, A. Ward, H.C. Angova, D.M. Vinkovic, H. Jhoti, Crystal structure of human cytochrome P4502C9 with bound warfarin, Nature 424 (2003) 464–468.
[36] M.R. Wester, J.K. Yano, G.A. Schoch, C. Yang, K.J. Griffin, C.D. Stout, E.F. Johnson, The structure of human cytochrome P4502C9 complexed with flurbiprofen at 2.0 angstrom resolution, J. Biol. Chem. 279 (2004) 35630–35637.
[37] W.D. Cornell, P. Cieplak, C.I. Bayly, P.A. Kollman, Application of RESP charges to calculate conformational energies, hydrogen-bond energies, and free-energies of solvation, J. Am. Chem. Soc. 115 (1993) 9620–9631.
[38] F. Autenrieth, E. Tajkhorshid, J. Baudry, Z. Luthey-Schulten, Classical force field parameters for the heme prosthetic group of cytochrome c, J. Comput. Chem. 25 (2004) 1613–1622.

[39] A. Oda, N. Yamaotsu, S. Hirono, New AMBER force field parameters of heme iron for cytochrome P450s determined by quantum chemical calculations of simplified models, J. Comput. Chem. 26 (2005) 818–826.
[40] A.D. Favia, A. Cavalli, M. Masetti, A. Carotti, M. Recanatini, Three-dimensional model of the human aromatase enzyme and density functional parameterization of the iron-containing protoporphyrin IX for a molecular dynamics study of heme-cysteinato cytochromes, Proteins: Struct, Funct. Bioinform. 62 (2006) 1074–1087.
[41] D.A. Case, T.E. Cheatham, T. Darden, H. Gohlke, R. Luo, K.M. Merz, A. Onufriev, C. Simmerling, B. Wang, R.J. Woods, The AMBER biomolecular simulation programs, J. Comput. Chem. 26 (2005) 1668–1688.
[42] J.M. Wang, P. Cieplak, P.A. Kollman, How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? J. Comput. Chem. 21 (2000) 1049–1074.
[43] G.A. Schoch, J.K. Yano, M.R. Wester, K.J. Griffin, C.D. Stout, E.F. Johnson, Structure of human microsomal cytochrome P4502C8 - Evidence for a peripheral fatty acid binding site, J. Biol. Chem. 279 (2004) 9497–9503.
[44] M. Otyepka, I. Bartova, Z. Kriz, J. Koca, Different mechanisms of CDK5 and CDK2 activation as revealed by CDK5/p25 and CDK2/Cyclin A dynamics, J. Biol. Chem. 281 (2006) 7271–7281.
[45] G. Vriend, WHAT IF, EMBL, Heidelberg, 1997.
[46] P. Anzenbacher, N. Bec, J. Hudecek, R. Lange, E. Anzenbacherova, High conformational stability of cytochrome P-450 1A2. Evidence from UV absorption spectra, Collect. Czech, Chem. Commun. 63 (1998) 441–448.
[47] J. Hudecek, E. Anzenbacherova, P. Anzenbacher, A.W. Munro, P. Hildebrandt, Structural similarities and differences of the heme pockets of various P450 isoforms as revealed by resonance Raman spectroscopy, Arch. Biochem. Biophys. 383 (2000) 70–78.
[48] A. Bonifacio, A.R. Groenhof, P.H.J. Keizers, C. de Graaf, J.N.M. Commandeur, N.P.E. Vermeulen, A.W. Ehlers, K. Lammertsma, C. Gooijer, G. van der Zwan, Altered spin state equilibrium in the T309V mutant of cytochrome P450 2D6: a spectroscopic and computational study, J. Biol. Inorg. Chem. 12 (2007) 645–654.
[49] J. Hudecek, P. Hodek, E. Anzenbacherova, P. Anzenbacher, Structural analysis of cytochromes P450 shows differences in flexibility of heme 2-and 4-vinyls, Biochim. Biophys. Acta 1770 (2007) 413–419.
[50] D. Fishelovitch, S. Shaik, H.J. Wolfson, R. Nussinov, Theoretical Characterization of Substrate Access/Exit Channels in the Human Cytochrome P450 3A4 Enzyme: Involvement of Phenylalanine Residues in the Gating Mechanism, J. Phys. Chem. B 113 (2009) 13018–13025.
[51] A.N. McCarthy, J.R. Grigera, Effect of pressure on the conformation of proteins. A molecular dynamics simulation of lysozyme, J. Mol. Graph. Model. 24 (2006) 254–261.
[52] M.D. Collins, G. Hummer, M.L. Quillin, B.W. Matthews, S.M. Gruner, Cooperative water filling of a nonpolar protein cavity observed by high-pressure crystallography and simulation, Proc. Natl Acad. Sci. U. S. A. 102 (2005) 16668–16671.
[53] R. Day, A.E. García, Water penetration in the low and high pressure native states of ubiquitin, Proteins: Struct, Funct. Bioinform. 70 (2008) 1175–1184.
[54] G. Hui Bon Hoa, M.A. McLean, S.G. Sligar, High pressure, a tool for exploring heme protein active sites, Biochim. Biophys. Acta 1595 (2003) 297–308.
[55] D.R. Davydov, J.R. Halpert, J.P. Renaud, G.H.B. Hoa, Conformational heterogeneity of cytochrome P450 3A4 revealed by high pressure spectroscopy, Biochem. Biophys. Res. Commun. 312 (2003) 121–130.

THE JOURNAL OF PHYSICAL CHEMISTRY A

ARTICLE

pubs.acs.org/JPCA

Membrane Position of Ibuprofen Agrees with Suggested Access Path Entrance to Cytochrome P450 2C9 Active Site

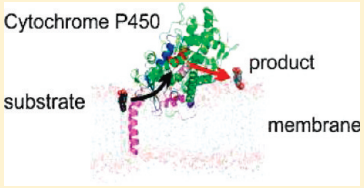
Karel Berka,[†] Tereza Hendrychová,[†] Pavel Anzenbacher,[‡] and Michal Otyepka*[,†]

[†]Regional Centre of Advanced Technologies and Materials, Department of Physical Chemistry, Faculty of Science, Palacky University, 17. listopadu 12, 771 46 Olomouc, Czech Republic

[‡]Department of Pharmacology, Faculty of Medicine and Dentistry, Palacky University, Hněvotínská 3, 775 15 Olomouc, Czech Republic

Supporting Information

ABSTRACT: Cytochrome P450 2C9 (CYP2C9) is a membrane-anchored human microsomal protein involved in the drug metabolism in liver. CYP2C9 consists of an N-terminal transmembrane anchor and a catalytic cytoplasmic domain. While the structure of the catalytic domain is well-known from X-ray experiments, the complete structure and its incorporation into the membrane remains unsolved. We constructed an atomistic model of complete CYP2C9 in a dioleoylphosphatidylcholine membrane and evolved it by molecular dynamics simulations in explicit water on a 100+ ns time-scale. The model agrees well with known experimental data about membrane positioning of cytochromes P450. The entry to the substrate access channel is proposed to be facing the membrane interior while the exit of the product egress channel is situated above the interface pointing toward the water phase. The positions of openings of the substrate access and product egress channels correspond to free energy minima of CYP2C9 substrate ibuprofen and its metabolite in the membrane, respectively.



INTRODUCTION

Cytochrome P450 enzymes (CYPs) are prominent family of biotransformation enzymes involved in the metabolism of xenobiotics as well as in pathways involving various endogenous compounds.[1] Metabolism of drugs by microsomal CYPs plays an important role in the pharmacological and toxicological effects of drugs as well as in related drug−drug interactions in humans.[2,3] Consequently, the mechanisms whereby CYPs accommodate various substrates and oxidize them in a stereo- and regio-specific manner are a topic of great interest in the biochemistry of the CYPs. Current insights into determinants of substrate specificity emerging from the available X-ray structures and biochemical data suggest that the mechanism of substrate specificity is complex; involving not only core active site residues but also other structural features, for example, access/egress paths and conformational flexibility.[4−10]

Human microsomal CYPs are anchored in the membrane of the endoplasmic reticulum by their N-terminal sequence.[11,12] However, the structural data available for the CYPs are based on analyses of engineered CYPs without the N-terminal membrane anchor and often with introduced mutations designed to enhance the protein's solubility.[7] As a result, the current X-ray data do not necessarily provide a comprehensive description of the complex structures of the membrane anchored CYPs.[8,9] Despite the growth in the body of data concerning the structural features of CYPs during the past decade, the orientation and position of CYP in the membrane at the atomic level remains unknown. Information about the membrane topology of CYPs is derived from indirect experiments, that is, from epitope search,[12,13] mutagenesis,[14−16] spectroscopy,[17,18] other experimental techniques[18−20] or from models[12,21,22] and calculations based on factors such as the hydrophobicity of amino acid residues.[7,23] These data suggests that other parts of CYP structure, besides the N-terminal membrane anchor, are immersed in the membrane; for example, part of N-terminal domain and the F/G loop.[12,24,25]

One aspect of CYP function involves the oxidation of hydrophobic substrates to make them more hydrophilic (for schematics of the enzyme action see Figure 1). In this respect, a question arises as to how the lipophilic substrates enter the buried active sites of CYPs? From mutation studies, it has been suggested that substrate access channels of CYP enzymes might be located in parts of the protein either in close contact or inside the membrane and that the lipophilic substrates may enter the active site from the membrane via an appropriate access channel.[26] Access channels from family 2 (or pw2, according to Wade's nomenclature[5]) seem to be suitable for passage of substrates to the active site. Similarly another important question is how and where the more hydrophilic products are released from the active site. The solvent channel of CYPs has been identified as a passage enabling active site solvation;[8,27] such a hydrophilic channel may also serve as a potential egress channel for the metabolites.[26] However, the exact positions of the openings of the access/egress channels in respect to the membrane remain to be determined.

Special Issue: Pavel Hobza Festschrift

Received: May 13, 2011
Revised: June 5, 2011

ACS Publications © XXXX American Chemical Society

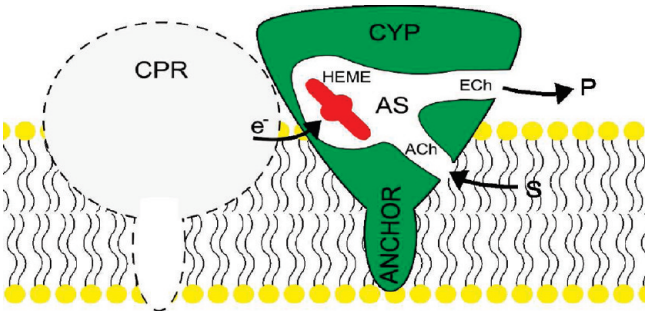


Figure 1. In this work we discuss several structural aspects of CYP2C9 anchored to the membrane at atomic resolution. The scheme shows that it is hypothesized that hydrophobic substrates (S) enter the CYP enzyme active site (AS) via an access channel (ACh). The substrate is oxidized in the active site to a more hydrophilic product (P) and leaves the active site via an egress channel (ECh) to cytosol. The oxidation involves heme (HEME) and requires electrons, which can be provided by cytochrome P450 reductase (CPR).

Here we present an atomic model of the membrane-anchored enzyme CYP2C9, the major CYP2C form in human liver, which is responsible for the biotransformation of both weakly acidic drugs such as ibuprofen and many endogenous compounds.[1,2,28] The model is cross-validated with the available experimental data and provides important information about membrane anchoring, the structural changes CYP2C9 undergoes upon interacting with the membrane and the orientations of the access/egress path within this structure. The calculated free energy profiles of ibuprofen and its metabolite (3-hydroxyibuprofen) show the disposal (penetration depth) of these compounds in the membrane. The mechanisms of substrate access and product release are further discussed with respect to the proposed positions of the substrate and product entrance in/on the membrane.

MATERIALS AND METHODS

Protein Model Building. We used the crystal structure of ligand-free CYP2C9 (PDB ID 1OG2) as a template for model building.[29] The sequence of the engineered CYP2C9 was aligned with the sequence of wild-type human CYP2C9 to detect inserted mutations and with sequences of rabbit CYP2B4 and rat CYP2B1 to compare the experimental data on membrane topology. Multiple alignment was performed using ClustalW2[30] within Jalview 2.6[31] (see Figure S1).

Membrane positioning data were collected mainly from epitope screening of rat CYP2B1 and rabbit CYP2B4,[12,13,32] and from knowledge of phosphorylation[12] and mutations in the CYP2C family.[15,25] Data from epitope screening revealed regions that are inaccessible to the antibodies (i.e., buried inside the protein or within the membrane).

Model building was done in Pymol 0.99rc6.[33] To begin with, the rescue mutations (E206K, V215I, Y216C, P220S, A221P, L222I, L223I) were applied to restore the wild-type sequence of CYP2C9. The transmembrane segment was modeled as α-helix and joined to the rest of the protein by a flexible arm. The protein was then immersed into a membrane modeled on known membrane segments derived from multiple alignment (see Figure S1 in Supporting Information). The transmembrane helix was rotated within the membrane to be approximately perpendicular to the membrane surface.

More than 50% of the membrane of the endoplasmic reticulum consists of phosphatidylcholines.[34] Due to this fact, we based our model membrane on an equilibrated dioleoylpalmitoylphosphatidylcholine (DOPC) membrane with a topology downloaded from http://www.bioinf.uni-sb.de/RB.[35] The membrane was oriented perpendicular to the *z*-axis of the box.

Protein/Membrane Simulation Protocol. Molecular dynamics simulations were used to study the stability of the model with the united-atom GROMOS96 53a6 force field[36] with Berger lipid addition[37] in Gromacs 4.0.7.[38] Parameters of heme were taken from the 53a6 force field. SPC water molecules and chloride and sodium ions were added to neutralize the system and to achieve a cellular salt concentration of 0.1 mM. After energy minimization, 200 ps long prerun was performed to equilibrate water molecules with simultaneous position constraints on protein heavy atoms. The system was then equilibrated in the NPT ensemble for a further 10 ns to equilibrate pressure and box size with an anisotropic Berendsen barostat with a time constant of 10 ps and pressure of 1 bar and V-rescale thermostat with a time constant of 0.1 ps and with temperature of 300 K. Three independent production simulations in the NPT ensemble were then run for a total time of 250 ns. The system equilibrated after 30 ns (for comparison between initial and final model see Figure S2 and for evolution of rmsd see Figure S3 in Supporting Information).

Access Channels Analysis. The access/egress channels were calculated using MOLE (http://mole.chemi.muni.cz) software.[39] MOLE outperforms the previously released CAVER,[40] which is widely used for studies of biomolecular channels and has been successfully applied to the study of CYPs.[5,8,41]

Drug Disposition Simulations. To establish the disposition of drugs in the membrane/water environment, the respective drug molecules were subjected to simulation in the same DOPC/water environment. Gromos96 parameters for the model drug (in this case ibuprofen, a typical substrate of CYP2C9) and its metabolite (3-hydroxyibuprofen) were obtained from the Dundee PRODRG 2.5 Beta server.[42] Ibuprofen (and also 3-hydroxyibuprofen) was modeled in both its uncharged and charged protonation states. Free energy profiles were calculated with the g_wham program.[43] Profiles were calculated by umbrella sampling of the potential of mean force on 45 windows with spacing of 1 Å defined on the membrane normal using a harmonic potential with a force constant of 1000 $kJ \cdot mol^{-1} \cdot nm^{-2}$. Each window was run twice for at least 3 ns: the windows with substantial gradients were run for a longer period with twice as high force constant to minimize error. Starting structures for each window were obtained from free simulations with the molecule starting in water or from pull simulations with force applied on the drug in the direction normal to the membrane.

RESULTS

Model Building. The N-terminal sequence missing from the initial X-ray structure PDB ID 1OG2[29] was modeled as an α-helix. The α-helix topology is typical for a single transmembrane segment as it is able to accommodate the hydrogen bonding of the main-chain atoms in a nonpolar environment. The α-helical topology was consistent with the findings of an NMR study of the N-terminal transmembrane segment from prostaglandin I2 synthase (also known as CYP8A1).[44] The transmembrane α-helix was terminated by a positively charged N-terminus.[1] The α-helix was attached to the CYP2C9 structure

dx.doi.org/10.1021/jp204488j |*J. Phys. Chem. A* XXXX, XXX, 000–000

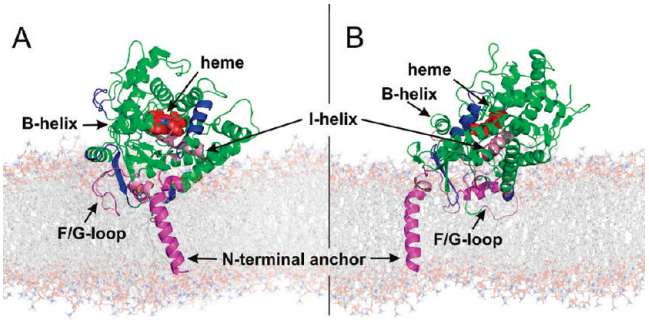


Figure 2. Snapshot of the structure of CYP2C9 in DOPC membrane taken at 50 ns. The protein is shown in green, while parts determined by epitope screening as not being exposed to water (i.e., in the membrane or inside the protein core) are shown in violet. Parts predicted to be accessible to water are shown in blue. Heme is shown by red spheres. The DOPC membrane is shown using the transparent stick representation with orange spheres at the positions of phosphorus atoms in the lipid beads. The image was prepared with Pymol 0.99rc6.[33].

and the mutations introduced into the 1OG2 structure were modified to the native wild-type sequence of CYP2C9. The initial orientation of the CYP2C9 structure on the membrane as well as the depth of the part embedded in the membrane were based on known experimental data from epitope labeling, which indicate parts of CYP inaccessible to the antibodies, that is, those inside the protein or within the membrane.[12,13,32] Arbitrary orientations of the CYP2C9 with respect to the membrane were chosen for molecular dynamics (MD) simulations and all of these systems relaxed to the topology resembling the orientation of the final model (data not shown). A snapshot taken from the equilibrated part of our simulation was used as a final model (Figure 2).

The model presented here provides some new insights and answers, but it should be noted that the model is based on classical MD simulations, which are limited mainly by the quality of the empirical force field and available time scales. The 100 ns long time scale of MD simulation enables limited sampling of the conformational space, that is, the molecules typically fluctuate close to their starting structures (this limitation is also known as the "local conformational trap"). Enhanced sampling can be in principle achieved, for example, by coarse-grained MD simulations using even more simplified coarse-grained potentials. The coarse-grained simulations, however, may suffer from the charge artifacts and also conformational changes of the protein backbone are not adequately described.[45] Both atomic and coarse-grained MD simulations should be always cross-validated with available experimental data and interpreted with care keeping in mind their inherent limitations.

Final Model Cross-Validation. The most important part of the model building is its cross-validation against available experimental data. In our analysis all of the simulations were stable and the protein was relatively rigid. The membrane topology did not change significantly from that used as an input. From this perspective it is not surprising that the information gathered from the epitope screening,[12,13,32] which was used as the main input information during the model construction (see Methods section for details), were also valid for the final model. The regions inaccessible to the antibodies are inside the protein (positions 93−98, B′-helix, and 287−299, I-helix) or within the membrane (positions 1−46, N-terminal anchor, and 207−220, F/G loop).

Protein rotation studies on several CYPs shows that the tilt angle between the heme plane and the membrane plane is variable for different CYPs from approximately 38 to 78°.[46] The tilt angle between the heme and the plane of the membrane was $(35 \pm 5)°$ in our simulations; however, this is still acceptably consistent with the experimental data.

A tryptophan fluorescence scanning study on CYP2C2 suggested that tryptophan residues within the A-helix and the β2-2 sheet (positions 36, 69, and 380) are positioned within the membrane as their fluorescence was quenched by the probes with nitroxide spin labels on fatty acid tails.[47] The authors concluded that the F/G-loop (position 225) is not localized within the membrane. However, the immersion of the F/G-loop in the membrane is based on observations from previous fluorescent studies,[48] mutagenesis,[15,25,29] and trypsinolysis.[49,50] This apparent paradox could possibly be resolved by considering the homodimerization of CYPs from the 2C family. The dimerization interface in crystal structures of truncated CYP2C8 (which is highly homologous to CYP2C9) is formed by the F/G-loop region even with lipids bound.[51] The F/G loop was also identified recently to be at the dimerization interface even in physiological membranes in addition to the N-terminal signaling domain.[52,53] Dimers are mainly formed at higher concentrations; at lower concentrations, CYP2C8 exists primarily as the monomer.[51] Because the concentrations of CYPs in the tryptophan fluorescence scanning study were quite high due to their increased expression in the insect cells,[47] it is possible that the majority of CYPs were dimers under these conditions and that these dimers were likely in contact by the F/G-loop region. The F/G-loop region also forms the dimeric surface with lipids present in the CYP2C8 crystal structure PDB ID 2NNH.[51] On the other hand, older data from trypsinolysis and epitope searching were mainly collected from studies on microsomes with lower concentrations of CYPs, where the CYPs may be present mainly in the form of monomers. Our model represents a monomeric CYP2C9 with an F/G loop and A′-helix inside the membrane. On the other hand, we cannot rule out a possibility that the membrane position of F/G-loop depends on local conditions (e.g., formation of oligomers) or dynamics equilibrium beyond our simulation time scale.

The position of CYPs in the membrane has also been studied by AFM[19,54] and the Langmuir−Blodgett monolayer technique.[20] It was found that the height of the CYP above the membrane is approximately 3.5 ± 0.9 nm and that it occupies an area of 6.8 ± 0.95 nm^2 in the membrane. Therefore, not only the transmembrane helix of the CYP but also a part of the catalytic domain is partially immersed in the membrane. The maximal height of CYP2C9 above the membrane was ~4 nm in our model and the area of the protein in the membrane was slightly more extensive than 8 nm^2. However, some of the lipid molecules were intercalated between the A′-helix and the F/G loop and they contributed to the enlargement of this area. Taken together, this information suggests that our model fits reasonably well with the available experimental data and is valid for further analyses.

Final Model Description. The protein was stable and relatively rigid in all simulations (see the plot of rmsd in Figure S3). The most flexible parts, according to the calculated root-mean-square fluctuations (RMSF, which can be transformed to B-factors as $B_f = 8/3\pi^2$ RMSF^2), were mostly those of the N-terminal transmembrane helix or the loops which were exposed to water. Interestingly, the F/G loop, flexible in some CYPs,[6] was more rigid in CYP2C9 on the membrane. The central part of the

dx.doi.org/10.1021/jp204488j |J. Phys. Chem. A XXXX, XXX, 000–000

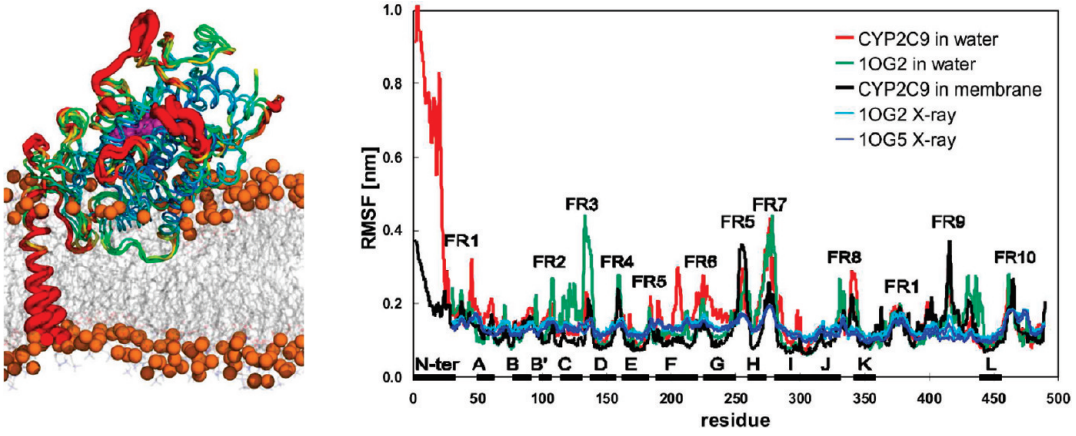


Figure 3. Left panel shows B-factors mapped onto the structure calculated from individual 100 ns long simulations of CYP2C9 in membrane. The B-factors are colored by spectrum from dark blue for rigid regions to red for the most flexible regions. Heme is shown in magenta and membrane lipids are shown in gray with orange spheres for phosphoric atoms. Right panel shows RMSF per residue for simulations of wt CYP2C9 in membrane (black line), wt CYP2C9 in water (green line), CYP2C9 based on crystal structure 1OG2 (without N-terminal anchor) in water (red line), as well as RMSF calculated from the B-factors of crystals structures 1OG2 and 1OG5 (cyan and blue lines, respectively). The figure shows that the most flexible parts of the molecule are the N-terminal anchor (FR1), G/H loop (FR5), K/L loop (FR9), and C-terminus (FR10). Interestingly, the F/G loop, which is flexible in some CYPs, is rather rigid.[6] Also, the central part of the protein close to the catalytic site is relatively rigid. The simulation of CYP2C9 in water shows highly enhanced flexibility of the N-terminal part (FR1) and of the F/G-loop (FR6) in comparison with simulation of CYP2C9 anchored to the membrane. The profiles of RMSF calculated from all simulations have similar trends, which also agree well with published data.[6−8]

protein close to the catalytic site was also rigid (see Figure 3). The flexible N-terminal anchor retained α-helical topology, with a bend at the Arg21 position (see Figure S4 in Supporting Information).

The membrane was rather flexible and it adapted to the presence of the hydrophobic tip of protein composed from the N-terminal transmembrane domain, the A′-helix and the F/G-loop. In the vicinity of the N-terminal anchor, the thickness of the membrane decreased from the initial value of ∼40 Å (the mean distance between two phosphoric groups on the opposite sides of the free DOPC membrane) to ∼33 Å. The adaptation to the protein also resulted in a thickening of the membrane to ∼46 Å in the contact region close to the F/G and K/L loops. After 10 ns of equilibration in the simulation, there were no other significant changes in the membrane topology.

To function, CYPs require CYP reductase (CPR) as a source of electrons (cf. Figure 1). CYP and CPR are both membrane-bound proteins[55] and the membrane facilitates contact between the proteins. The final model should therefore enable CYP/CPR complex formation. The positions of interacting residues were detected experimentally by mutational study on the affinity of CYP2B4 toward CPR.[56] The corresponding residues in CYP2C9 are K121, R125, R132, F134, M136, K138, K432, and G442 (see Figure S1). All of these residues on the surface of the CYP2C9 are accessible to water and available for interaction with CPR in our model (Figure 4).

Effects of the Membrane on the CYP Structure and Flexibility. The RMSF profile calculated from the simulation of CYP2C9 based on 1OG2 X-ray structure agrees well with published results and displays the previously identified common flexible regions.[6,8] The simulation of CYP2C9 with the attached N-terminal anchor displays a similar RMSF pattern besides the extremely flexible N-terminal anchor. The simulation of CYP2C9 in the membrane also shows the same RMSF pattern as simulations of CYP2C9 in water. However, the flexibility of the part of the enzyme in contact with the membrane (specifically the N-terminal part and the C/D and F/G loops) is somewhat decreased. The diffusion of CYP2C9 is also slowed by the presence of the membrane (see Table S1 in Supporting Information).

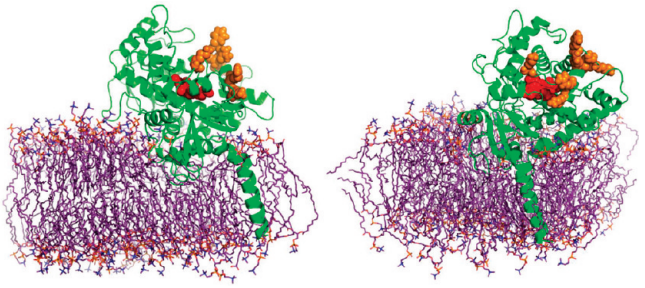


Figure 4. Proximal side of CYP2C9 anchored to DOPC membrane is accessible to water and CYP reductase (CPR). These two views of CYP2C9 show amino acids responsible for binding to CPR (K121, R125, R132, F134, M136, K138, K432, and G442 shown in orange and heme shown in red). These amino acids are present on the CYP proximal side surface.[56] CPR-binding amino acids are accessible to water in our model.

DISCUSSION

Comparison to Empirical Membrane Topology Models from the Literature. Previous studies have attempted to predict CYP membrane topology on the basis of the surface hydrophobicity and available experimental data; our model of wt CYP2C9 is the first atomistic model of the membrane topology. Figure 5 shows a comparison of these predictions and their variabilities. The most similar positioning to our atomistic model provided by the PPM 2.0 computational approach[23] is applied to the complete wt CYP2C9 structure with the N-terminal anchor (cf. white and orange membrane planes in Figure 5). The main difference is in the change in the membrane thickness in the vicinity of the CYP, which is not accounted for in PPM calculations.

dx.doi.org/10.1021/jp204488j |J. Phys. Chem. A XXXX, XXX, 000–000

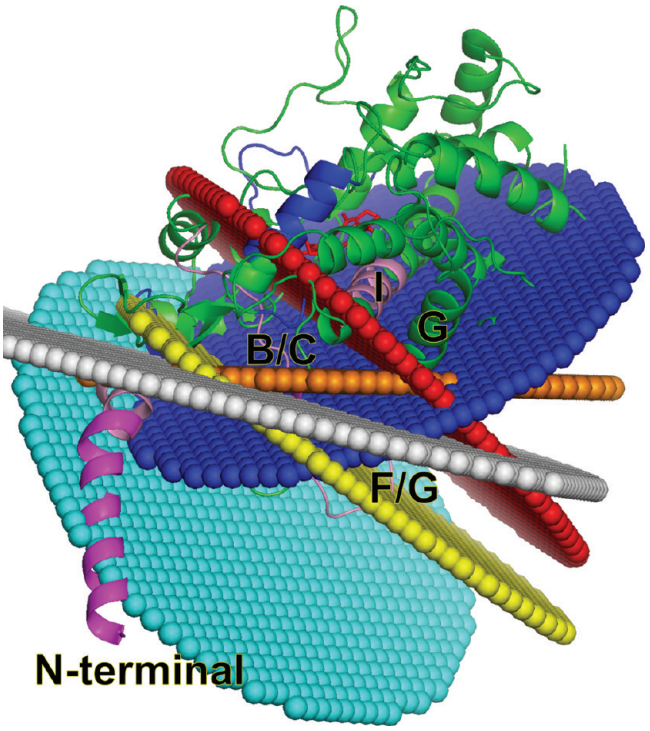


Figure 5. Comparison of models of embedding the CYP2C9 protein into the membrane. Planes of membrane upper layer (approximately at the glycerol level) are colored as follows: This work (orange), Williams et al.[25] (yellow), Wade et al.[57] (cyan), Zhao et al.[59] (red), Lomize et al. PPM 2.0 protein membrane positioning approach for 1OG5 from OPM database http://opm.phar.umich.edu/[23] (blue), and the same PPM 2.0 approach applied to on our wt CYP2C9 model with the N-terminal anchor present (white). As the OPM database shows the planes of positions for the hydrophobic slab, PPM planes were shifted by 10 Å upward to be consistent with other planes at the glycerol level.

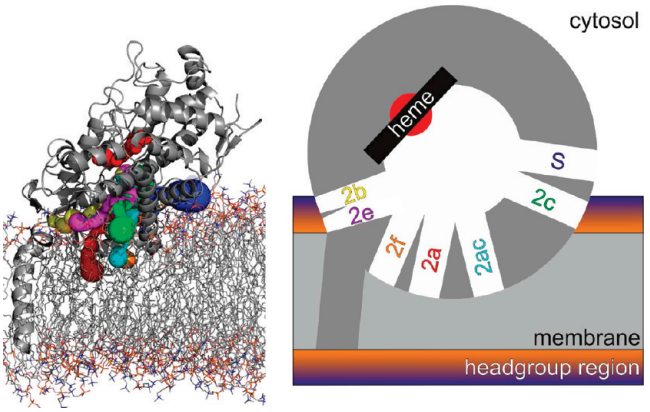


Figure 6. Positions of access/egress channels within membrane anchored CYP2C9. Left panel shows the model of CYP2C9 in DOPC membrane. Protein is shown in dark gray cartoon representation. Heme is shown by red balls. DOPC membrane is shown in wires. The most widely open active site access channel is solvent channel (blue). The mouth opening of the solvent channel is slightly above the membrane-water interface. Channels from 2x family are shown in an order following their opening during the simulation (see Figure S5 in Supporting Information) as follows: 2c (green), 2ac (cyan), 2b (yellow), 2e (magenta), 2f (orange), and 2a (red). All of the openings of these channels are pointing into the membrane; however, the 2b and 2c channel openings are pointing into the headgroup region interface. Access channels were analyzed by MOLE[39] with Pymol 0.99rc6.[33] Right panel shows schematical description of the positions of above-mentioned channels of our model.

If the transmembrane segment is not present (1OG2), the prediction incorrectly exposes residues around the N-terminus to the solvent (Figure 5, blue). The other topologies adopted from the literature indicate that the N-terminus should be inside the membrane, but they show either too shallow immersion of CYP, as shown in Figure 5, in cyan,[57] in yellow,[58] or deeper immersion.[59]

Effects on Access and Egress Channels of CYP2C9. An intriguing question of CYP biochemistry is how substrates enter and products leave the active site, which is buried deep inside the CYP structure. The active site is connected to the outside environment by a network of access/egress channels, which may enable passage of a substrate to, and a product from, the active site.[5,57,60−62] However, there are at least two important and unanswered questions. Does the CYP molecule undergo a large conformational change during substrate binding and product release and which channels enable access and which enable egress?

Our MD simulations do not show any large conformational changes in the CYP2C9 anchored to the membrane. CYP2C9 retains its topology and RMSF indicates decreased fluctuations of CYP2C9 in the membrane in respect to CYP2C9 in water. This implies that common structural features and variations of CYP, derived from X-ray structures of engineered CYPs,[7] are also valid for wild-type CYPs. However, some larger structural changes (namely, in the active site upon binding of larger substrates[63]) of membrane-anchored CYP cannot be ruled out because of the above-mentioned limitation of our MD simulations.

The access/egress channels were analyzed by MOLE software.[39] We adopted the nomenclature of access/egress channels introduced by Wade and co-workers.[5] The X-ray structure of 1OG2 indicates the possible active site access paths, the solvent channel and the 2c channel (between the B′ and G helices). During the MD simulations of CYP2C9 in water the solvent channel closes and the 2c channel opens wider. The analysis of the final model of CYP2C9 in the membrane identified the solvent channel as the widest active site access channel (for further analysis, see Figure S5 in Supporting Information). Solvent channel points slightly above the interface between the membrane and water. Channels from the 2x family (which are mostly closed having bottleneck radii below 1.5 Å) are pointing into the membrane; whereas 2b and 2c channel openings are facing the interface (see Figure 6).

Effects on Drug Disposition and Metabolism. To shed some light to the question how substrates enter and products leave the CYP buried active site, we investigated the position (penetration depth) of a typical CYP2C9 substrate, ibuprofen and its metabolite (3-hydroxyibuprofen), in the membrane. The positions were identified as minima on the free energy profiles along the *z*-axis, which was perpendicular to the membrane (see Figure 7). Potential of mean force simulations on the DOPC membrane/water boundary showed that the free energy profiles of both molecules have different positions of minima (see Table 1). While negatively charged ibuprofen prefers to be inside the membrane perpendicular to the membrane plane with its carboxyl group attached to the positively charged choline group in the DOPC bead, the neutral ibuprofen prefers the hydrophobic membrane interior (see Figure 7). The position of ibuprofen agrees with previous work by Boggara and Krishnamoorti.[64] In contrast, the hydroxylated metabolite is only weakly bound to the

dx.doi.org/10.1021/jp204488j |*J. Phys. Chem. A* XXXX, XXX, 000–000

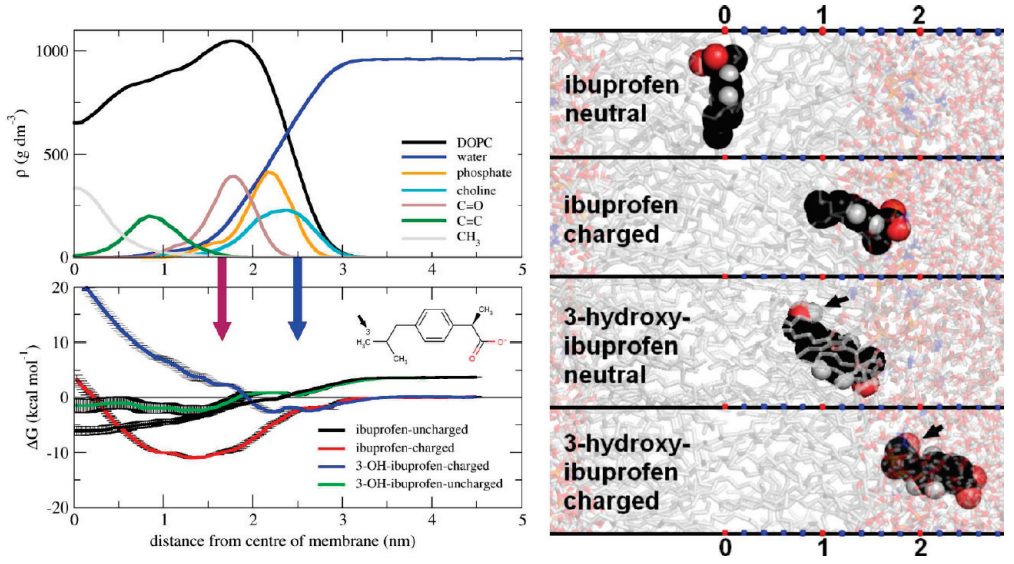


Figure 7. Positions of CYP2C9 substrate ibuprofen and its metabolite on membrane. The upper left part shows partial densities of selected groups on the DOPC membrane/water boundary, while the lower part shows free energy profiles for ibuprofen and 3-hydroxyibuprofen. Free energies were calculated by the potential of mean force imposed on the distance in the *z*-axis between the center-of-mass of the drug and membrane. Zero free energy was arbitrarily selected in the distance of 3.5 nm from the center of the membrane (i.e., in water). The free energy profile for uncharged ibuprofen was shifted by the free energy calculated from the ibuprofen acidity constant in water and for pH 7 by 3.5 kcal/mol (14.6 kJ/mol). The scheme of protonated ibuprofen highlighted with an arrow shows the oxygenation site at position 3 and approximate positions of the opening for 2x and solvent channels are shown by magenta and blues arrows, respectively. The structures on the right side were selected from the window taken from umbrella sampling with minimal free energy. Note that while neutral ibuprofen is in the center of the membrane without any contact with the water environment, both charged molecules are in contact with water molecules mediated by their charged carboxylic group (shown in red). The more polar 3-hydroxyibuprofen is however more exposed to water by its hydroxyl group as shown by black arrow. The figure was prepared in Pymol 0.99rc6.[33].

Table 1. Free Energy Minima for Molecules in DOPC Membrane[a]

	minimum of free energy	
molecule	distance from center of membrane [nm]	ΔG [kcal/mol] (kJ/mol)
ibuprofen (neutral, charge 0)	0.1	−9.6 ± 0.7 (−40.2 ± 2.9)
ibuprofen (charge −1)[b]	1.4	−11.0 ± 0.1 (−46.0 ± 0.4)
3-hydroxyibuprofen (neutral, charge 0)	1.4	−2.4 ± 0.8 (−10.0 ± 3.3)
3-hydroxyibuprofen (charge −1)[b]	2.2	−2.7 ± 0.1 (−11.3 ± 0.4)

[a] Data were calculated by the potential of mean force simulations on DOPC/water membrane. The distance of the carboxylic group to the center-of-mass of the ibuprofen is approximately 5 Å. See Figure 7 for structures corresponding to respective minima. [b] Preferential form in water at pH = 7 (ibuprofen pK_a = 4.44, the same value was also used as a guess for 3-hydroxyibuprofen).[67]

membrane surface and its incorporation into the membrane is energetically highly disadvantageous (see Figure 7). The incorporation of small and rigid molecules (such as cholesterol) is connected with changes in the membrane flexibility and phase transitions.[65] In this respect, it can be hypothesized that CYPs may "sweep out" the membrane by increasing the solubility of low-polar xenobiotics in water, for example, ibuprofen.

It should be noted that there is a significant energy penalty for charged forms for crossing the membrane, whereas uncharged forms prefer the central part of the membrane. The penalty is connected with the introduction of a water-filled hole in the membrane structure, which is filled by the water molecules attached to the charged carboxylic group of ibuprofen. Therefore, a possible model of transfer of ionizable drug across the membrane consists of these steps: (i) the drug is attached into the membrane interior with part of the molecule, (ii) the drug is neutralized, (iii) the drug is transferred across the membrane center, and (iv) the drug is recharged on the other side of membrane. The height of the energy barrier across the membrane and depth of free energy minima define the drug penetration across the membranes and drug partitioning between water and membrane.[66] Both of these properties contribute to determining the drug disposition in the organism.

The minimum of free energy for ibuprofen is located at the position bellow the phosphate groups. This position corresponds approximately to the entrance to the 2x channels. On the other hand, the minimum for 3-hydroxyibuprofen is located approximately at the position of the solvent channel mouth opening (Figure 7). Hence, we can further hypothesize that the 2x channels are involved in the substrate access into the CYP active site and the solvent channel is used for the metabolite egress. This also coincides with information from studies that show that the solvent channel is lined by polar residues enabling active site solvation[8,27] and the 2x channels are in this respect considerably less polar. On the other hand, we cannot rule out the possibility that other (e.g., more polar) substrates can enter the active site via solvent channel and also that some substrates can share the same path for substrate access and product egress.

dx.doi.org/10.1021/jp204488j |*J. Phys. Chem. A* XXXX, XXX, 000–000

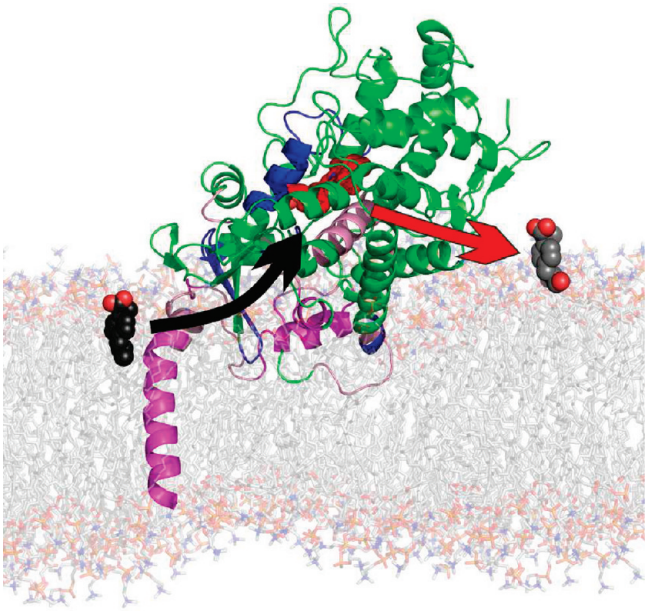


Figure 8. Proposed mechanism of CYP2C9 substrate access and product release. Ibuprofen is accumulated in the membrane, from which it can be effectively "hoovered" by access channels leading to the CYP active site, and following the enzymatic reaction the resulting metabolite may be released via egress channels leading to the cytosol.

CONCLUSIONS

We have constructed an atomic model of CYP2C9 in DOPC membrane; its features were cross-validated with available experimental data from epitope screening, atomic force microscopy measurements, mutagenesis, and fluorescence studies. This model is also consistent with homodimerization and interaction of CYP2C9 with the CYP reductase. The transmembrane helix, 1−1 sheet, A′, G′ helices are located within the membrane, while A, B′, F′, G helices are located on the membrane surface. The membrane anchor is mostly α-helical and is kinked in the membrane at position R21.

The entrance (mouth opening) of the 2x channels points toward the hydrophobic part of the membrane and position of their openings corresponds to the free energy minimum (∼1 nm deep in the membrane) of ibuprofen, which is a CYP2C9 substrate. The mouth opening of solvent channel is positioned in the membrane/solvent interface, corresponding to the minima of free energy profile of the ibuprofen oxidation product on the membrane/water environment (see Figure 8). Therefore, 2x channels (possibly with exception of 2b, 2c, and 2e channels) may be involved in substrate binding (at least of low polar substrates), and the solvent channel is likely to be involved in the product release. The proposed mechanism of drug-binding may be shared among all membrane-bound microsomal CYPs.

ASSOCIATED CONTENT

Supporting Information. Multiple alignment with aggregated information about membrane insertion, table with diffusion coefficient changes, and other figures. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author
*E-mail: michal.otyepka@upol.cz.

ACKNOWLEDGMENT

Analysis of channel opening was kindly provided by Vlad Cojocaru from Department of Cell and Developmental Biology, Max Planck Institute for Molecular Biomedicine in Muenster. K. B. would like to thank Kateřina Holá for her help with schematic figures. Support through GACR grants 303/09/1001 and 203/09/H046 and by Student Project PrF_2011_020 of Palacky University are gratefully acknowledged. This work was supported by the Operational Program Research and Development for Innovations - European Social Fund (CZ.1.05/2.1.00/03.0058, CZ.1.05/2.1.00/01.0030 and CZ.1.07/2.3.00/20.0017).

REFERENCES

(1) Ortiz de Montellano, P. R. *Cytochrome P450: Structure, Mechanism, and Biochemistry*, 3rd ed.; Kluwer Academic/Plenum Publishers: New York, 2005.
(2) Anzenbacher, P.; Anzenbacherova, E. *Cell. Mol. Life Sci.* **2001**, *58* (5−6), 737–47.
(3) Coon, M. J. *Annu. Rev. Pharmacol. Toxicol.* **2005**, *45*, 1–25.
(4) Graham, S. E.; Peterson, J. A. *Arch. Biochem. Biophys.* **1999**, *369* (1), 24–9.
(5) Cojocaru, V.; Winn, P. J.; Wade, R. C. *Biochim. Biophys. Acta* **2007**, *1770* (3), 390–401.
(6) Hendrychova, T.; Anzenbacherova, E.; Hudecek, J.; Skopalik, J.; Lange, R.; Hildebrandt, P.; Otyepka, M.; Anzenbacher, P. *Biochim. Biophys. Acta* **2011**, *1814* (1), 58–68.
(7) Otyepka, M.; Skopalik, J.; Anzenbacherova, E.; Anzenbacher, P. *Biochim. Biophys. Acta* **2007**, *1770* (3), 376–389.
(8) Skopalik, J.; Anzenbacher, P.; Otyepka, M. *J. Phys. Chem. B* **2008**, *112* (27), 8165–73.
(9) Pochapsky, T. C.; Kazanis, S.; Dang, M. *Antioxid. Redox Signaling* **2010**, *13* (8), 1273–96.
(10) Fishelovitch, D.; Hazan, C.; Shaik, S.; Wolfson, H. J.; Nussinov, R. *J. Am. Chem. Soc.* **2007**, *129* (6), 1602–11.
(11) Sakaguchi, M.; Mihara, K.; Sato, R. *EMBO J.* **1987**, *6* (8), 2425–31.
(12) Black, S. D. *FASEB J.* **1992**, *6* (2), 680–5.
(13) Black, S. D.; Martin, S. T.; Smith, C. A. *Biochemistry* **1994**, *33* (22), 6945–51.
(14) Szklarz, G. D.; Halpert, J. R. *Life Sci.* **1997**, *61* (26), 2507–20.
(15) Cosme, J.; Johnson, E. F. *J. Biol. Chem.* **2000**, *275* (4), 2545–53.
(16) Nakayama, K.; Puchkaev, A.; Pikuleva, I. A. *J. Biol. Chem.* **2001**, *276* (33), 31459–65.
(17) Fernando, H.; Halpert, J. R.; Davydov, D. R. *Biochemistry* **2006**, *45* (13), 4199–209.
(18) Lepesheva, G. I.; Seliskar, M.; Knutson, C. G.; Stourman, N. V.; Rozman, D.; Waterman, M. R. *Arch. Biochem. Biophys.* **2007**, *464* (2), 221–7.
(19) Bayburt, T. H.; Sligar, S. G. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99* (10), 6725–30.
(20) Shank-Retzlaff, M. L.; Raner, G. M.; Coon, M. J.; Sligar, S. G. *Arch. Biochem. Biophys.* **1998**, *359* (1), 82–8.
(21) Hudecek, J.; Anzenbacher, P. *Biochim. Biophys. Acta* **1988**, *955* (3), 361–70.
(22) Friedman, F. K.; Robinson, R. C.; Dai, R. *Front. Biosci.* **2004**, *9*, 2796–806.
(23) Lomize, A. L.; Pogozheva, I. D.; Lomize, M. A.; Mosberg, H. I. *Protein Sci.* **2006**, *15* (6), 1318–33.
(24) Scott, E. E.; He, Y. Q.; Halpert, J. R. *Chem. Res. Toxicol.* **2002**, *15* (11), 1407–13.

(25) Williams, P. A.; Cosme, J.; Sridhar, V.; Johnson, E. F.; McRee, D. E. *Mol. Cell* **2000**, *5* (1), 121–131.
(26) Conner, K. P.; Woods, C. M.; Atkins, W. M. *Arch. Biochem. Biophys.* **2011**, *507* (1), 56–65.
(27) Rydberg, P.; Rod, T. H.; Olsen, L.; Ryde, U. *J. Phys. Chem. B* **2007**, *111* (19), 5445–5457.
(28) Evans, W. E.; Relling, M. V. *Science* **1999**, *286* (5439), 487–91.
(29) Williams, P. A.; Cosme, J.; Ward, A.; Angove, H. C.; Matak Vinkovic, D.; Jhoti, H. *Nature* **2003**, *424* (6947), 464–8.
(30) Larkin, M. A.; Blackshields, G.; Brown, N. P.; Chenna, R.; McGettigan, P. A.; McWilliam, H.; Valentin, F.; Wallace, I. M.; Wilm, A.; Lopez, R.; Thompson, J. D.; Gibson, T. J.; Higgins, D. G. *Bioinformatics* **2007**, *23* (21), 2947–8.
(31) Waterhouse, A. M.; Procter, J. B.; Martin, D. M.; Clamp, M.; Barton, G. J. *Bioinformatics* **2009**, *25* (9), 1189–91.
(32) von Wachenfeldt, C.; Johnson, E. F. Structures of Eukaryotic Cytochrome P450 Enzymes - Membrane Topology. In *Cytochrome P450: Structure, Mechanism and Biochemistry*, 2nd ed.; Plenum Press: New York, 1995; pp 183−223.
(33) DeLano, W. L. *The PyMOL Molecular Graphics System*, 0.99rc6; DeLano Scientific: Palo Alto, CA, 2002.
(34) van Meer, G.; Voelker, D. R.; Feigenson, G. W. *Nat. Rev. Mol. Cell Biol.* **2008**, *9* (2), 112–24.
(35) Siu, S. W.; Vacha, R.; Jungwirth, P.; Bockmann, R. A. *J. Chem. Phys.* **2008**, *128* (12), 125103.
(36) Oostenbrink, C.; Villa, A.; Mark, A. E.; van Gunsteren, W. F. *J. Comput. Chem.* **2004**, *25* (13), 1656–76.
(37) Berger, O.; Edholm, O.; Jahnig, F. *Biophys. J.* **1997**, *72* (5), 2002–13.
(38) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. *J. Chem. Theory Comput.* **2008**, *4* (3), 435–447.
(39) Petrek, M.; Kosinova, P.; Koca, J.; Otyepka, M. *Structure* **2007**, *15* (11), 1357–63.
(40) Petrek, M.; Otyepka, M.; Banas, P.; Kosinova, P.; Koca, J.; Damborsky, J. *BMC Bioinf.* **2006**, *7*, 316.
(41) Porubsky, P. R.; Battaile, K. P.; Scott, E. E. *J. Biol. Chem.* **2010**, *285* (29), 22282–90.
(42) Schuttelkopf, A. W.; van Aalten, D. M. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004**, *60* (Pt 8), 1355–63.
(43) Hub, J. S.; de Groot, B. L.; van der Spoel, D. *J. Chem. Theory Comput.* **2010**, *6* (12), 3713–3720.
(44) Ruan, K. H.; So, S. P.; Zheng, W.; Wu, J.; Li, D.; Kung, J. *Biochem. J.* **2002**, *368* (Pt 3), 721–8.
(45) Bond, P. J.; Wee, C. L.; Sansom, M. S. *Biochemistry* **2008**, *47* (43), 11321–31.
(46) Ohta, Y.; Kawato, S.; Tagashira, H.; Takemori, S.; Kominami, S. *Biochemistry* **1992**, *31* (50), 12680–7.
(47) Ozalp, C.; Szczesna-Skorupa, E.; Kemper, B. *Biochemistry* **2006**, *45* (14), 4629–37.
(48) Headlam, M. J.; Wilce, M. C.; Tuckey, R. C. *Biochim. Biophys. Acta* **2003**, *1617* (1−2), 96–108.
(49) Pikuleva, I. A.; Mast, N.; Liao, W. L.; Turko, I. V. *Lipids* **2008**, *43* (12), 1127–32.
(50) Mast, N.; Liao, W. L.; Pikuleva, I. A.; Turko, I. V. *Arch. Biochem. Biophys.* **2009**, *483* (1), 81–9.
(51) Schoch, G. A.; Yano, J. K.; Wester, M. R.; Griffin, K. J.; Stout, C. D.; Johnson, E. F. *J. Biol. Chem.* **2004**, *279* (10), 9497–9503.
(52) Hu, G.; Johnson, E. F.; Kemper, B. *Drug Metab. Dispos.* **2010**, *38* (11), 1976–83.
(53) Szczesna-Skorupa, E.; Mallah, B.; Kemper, B. *J. Biol. Chem.* **2003**, *278* (33), 31269–76.
(54) Nussio, M. R.; Voelcker, N. H.; Miners, J. O.; Lewis, B. C.; Sykes, M. J.; Shapter, J. G. *Chem. Phys. Lipids* **2010**, *163* (2), 182–9.
(55) Kida, Y.; Ohgiya, S.; Mihara, K.; Sakaguchi, M. *Arch. Biochem. Biophys.* **1998**, *351* (2), 175–9.
(56) Bridges, A.; Gruenke, L.; Chang, Y. T.; Vakser, I. A.; Loew, G.; Waskell, L. *J. Biol. Chem.* **1998**, *273* (27), 17036–49.
(57) Wade, R. C.; Motiejunas, D.; Schleinkofer, K.; Sudarko; Winn, P. J.; Banerjee, A.; Kaniakin, A.; Jung, C. *Biochim. Biophys. Acta* **2005**, *1754* (1−2), 239–244.
(58) Williams, P. A.; Cosme, J.; Sridhar, V.; Johnson, E. F.; McRee, D. E. *J. Inorg. Biochem.* **2000**, *81* (3), 183–90.
(59) Zhao, Y.; White, M. A.; Muralidhara, B. K.; Sun, L.; Halpert, J. R.; Stout, C. D. *J. Biol. Chem.* **2006**, *281* (9), 5973–81.
(60) Schleinkofer, K.; , Sudarko; Winn, P. J.; Ludemann, S. K.; Wade, R. C. *EMBO Rep.* **2005**, *6* (6), 584–589.
(61) Li, W. H.; Shen, J.; Liu, G. X.; Tang, Y.; Hoshino, T. *Proteins: Struct., Funct., Bioinf.* **2011**, *79* (1), 271–281.
(62) Fishelovitch, D.; Shaik, S.; Wolfson, H. J.; Nussinov, R. *J. Phys. Chem. B* **2009**, *113* (39), 13018–25.
(63) Ekroos, M.; Sjogren, T. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103* (37), 13682–7.
(64) Boggara, M. B.; Krishnamoorti, R. *Biophys. J.* **2010**, *98* (4), 586–95.
(65) McMullen, T. P. W.; Lewis, R. N. A. H.; McElhaney, R. N. *Curr. Opin. Colloid Interface Sci.* **2004**, *8* (6), 459–468.
(66) Orsi, M.; Essex, J. W. *Soft Matter* **2010**, *6* (16), 3797–3808.
(67) Watkinson, R. M.; Herkenne, C.; Guy, R. H.; Hadgraft, J.; Oliveira, G.; Lane, M. E. *Skin Pharmacol. Physiol.* **2009**, *22* (1), 15–21.

Supplemental information

Membrane Position of Ibuprofen Agrees with Suggested Access Path Entrance to Cytochrome P450 2C9 Active Site

Karel Berka,[a] Tereza Hendrychová,[a] Pavel Anzenbacher,[b] Michal Otyepka*[a]
[a] Regional Centre of Advanced Technologies and Materials, Department of Physical Chemistry, Faculty of Science, Palacky University, 17. listopadu 12, 771 46 Olomouc, Czech Republic
[b] Department of Pharmacology, Faculty of Medicine and Dentistry, Palacky University, Olomouc, Czech Republic

updated 4. July 2011

Figure S1. Multiple alignment of the human CYP2C9 with sequences used in the experimental topology studies.

Multiple alignment of the human wt CYP2C9 (swissprot accession number P11712), crystal structure of the engineered CYP2C9 (pdbid: 1OG2 and 1OG5) and with rat CYP2B1 (P00176) and rabbit CYP2B4 (P00178) with annotated membrane topology and mutations. Multiple alignment was calculated by ClustalW2[1] and visualized with Jalview 2.6[2]. Red color denotes mutations used for preparation of the engineered crystal structure. Green color shows the amino acids which are responsible for the interactions with cytochrome P450 reductase[3]. Annotations for membrane topology are colored as follows: blue for the sequence accessible, violet for the sequence inaccessible and pink for regions where the data are mixed. Data were taken from Refs. [4-8].

Table ST1 – The diffusion coefficient changes upon membrane insertion.

The diffusion coefficients (D) were calculated with Einstein-Stokes equation from the linear part of the mean square displacement of protein within the simulations. The diffusion coefficients given by this study seem to be too big in comparison with experimental data for several reasons: (i) the use of homogenous DOPC membrane, and (ii) a simplified united-atom Berger force field model of the DOPC/SPC membrane/water environment. Even though the calculated diffusion is by order of magnitude bigger than experimental data, we still can conclude that the diffusion of the solubilized CYP in water is quicker than the diffusion of the CYP bound on the membrane.

Diffusion Type and System	Source	D 10^{-7} cm^2/s
Total, model in water	This study	7.0 ± 0.26
Total, model in DOPC membrane	This study	1.6 ± 0.14
Lateral, model in DOPC membrane	This study	2.1 ± 0.23
Lateral, WALP23 †	Ref. [9]	0.5 ± 0.1
Lateral, CYP2C2 in ER membrane	Ref. [10]	$5.8 \cdot 10^{-3} \pm 0.2 \cdot 10^{-3}$

†artificial helical TM segment in DOPC/DOPG 3:1 membrane.

Figure S2 – Comparison between initial (green) and final (blue) model of wt CYP2C9.

Panel A and B show the protein and membrane changes upon simulation, respectively. The protein is shown in cartoon representation and phosphorous atoms from the membrane are shown as spheres. While the protein core did not changed significantly during the simulation, the membrane topology and parts of protein responsible for the interaction with the membrane show slight rearrangement as discussed in paper. The figure was prepared in Pymol 0.99rc6[11].

A B

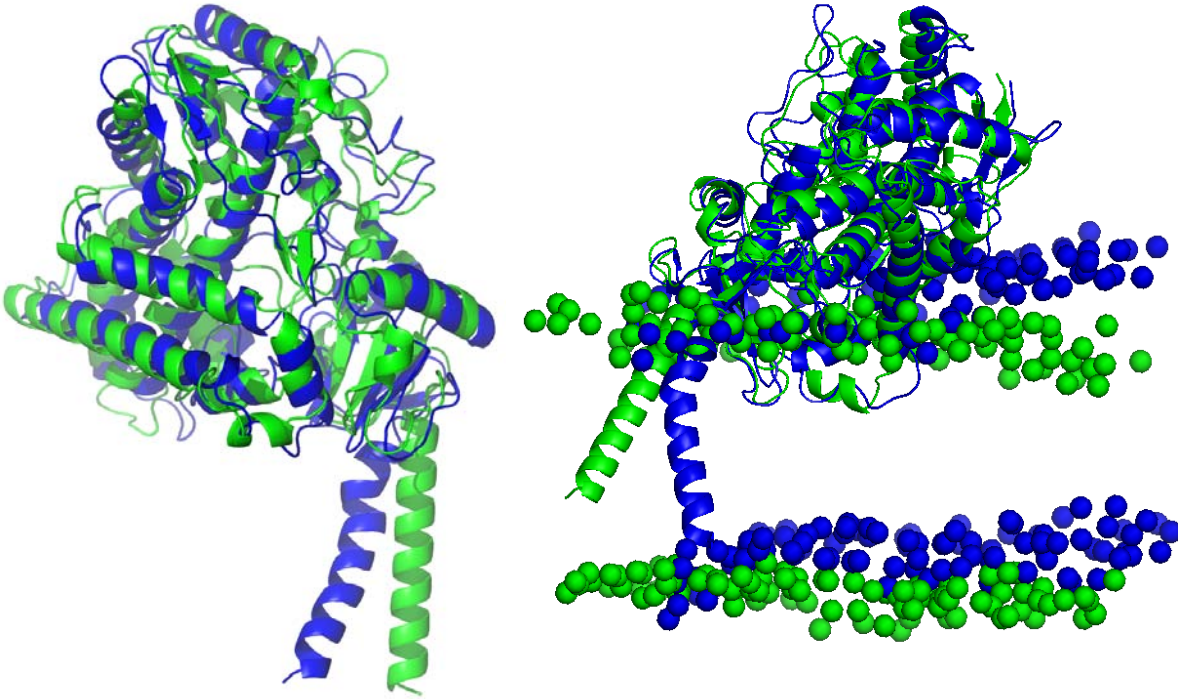


Figure S3 – The evolution of RMSD of CYP2C9 on the membrane in three individual simulations used in selection of the appropriate model.

The simulation was triplicated after first 10 ns. The simulations were then equilibrated after approximately 20 ns.

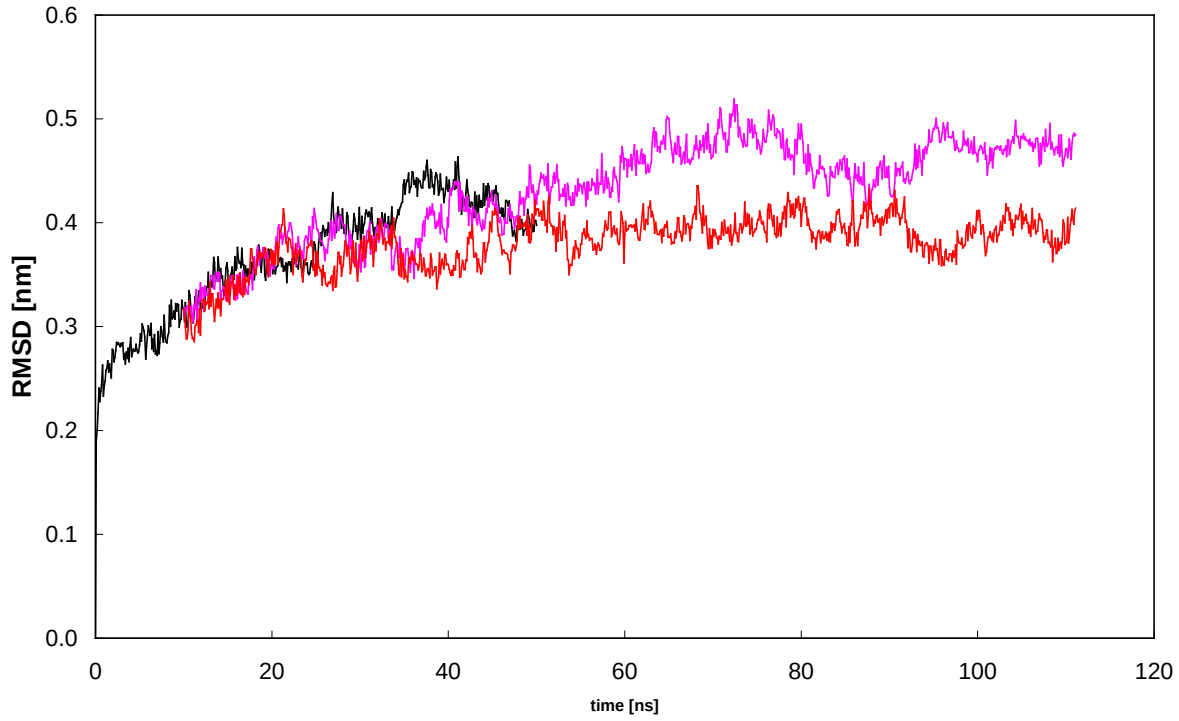


Figure S4 – Detailed structure of the transmembrane helical segment of CYP2C9.

The transmembrane helix is shown in magenta. The N-terminal part is not perpendicular to the membrane, nor is penetrating beyond charged groups in the luminal part of membrane. The helix is kinked in the Arg21 position. The figure was prepared in Pymol 0.99rc6[11].

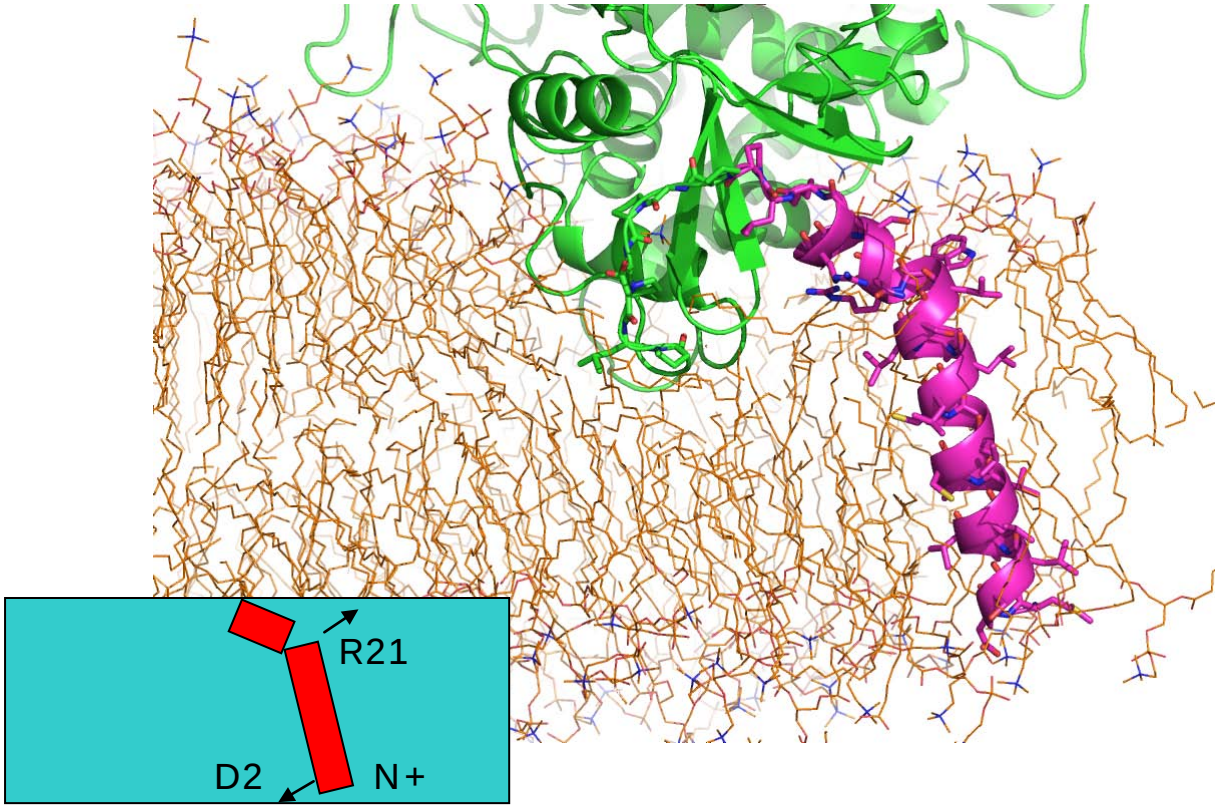


Figure S5 – Channel opening and closing during membrane simulations of CYP2C9.

Channel opening analysis provided by Vlad Cojocaru on our two independent simulations. The most open channel is solvent channel followed to some extent by channels 2c and 2ac as seen in coloring in the picture.

.

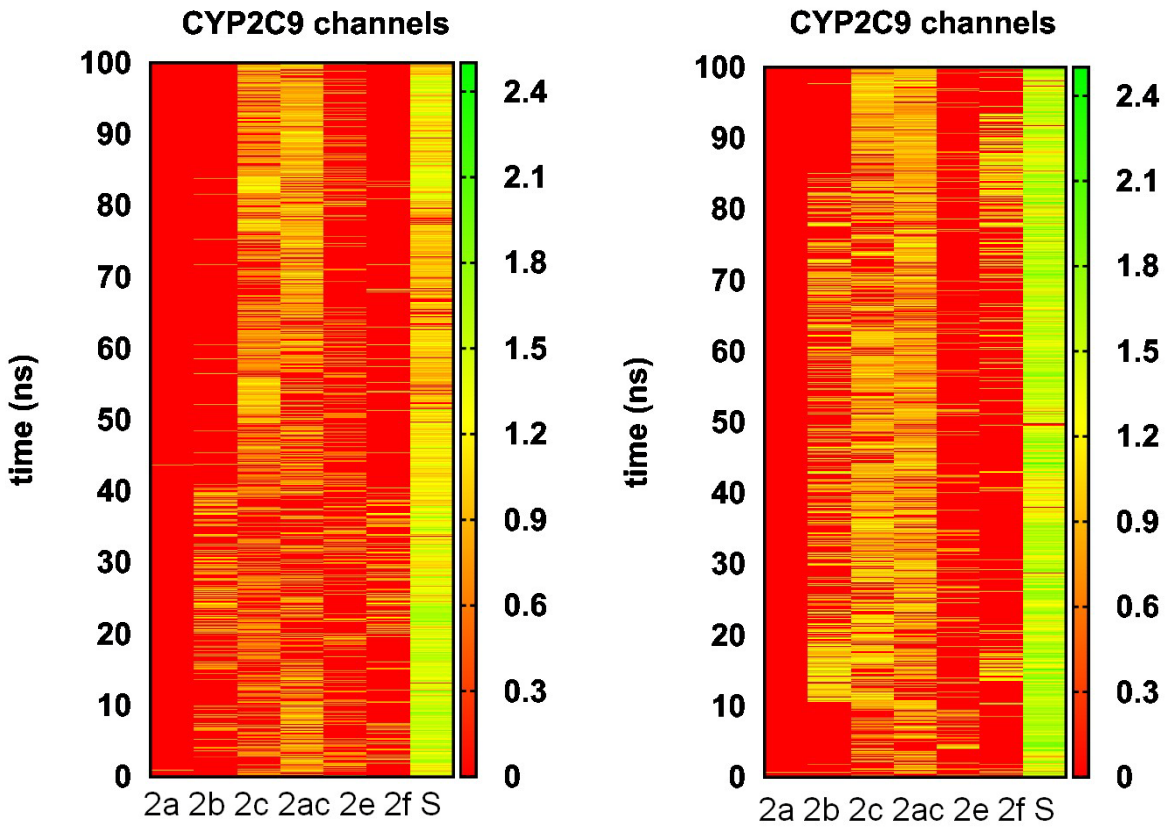
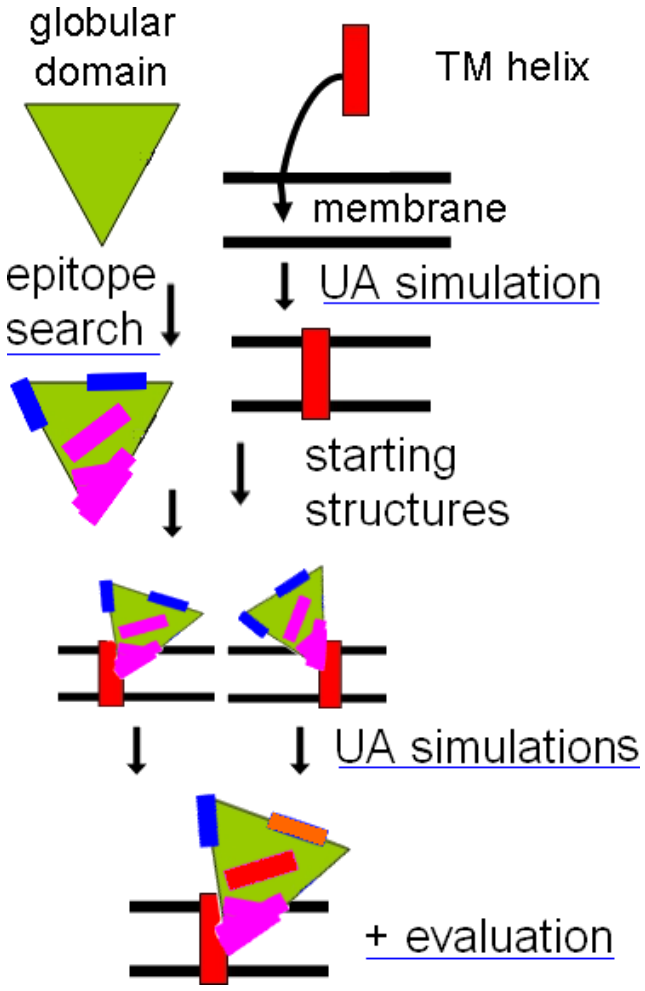


Figure S6 – Schematics of model preparation.

First, position of transmembrane helix in membrane has been evaluated by short 10 ns united atom simulation. The transmembrane helix was shown to be almost perpendicular to the membrane normal. Secondly, the starting positions of the globular domain were selected upon epitope labeling and the globular domain was merged to the transmembrane helix. After that three parallel simulations were produced with united atom force field for total time of 250 ns. Final model was then compared to all known experimental data as shown in the section *Final model cross-validation* in the Results.



References

(1) Larkin, M. A.; Blackshields, G.; Brown, N. P.; Chenna, R.; McGettigan, P. A.; McWilliam, H.; Valentin, F.; Wallace, I. M.; Wilm, A.; Lopez, R.; Thompson, J. D.; Gibson, T. J.; Higgins, D. G. *Bioinformatics*. **2007**, *23*, 2947-8.

(2) Waterhouse, A. M.; Procter, J. B.; Martin, D. M. a; Clamp, M.; Barton, G. J. *Bioinformatics*. **2009**, *25*, 1189-91.

(3) Bridges, A.; Gruenke, L.; Chang, Y. T.; Vakser, I. A.; Loew, G.; Waskell, L. *J. Biol. Chem.* **1998**, *273*, 17036-49.

(4) Black, S. D.; Martin, S. T.; Smith, C. A. *Biochemistry*. **1994**, *33*, 6945-51.

(5) Black, S. D. *The FASEB Journal*. **1992**, *6*, 680-685.

(6) Wachenfeldt, C. von; Johnson, E. F. In *Cytochrome P450: Structure, Mechanism, and Biochemistry*; Plenum Press: New York, 1995; p. 183–244.

(7) Cosme, J.; Johnson, E. F. *J. Biol. Chem.* **2000**, *275*, 2545-2553.

(8) Williams, P. A.; Cosme, J.; Ward, A.; Angove, H. C.; Matak Vinković, D.; Jhoti, H. *Nature*. **2003**, *424*, 464-8.

(9) Ramadurai, S.; Holt, A.; Krasnikov, V.; Bogaart, G. van den; Killian, J. A.; Poolman, B. *J. Am. Chem. Soc.* **2009**, *131*, 12650-6.

(10) Szczesna-Skorupa, E.; Chen, C. D.; Rogers, S.; Kemper, B. *Proc. Natl. Acad. Sci. USA.* **1998**, *95*, 14793-8.

(11) DeLano, W. L. *The PyMOL Molecular Graphics System*, http://pymol.org, 0.99rc6; DeLano Scientific: Palo Alto, CA, USA, **2002**.

Journal of Inorganic Biochemistry 110 (2012) 46–50

Contents lists available at SciVerse ScienceDirect

Journal of Inorganic Biochemistry

journal homepage: www.elsevier.com/locate/jinorgbio

Binding of quinidine radically increases the stability and decreases the flexibility of the cytochrome P450 2D6 active site

Karel Berka [a], Eva Anzenbacherová [b], Tereza Hendrychová [a], Reinhard Lange [c], Vlastimil Mašek [d], Pavel Anzenbacher [d,*], Michal Otyepka [a,**]

[a] Department of Physical Chemistry, Regional Centre of Advanced Technologies and Materials, Faculty of Science, Palacky University, tr. 17. listopadu 12, 771 46 Olomouc, Czech Republic
[b] Department of Medical Chemistry and Biochemistry, Faculty of Medicine and Dentistry, Palacky University, Hnevotinska 3, 775 15 Olomouc, Czech Republic
[c] INSERM U710, Université Montpellier 2, F-34095 Montpellier cédex 5, France
[d] Department of Pharmacology, Faculty of Medicine and Dentistry, Palacky University, Hnevotinska 3, 775 15 Olomouc, Czech Republic

ARTICLE INFO

Article history:
Received 17 October 2011
Received in revised form 13 January 2012
Accepted 15 February 2012
Available online 22 February 2012

Keywords:
CYP2D6
Quinidine
Molecular dynamics
Flexibility
Cytochrome P450
High pressure

ABSTRACT

Human cytochrome P450 2D6 (CYP2D6) is an enzyme of the CYP superfamily responsible for biotransformation of about 20% of drugs of known metabolism containing a basic nitrogen and a planar aromatic ring. Here, we present a combined experimental and computational study on the compressibility and flexibility of unliganded and quinidine-bound CYP2D6. Experimentally, high-pressure induced Soret band shifts of the enzyme were measured by UV/VIS spectroscopy, while 100 ns all atomic molecular dynamics (MD) simulations in explicit water were used in the computational analysis. We identified sharp differences between ligand-free and quinidine-bound CYP2D6 forms in compressibility, flexibility parameters and active site solvation. While the unliganded CYP2D6 is compressible, quinidine binding significantly rigidifies the CYP2D6 active site. In addition, MD simulations show that quinidine binding results in pronounced reductions in active site flexibility and solvation.

© 2012 Elsevier Inc. All rights reserved.

1. Introduction

In humans, most drugs are metabolized by cytochrome P450 (CYP) enzymes that are mainly present in the microsomal fraction of the liver and other tissues, such as brain, intestine, lung and kidney [1]. Microsomal CYPs are quite promiscuous, but vary substantially in their substrate preferences [2,3]. Key determinants of P450 substrate specificity have been analyzed from several perspectives and various structural features responsible for their promiscuity have been identified. Among other factors, the size and flexibility of the active site [4–9], together with properties of the substrate access and egress channels [10,11], have been shown to strongly influence the enzyme's substrate preferences. Techniques that can provide valuable information about their flexibility include molecular dynamics (MD) and UV/visible (UV/VIS) high pressure spectroscopy [7,8,12].

CYP2D6 is the enzyme responsible for biotransformation of about 20% of drugs of known metabolism that contain a basic nitrogen and a planar aromatic ring, including antiarrythmics (e.g. propafenone), antidepressants (e.g. venlafaxine, fluoxetine), antipsychotics (e.g. haloperidol, clozapine), beta-blockers (e.g. metoprolol, atenolol) and analgetics (e.g. codeine, tramadol) [13]. The rate of metabolism of drugs by CYP2D6 is known to reflect genetic variations of the enzyme; the metabolism of certain drugs is much faster in some patients than others, and even absent in carriers of several alleles of the encoding gene, with adverse effects in some cases [2,3]. To eliminate this variation, and associated therapeutic complications, there is great interest in designing drug candidates that are not metabolized by CYP2D6. Understanding the interactions between CYP2D6 and its ligands could greatly facilitate the rational design of potential drug candidates with such properties, or other desirable features [14].

Quinidine (see Fig. 1) is an antiarrhythmic agent (Pubchem CID: 441074) that is a substrate of another CYP enzyme, CYP3A4 [12]. It also has several typical features of CYP2D6 substrates (a planar aromatic ring and a polar nitrogen) and hence binds strongly to CYP2D6 in a similar fashion to its chiral isomer quinine, a CYP2D6 substrate, yielding a typical type I binding spectrum characteristic for binding of substrates [15]. However, in comparison to quinine, quinidine binds in reverse orientation to the CYP2D6 active site [16] and instead of being degraded it blocks the enzyme's active site and thus is a potent competitive inhibitor of CYP2D6 [17,18].

Recent analysis of the CYP2D6 active site and mutagenesis experiments have shown that two carboxylate groups (E216, D301) play key roles in its substrate recognition [14,19–21], possibly due to their interaction with the basic nitrogen of its substrates during entry or binding to the active site. It has also been suggested that

* Corresponding author.
** Corresponding author. Tel.: +420 585634764; fax: +420 585634761.
E-mail addresses: anzen@seznam.cz (P. Anzenbacher), michal.otyepka@upol.cz (M. Otyepka).

0162-0134/$ – see front matter © 2012 Elsevier Inc. All rights reserved.
doi:10.1016/j.jinorgbio.2012.02.010

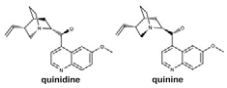


Fig. 1. Structure of quinidine and its diastereomer quinine for comparison.

aromatic residues (F120, F481, F483) may be involved in substrate binding through π–π interactions with substrates' aromatic rings [21]. Strong evidence supporting the putative roles of these residues in the enzyme's substrate recognition includes findings that (in marked contrast to the wild type CYP2D6 enzyme) E216F, E216Q/D301Q and F120A mutants are even capable of metabolizing quinidine [19].

In a previous study [7], we analyzed the flexibility of five human drug-metabolizing CYP enzymes (CYP1A2, 2A6, 2 C9, 2D6 and 3A4) and changes in properties of their active sites imposed by high pressure using both absorption spectroscopy and MD simulations. CYP2D6 showed an intermediate Soret band red shift under pressure, indicating that its active site is quite compressible, and more flexible than that of CYP1A2 and CYP2A6, but less flexible than that of CYP3A4 [7]. In addition, relatively facile denaturation of the native P450 form of CYP2D6 to the P420 form was observed, even at fairly low hydrostatic pressures (100-150 MPa).

Here, we describe changes in the structure, flexibility and compressibility of the CYP2D6 active site induced by quidinine binding observed by high-pressure UV/VIS spectroscopy and in MD simulations. Our results show that quinidine binding profoundly affects the properties of CYP2D6's active site, substantially reducing its flexibility and strongly suppressing the enzyme's denaturation to the P420 form under pressure. The large observed change in compressibility of CYP2D6 indicates that ligand binding to the active site can dramatically change enzyme properties.

2. Materials and methods

2.1. UV/VIS spectroscopy at high hydrostatic pressure

Absorption spectra were obtained with a Cary 3E spectrophotometer (Varian, Palo Alto, CA) equipped with a high-pressure cell. Details of the instrument and the applications of the method in protein generally, and CYP studies specifically, have been previously presented [12,22]. The enzyme was diluted to 2 μM concentration in a 20 mM K/PO_4 buffer (pH 7.4) containing 20% (v/v) glycerol. Quinidine was subsequently added as a 1 mM solution in the K/PO_4 buffer to the reaction mixture (final concentration, 20 μM) then the P450 was reduced by adding solid dithionite and complexed with carbon monoxide. The Soret absorption band position was determined by spectral analysis and its shift with applied pressure was plotted according to the equation $S = S_0 + \alpha P$, where S is the Soret band position at a given pressure P, S_0 is its position at normal pressure and the coefficient α is related to compressibility [23]. The results were processed using LabSpec v. 4.04 (JobinYvon Horiba, Villeneuve d'Ascq, FR) and Sigma Plot v. 8.0 (SPSS, Chicago, IL, USA).

2.2. Molecular dynamics simulations

We conducted MD simulations of CYP2D6 at normal (NP, 0.1 MPa) and high pressure (HP, 300 MPa) with/without quinidine present in the active site. The crystal structure (mean resolution 3.00 Å) of ligand-free CYP2D6 was taken from the PDB database (CYP2D6, PDB ID: 2F9Q [13]), while the structure with bound quinidine (CYP2D6qui) was kindly provided by Prof. Gordon Roberts (Department of Biochemistry, University of Leicester, UK). The positions of the backbone atoms of the missing residues (between P41 and Q52) of the CYP2D6 structure were modeled on those of the analogous residues in CYP2C8 (PDB ID: 1PQ2 [24]); the side chains of these residues were built using the Leap program of the AMBER package. The N-terminal transmembrane domains of CYP2D6 were not included in the simulations.

The force field parameters for non-standard heme residue derived according to Cornell et al. [25] were taken from our earlier articles [7,9]. Two axial ligand positions of the heme iron are occupied by two ligands: (i) cysteinate sulfur as the proximal ligand, and (ii) either a water molecule (CYP2D6) or carbon monoxide (CYP2D6qui) as the distal ligand. All structures were solvated by the TIP3P water box with three (CYP2D6) or six (CYP2D6qui) Na + counter ions to neutralize the systems. Each system was minimized, heated (0.5 ns) and after equilibration (5 ns) we ran a 100 ns long MD production simulation under periodic boundary conditions in the NpT ensemble at normal (NP, 0.1 MPa) and high (HP, 300 MPa) pressures, in both cases at 298.15 K, with 2 fs long integration steps using the AMBER 9.0 package and the parm99 force field [26]. The cut-off for noncovalent interactions was 1 nm and electrostatics were treated by the particle mesh Ewald method. Bonds to hydrogen atoms were constrained using the SHAKE algorithm.

All MD trajectories were analyzed using the ptraj module of the AMBER package. Radius of gyration and temperature B factors were calculated over backbone atoms. The solvation of the active sites of the CYP2D6 enzyme was monitored throughout long simulations at normal and high pressure. Solvation was analyzed using the radial distribution functions (RDF) for the water molecules and backbone atoms around the heme. The space-resolved distribution of water molecules was further evaluated on a grid at 1 $Å^3$ resolution in a similar fashion to the analysis of waters within CYP2D6 presented by Santos et al. [27]. Snapshots taken every 500 ps were superimposed on the backbone atoms and the positions of the water oxygen atoms on the grid were recorded. Sites within the enzyme were identified as water coordination positions if their average water occupancy was at least three times greater than the average values.

3. Results and discussion

3.1. Spectroscopy at high hydrostatic pressure shows strengthening of CYP2D6 with quinidine

Initial absorption spectroscopic analysis revealed that under high hydrostatic pressure the behavior of CYP2D6 is strikingly different in the presence and absence of bound quinidine. As shown in our preceding paper [7], unbound CYP2D6 is prone to denaturation to the inactive form, P420, even at pressures of about 100 MPa. At 250 MPa the P420 form predominates and the native P450 population represents only 15% of the starting value. However, following binding of quinidine formation of the denatured form, P420, is minimal and becomes apparent only at pressures of 250 MPa (Fig. 2).

In the next step, the red shift of the Soret band at 450 nm, characteristic of native P450 enzymes, was monitored to obtain information on the flexibility of the active site [7,19,20,28]. Our previous results indicated that in the absence of bound ligand the active site of CYP2D6 has considerable flexibility, comparable to that of CYP2 C9 [7], substantially more than that of CYP2A6, but less than that of CYP2E1 [28] and CYP3A4 [7] (for values of the compressibility-related coefficients, -α, see Table 1). However, the binding of quinidine reduces the flexibility of the CYP2D6 active site, in line with the stabilization of the CYP2D6 structure against denaturation described in the previous paragraph, and reduced the -α value to about half that of the unliganded protein (Table 1). This finding should not be generalized to other enzyme:substrate pairs as quinidine binds to CYP2D6 in an exceptionally specific and well characterized manner; the effects of other substrates binding to other enzymes may be different [13,18,19].

Overall, the experimental results indicate that quinidine binding stiffens the CYP2D6 structure and/or induces large changes in heme

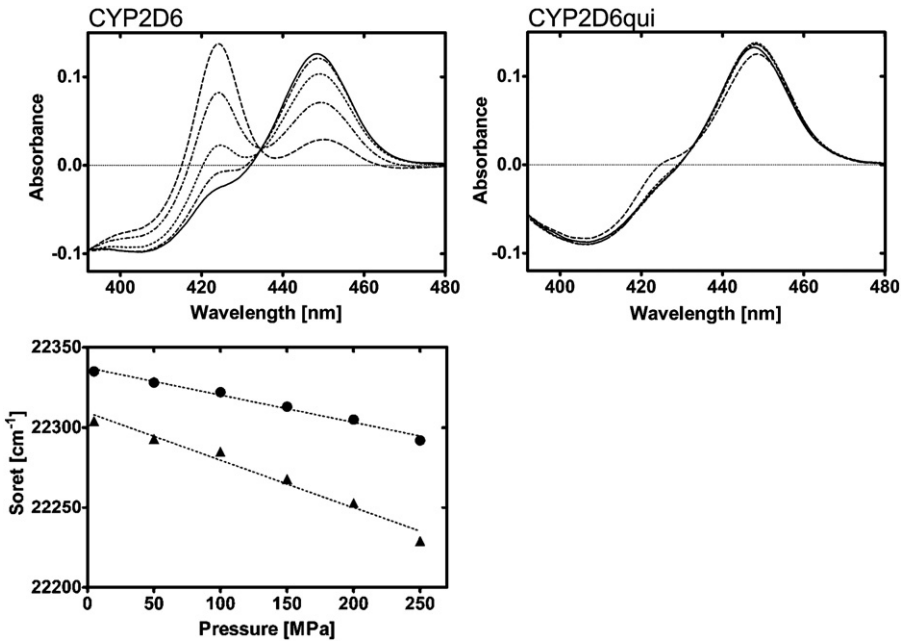


Fig. 2. Difference absorption spectra of reduced complexes of CYP2D6 in the absence (upper left panel) and presence of quinidine (upper right panel) at hydrostatic pressures of 50 to 250 MPa, with protein and quinidine concentrations of 2 and 20 μM, respectively, in 100 mM potassium phosphate buffer (pH 7.4) with 20% (v/v) glycerol. Lower panel: Soret band red shift of CO-reduced CYP2D6 with pressure in the presence (circles) and absence (triangles) of quinidine.

solvation. The change in heme solvation was recently identified as important contribution involved in pressure induced Soret band red shift of ligand free CYP forms [7]. To evaluate the abovementioned possibilities, changes in the flexibility of the CYP2D6 protein and solvation of its active site following quinidine binding, at normal and high hydrostatic pressure (NP and HP, respectively), were explored by MD simulations.

3.2. MD simulations, structural changes and fixation by quinidine at high pressure

MD simulations yielded information on two characteristics that reflect, to some extent, the plasticity of the protein molecule and flexibility of its active site or flexible regions (FR, for definition, see [28]). Firstly, the radius of gyration (R_g) of the protein, reflecting the effective radius (size) of the whole macromolecule, and secondly temperature B-factors (B_f), reflecting the fluctuation of atoms about their average positions due to thermal motion.

MD simulations under NP and HP showed that R_g values of the bound and ligand-free enzyme differ. Further, HP induces substantial (0.46 Å) reductions in R_g of the ligand-free enzyme, but only very minor (0.02 Å) changes in the R_g of the quinine-bound form (Table 2, Fig. 3). Thus, the unbound CYP2D6 molecule as a whole is quite compressible, but the quinidine-bound form, CYP2D6qui, appears to be substantially more rigid.

Temperature B factors obtained from the MD simulations allowed evaluation of the differences in flexibility of CYP2D6 with and without ligand bound in the active site. In both cases, the total B_f, i.e., flexibility of the enzyme, is lowered under HP (Table 2). However, the MD also enabled examination of the flexibility pattern of atoms at the active site by calculating the B_f of atoms in the active site, namely, closer than 10 Å from the heme. The results show that the core of the protein around heme is more rigid following binding of quinidine as the B_f of the active site atoms are considerably lower after its binding (cf. 5.44 Å vs. 4.54 Å, see Table 2). When high hydrostatic pressure is imposed, the B_f values follow the same pattern (4.41 Å vs. 3.74 Å), again indicating that CYP2D6 is more flexible without bound quinidine, as the difference between B_f values at high vs. normal pressure is 1.03 Å for the unliganded and 0.80 Å for the ligand (quinidine) bound enzyme. Taken together, the binding of quinidine is reflected in MD simulation by lower freedom of the active site atoms and their lower flexibility (determined as the change in B_f between normal and high pressure for unliganded and quinidine-bound CYP2D6).

Interestingly, when the average B_f for the backbone atoms of the whole molecule were calculated, the results revealed that whereas the active site stiffens, the loops become more flexible as the total B_f values increase upon quinidine binding (Table 2). Such behavior has also been described for other proteins, e.g. arabinose-binding protein [29].

Analysis of the active site solvation, which was also recently shown to be pressure sensitive [7], showed that more water molecules remained

Table 1
Compressibility-related coefficients of the Soret band shift ($-\alpha$) and regression (R^2) coefficients of selected cytochrome P450 enzymes obtained from the UV/VIS high pressure experiments.

CYP	$-\alpha$ [cm^{-1}/MPa]	R^2
2D6	0.297 ± 0.045	0.917
2D6qui	0.171 ± 0.012	0.981
1A2	0.253 ± 0.003 [4]	0.981
2A6	0.233 ± 0.014 [7]	0.986
2 C9	0.273 ± 0.039 [7]	0.926
2E1	0.391 ± 0.003 [28]	0.977
3A4	0.449 ± 0.014 [7]	0.996

Table 2
Results of MD simulations.

CYP	PDB	NP (0.1 MPa)			HP (300 MPa)		
		R_g[a]	B_f[b]		R_g[a]	B_f[b]	
		(Å)	<10 Å	Total	(Å)	<10 Å	Total
2D6	2F9Q	22.70 ± 0.06	5.44	11.37	22.24 ± 0.04	4.41	11.06
2D6qui	Roberts [19]	23.07 ± 0.05	4.54	13.36	23.05 ± 0.05	3.74	11.49

[a] Radius of gyration (R_g) calculated over the last 3 ns of the simulations on backbone atoms.
[b] B factors (B_f) calculated over the last 1 ns of the simulations on backbone atoms and around 10 Å from heme.

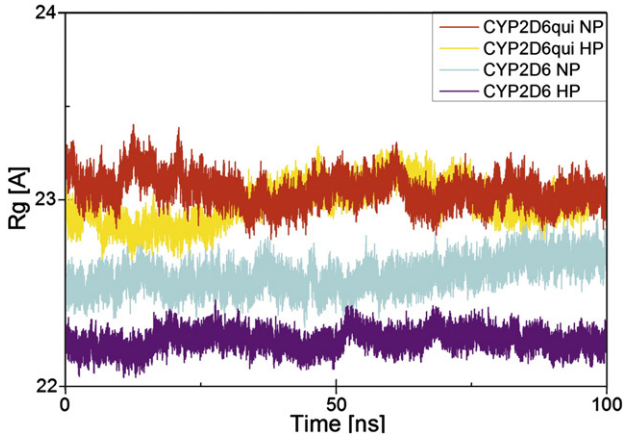


Fig. 3. Changes with time in the radius of gyration (R_g) of unbound and liganded CYP2D6 during MD simulations. Unliganded CYP2D6 is compressed upon application of pressure, as shown by the results obtained under normal pressure (CYP2D6NP, cyan, 2nd band from bottom) and high pressure (CYP2D6HP, violet, bottom band). The R_g of CYP2D6 with bound quinidine is larger than that of the unbound enzyme, as the enzyme has to accommodate the quinidine. However, its R_g is almost identical in simulations at both normal pressure (CYP2D6quiNP, red, top darker band) and high pressure (CYP2D6quiHP, yellow, top lighter band).

close to heme in simulations without quinidine. Generally, under HP more water molecules are observed in structures of both systems (CYP2D6 and CYP2D6qui). These HP injected water molecules likely constrain the flexibility of enzyme under HP, as indicated by lower B_f under HP [7] (Table 2). Quinidine occupies water positions in the vicinity of D301 in the B/C-loop, which were densely occupied in the ligand-free simulations (Fig. 4). Furthermore, the presence of quinidine in the active site of CYP2D6qui seems to block entry of water molecules into the active site, which may contribute to the high, experimentally observed, denaturation resistance of the quinidine-bound CYP2D6 at higher pressure.

The structural changes of the CYP2D6 protein upon quinidine binding under HP were detectable mainly in the distal part of the enzyme. However, the overall fold of CYP was conserved during all MD simulations. NP simulation of the unliganded structure showed a contraction of the F/G loop (compared to the crystal structure) and unwinding of the β4 sheet, while the structure of other parts remained similar to the crystal structure (Fig. 5, left panel). Presence of quinidine in the CYP2D6 active site led to preservation of both the expanded crystal-like F/G loop and complete β4 sheet, but at the cost of loss of B' helix in the B/C loop connected with straightening of the first two strands from the β1 sheet. Another visible difference was a motion of the F120 side chain to the vicinity of the heme propionate group while quinidine occupied its unliganded position on top of the porphyrin ring. In addition, the F helix was disrupted at the E211 position, which led to rotation of the end of the F helix by approximately 90°, thus stabilizing the expanded F/G loop (Fig. 5, right panel). Similar contraction of the F/G loop and unwinding of the β4 sheet identified in the NP simulation of unliganded CYP2D6 was also observed in HP simulation of CYP2D6. In addition, it led to a contraction of the active site cavity under pressure as F120 in the B/C loop and the F/G loop moved closer to each other.

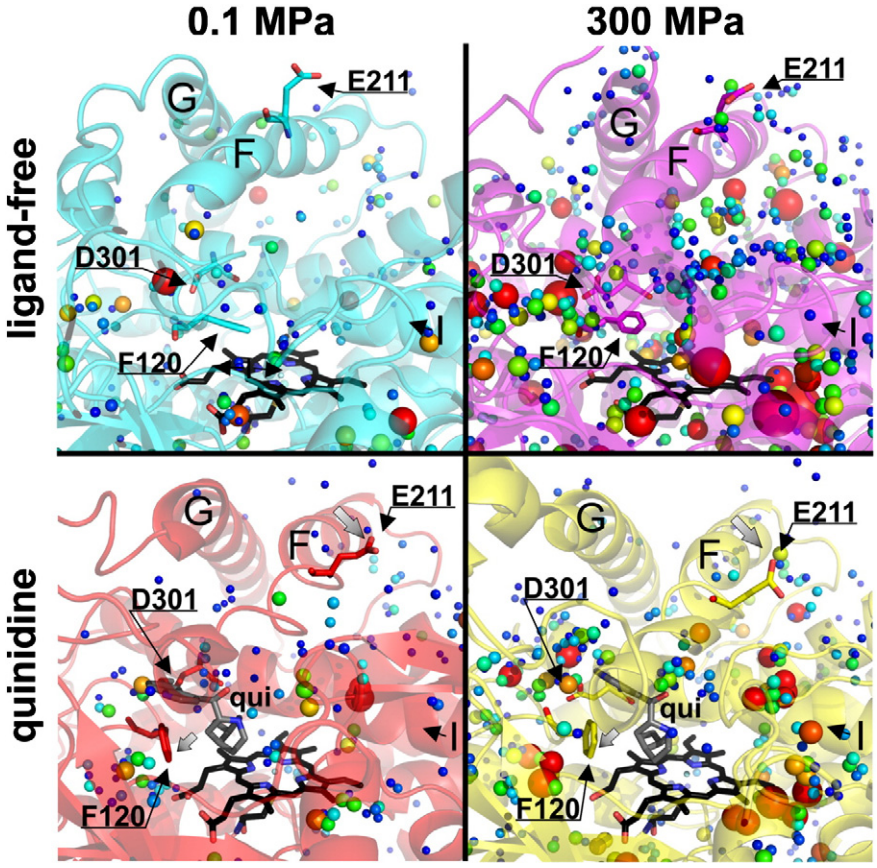


Fig. 4. Positions occupied by water molecules inside the CYP2D6 structures. Positions occupied by water molecules during 100 ns long simulations are shown on average structures. The structures from high pressure simulations contain more water molecules; however the presence of quinidine blocks the entry of water molecules to the active site in the vicinity of the probe (heme, shown in black). Quinidine (shown in gray) occupies otherwise conserved water positions in the vicinity of D301 in the B/C-loop. Changes of positions of F120 and E211 in adaptation to quinidine are shown by gray arrows. The positions of water molecules are shown by spheres of differing size and color. The positions where water molecules remained at least 1/10 of the simulation time (corresponding approximately to three times longer resistance times than in bulk water) are shown by small blue spheres. Longer residence times are shown by larger spheres and warmer colors and the largest red spheres correspond to the positions where water molecules remained for more than 1/3 of the simulation time.

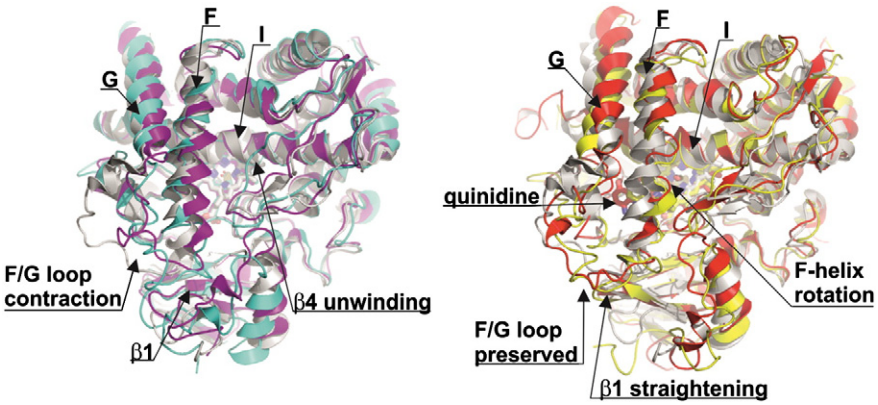


Fig. 5. Structural changes of CYPD26 in the presence/absence of quinidine. The crystal structure is shown in white; other structures are taken from 100 ns MD simulations. Left panel: structures without ligand at normal (cyan) and high pressure (violet). Right panel: corresponding structures with bound quinidine at normal (red) and high pressure (yellow). Positions of important structural elements and visible differences are depicted by black arrows.

In contrast, the active site of quinidine-bound structures remained almost unchanged under both pressures (Fig. 5, right panel) in line with its stiffening. In both simulations (CYP2D6quiNP, CYP2D6quiHP) the quinidine molecule itself is positioned in the vicinity of two key residues, F120 (in the B/C loop) and D301 (in helix I) and its position remained stable during MD. Consequently, quinidine plugs into the cavity between the B/C loop, F/G loop, I helix, β1 sheet and heme, thereby significantly contributing to prevention of compression of the active site.

4. Conclusion

Results obtained using two techniques, UV/VIS spectroscopy under high hydrostatic pressure and MD simulation collectively indicate that quinidine binding to CYP2D6 significantly affects the flexibility of the enzyme's active site. The experiments show that ligand-free CYP2D6 is rather compressible and denaturation-prone under high pressure, while quinidine-bound CYP2D6 is rigid and resistant to denaturation. MD simulations confirm this difference in behavior, and indicate that both active site flexibility and solvation are reduced by quinidine binding. Such information may facilitate the design of substrates or inhibitors of this and other (notably drug-metabolizing) enzymes that modify the flexibility and/or actively block and rigidify their active sites.

Abbreviations

HP	high pressure (300 MPa)
NP	normal pressure (0.1 MPa)
CYP	cytochrome P450
RDF	radial distribution function
Bf	B-factor
Rg	radius of gyration
qui	quinidine
CN	coordination number

Acknowledgments

The authors gratefully acknowledge support from the GACR (grant 303/09/1001 and 203/09/H046) and Student Project PrF_2011_020 of Palacky University. This work was also supported by the Operational Program Research and Development for Innovations – European Regional Development Fund (CZ.1.05/2.1.00/03.0058 and CZ.1.05/2.1.00/01.0030) and European Social Fund (CZ.1.07/2.3.00/20.0017).

References

[1] P.R. Ortiz de Montellano (Ed.), Cytochrome P450: Structure, Mechanism, and Biochemistry, third ed., Kluwer Academic / Plenum Publishers, New York, 2005.
[2] P. Anzenbacher, E. Anzenbacherova, Cell. Mol. Life Sci. 58 (2001) 737–747.
[3] F.P. Guengerich, in: P.R. Ortiz de Montellano (Ed.), Cytochrome P450: Structure, Mechanism, and Biochemistry, Kluwer Academic / Plenum Publishers, New York, 2005, pp. 377–530.
[4] E. Anzenbacherova, N. Bec, P. Anzenbacher, J. Hudecek, P. Soucek, C. Jung, A.W. Munro, R. Lange, Eur. J. Biochem. 267 (2000) 2916–2920.
[5] P. Anzenbacher, J. Hudecek, J. Inorg. Biochem. 87 (2001) 209–213.
[6] M. Ekroos, T. Sjogren, Proc. Natl. Acad. Sci. U. S. A. 103 (2006) 13682–13687.
[7] T. Hendrychova, E. Anzenbacherova, J. Hudecek, J. Skopalik, R. Lange, P. Hildebrandt, M. Otyepka, P. Anzenbacher, Biochim. Biophys. Acta 1814 (2011) 58–68.
[8] M. Otyepka, K. Berka, P. Anzenbacher, Curr. Drug Metab. 13 (2012) 130–142.
[9] T. Hendrychova, K. Berka, V. Navratilova, P. Anzenbacher, M. Otyepka, Curr. Drug Metab. 13 (2012) 177–189.
[10] R.C. Wade, P.J. Winn, E. Schlichting, Sudarko, J. Inorg. Biochem. 98 (2004) 1175–1182.
[11] V. Cojocaru, P.J. Winn, R.C. Wade, Biochim. Biophys. Acta 1770 (2007) 390–401.
[12] F. Bancel, G. Hui Bon Hoa, P. Anzenbacher, C. Balny, R. Lange, Methods Enzymol. 357 (2002) 145–157.
[13] P. Rowland, F.E. Blaney, M.G. Smyth, J.J. Jones, V.R. Leydon, A.K. Oxbrow, C.J. Lewis, M.G. Tennant, S. Modi, D.S. Eggleston, R.J. Chenery, A.M. Bridges, J. Biol. Chem. 281 (2006) 7614–7622.
[14] J.D. Marechal, C.A. Kemp, G.C. Roberts, M.J. Paine, C.R. Wolf, M.J. Sutcliffe, Br. J. Pharmacol. 153 (Suppl. 1) (2008) S82–S89.
[15] G.P. Hayhurst, J. Harlow, J. Chowdry, E. Gross, E. Hilton, M.S. Lennard, G.T. Tucker, S.W. Ellis, Biochem. J. 355 (2001) 373–379.
[16] L. Yang, C.Z. Ai, Y. Li, Y.H. Wang, Y.D. Chen, Bioorg. Med. Chem. Lett. 19 (2009) 803–806.
[17] F.P. Guengerich, G.P. Miller, I.H. Hanna, H. Sato, M.V. Martin, J. Biol. Chem. 277 (2002) 33711–33719.
[18] R.A. Branch, A. Adedoyin, R.F. Frye, J.W. Wilson, M. Romkes, Clin. Pharmacol. Ther. 68 (2000) 401–411.
[19] L.A. McLaughlin, M.J. Paine, C.A. Kemp, J.D. Marechal, J.U. Flanagan, C.J. Ward, M.J. Sutcliffe, G.C. Roberts, C.R. Wolf, J. Biol. Chem. 280 (2005) 38617–38624.
[20] M.J. Paine, L.A. McLaughlin, J.U. Flanagan, C.A. Kemp, M.J. Sutcliffe, G.C. Roberts, C.R. Wolf, J. Biol. Chem. 278 (2003) 4021–4027.
[21] J. Venhorst, A.M. ter Laak, J.N. Commandeur, Y. Funae, T. Hiroi, N.P. Vermeulen, J. Med. Chem. 46 (2003) 74–86.
[22] R. Lange, N. Bec, P. Anzenbacher, A.W. Munro, A.C. Gorren, B. Mayer, J. Inorg. Biochem. 87 (2001) 191–195.
[23] C. Jung, G. Hui Bon Hoa, D. Davydov, E. Gill, K. Heremans, Eur. J. Biochem. 233 (1995) 600–606.
[24] G.A. Schoch, J.K. Yano, M.R. Wester, K.J. Griffin, C.D. Stout, E.F. Johnson, J. Biol. Chem. 279 (2004) 9497–9503.
[25] W.D. Cornell, P. Cieplak, C.I. Bayly, P.A. Kollman, J. Am. Chem. Soc. 115 (1993) 9620–9631.
[26] J.M. Wang, P. Cieplak, P.A. Kollman, J. Comput. Chem. 21 (2000) 1049–1074.
[27] R. Santos, J. Hritz, C. Oostenbrink, J. Chem. Inf. Model. 50 (2010) 146–154.
[28] J. Skopalik, P. Anzenbacher, M. Otyepka, J. Phys. Chem. B 112 (2008) 8165–8173.
[29] C.A. MacRaild, A.H. Daranas, A. Bronowska, S.W. Homans, J. Mol. Biol. 368 (2007) 822–832.

Dynamics and Hydration of the Active Sites of Mammalian Cytochromes P450 Probed by Molecular Dynamics Simulations

Tereza Hendrychová [a], Karel Berka [a], Veronika Navrátilová[a], Pavel Anzenbacher[b] and Michal Otyepka[a,*]

[a]*Department of Physical Chemistry, Regional Centre of Advanced Technologies and Materials, Faculty of Science, Palacky University, Olomouc, Czech Republic;* [b]*Department of Pharmacology, Faculty of Medicine and Dentistry, Palacky University, Olomouc, Czech Republic*

Abstract: The flexibility, active site volume, solvation, and access path dynamics of six metabolically active mammalian cytochromes P450 (human 2A6, 2C9, 2D6, 2E1, 3A4 and rabbit 2B4) are extensively studied using molecular dynamics (MD) simulations. On average, the enzymes' overall structures equilibrate on a 50+ ns timescale. The very open CYP2B4 structure closes slowly over the course of the simulation. The volumes of the active sites fluctuate by more than 50% during the MD runs; these fluctuations are mainly due to movements of the main chains, with only a handful of amino acid residues in CYP2B4, CYP2D6, CYP2A6 and CYP2C9 showing significant independent side chain movement. The volume of the active site of CYP2E1 fluctuates heavily, ranging from 220 to 1310 $Å^3$, due to the opening and closing of gates to two adjacent cavities. CYP2E1 has the least hydrated active site of the studied CYPs; this is consistent with its preference for non-polar substrates. The CYP2A6 and CYP2E1 active sites are deeply buried, with access paths that are narrower than the radius of a water molecule. However, waters are still able to access these active sites due to local adaptations of the channel to accommodate their passage. This finding may imply that the access paths of the CYPs never fully open prior to contact with the substrate; instead, the substrate may induce adaptive conformational changes during its passage to the active site. This may also explain why some substrate recognition sites are localized along individual enzymes' access paths.

Keywords: Molecular dynamics, flexibility, cytochrome P450, channels, solvation, main chain, side chain, active site.

INTRODUCTION

In humans, the majority of commercial drugs are metabolized by the cytochrome P450 (CYP) enzymes in the liver [1-4]. The active site cofactor heme is directly involved in the catalytic cycle, which typically results in substrate oxidation [5-8]. Recently, analysis of the X-ray structures of some microsomal CYPs has provided new insights into certain aspects of substrate binding and catalysis in these enzymes, in atomic detail [9-38] (see Table **1**). In particular, these crystallographic analyses led to the identification of the enzymes' most important common structural features and variations [39]. Mammalian CYPs exhibit a common, predominantly α-helical CYP fold and a deeply buried active site housing the heme cofactor. However, there is considerable variation in the nature of the active site access paths [40] and in the sizes of the active sites of mammalian CYPs.

Experimental and computational studies have also demonstrated that there is also some variation in the enzymes' flexibility [41, 42]. Flexibility seems to play an important role in modulating enzyme specificity [29, 43] besides other important structural determinants of CYP substrate specificity, such as variations of amino acid composition of the active site and of access/egress channels [44-47]. The potential connection between CYP flexibility and promiscuity is summarized in a separate review in this special issue [48]. The CYPs must exhibit sufficient structural flexibility to allow the substrate to access the active site and the product to leave, *via* the access/egress channels. The active site access channels (in this study, both access and egress channels are referred to as access channels) identified in X-ray structures are mostly not wide enough to enable passage of substrates, because they contain bottlenecks with diameters of less than 3 Å [40]. However, while the bottleneck diameter is often observed to fluctuate, no spontaneous opening of the active site access paths to an extent that would allow the passage of a substrate has yet been detected in any conventional molecular dynamics (MD) simulation run in the absence of a ligand [42, 49-52] or with the substrate bound to the active site [41, 51, 53]. It is worth noting that fluctuations in the diameter of the access paths bottlenecks enables the solvation of the active site, i.e. the exchange of water molecules between the active site and bulk solvent; this typically occurs *via* the solvent channel [51, 52].

On the other hand, the very open structure of ligand-free CYP2B4 [54] illustrates the upper limit of conformational changes that the CYP structure can undergo [39]. However, the wide opening of CYP2B4 might be an artifact of its crystallization and its relevance under physiological conditions is questionable. The binding of erythromycin to CYP3A4 has also been shown to cause large conformational changes in the F/G loop [29]. The conformational changes induced by ligand passage through the channel have been the subject of several studies using various molecular dynamics techniques, including random acceleration molecular dynamics [47, 55-57] and steered molecular dynamics [50, 53, 57-59]. Despite this effort, the detailed mechanism by which substrates access and bind to CYP active sites remains unknown.

Molecular dynamics simulations were also used in the analysis of the overall conformational changes in the CYP structure [39, 42, 53, 60-65]. The analysis of active site solvation has been successfully carried out by MD simulations [51, 52, 66], which identified several positions with higher water occupancy within the active site. Such water molecules might be important for substrate binding or electron transfer [6, 66-68]. Thermal stability was also probed by MD simulations [69], as well as the interactions between substrate and CYP enzyme [29, 39, 46, 49, 57, 59, 70-72] including the effects of polymorphisms and mutations [73-80]. The modeling of the membrane topology of the CYP2C9 and its implications to the CYP function is a recent example of application of MD simulation in CYP structural biochemistry [81, 82].

In this paper, we have two key objectives: to identify the timescales on which selected structural features converge and to examine the structural variation of individual CYPs, primarily in terms of changes in the active site volume, solvation, and the opening of access channels during the simulation.

The five human liver microsomal CYP enzymes examined in this work, CYP2A6, 2C9, 2D6, 2E1, and 3A4, exemplify the most

*Address correspondence to this author at the Regional Centre of Advanced Technologies and Materials, Department of Physical Chemistry, Faculty of Science, Palacky University, 17. listopadu 12, 771 46 Olomouc, Czech Republic; Tel: +420585634764; Fax: +420585634761;
E-mail: michal.otyepka@upol.cz

1389-2002/12 $58.00+.00 © 2012 Bentham Science Publishers

Table 1. Summary of Solved Structures of Human CYPs.

CYP	pdbid	Ligand	Res [Å]	Ref	CYP	pdbid	Ligand	Res [Å]	Ref
1A2	2HI4	α-naphthoflavone	1.95	[9]	2E1	3LC4	12-(1H-imidazol-1-yl)	3.10	[23]
1B1	3PM0	α-naphthoflavone	2.70	[10]			dodecanoic acid		
2A6	2PG6		2.53	[11]		3KOH	8-(1H-imidazol-1-yl) octanoic acid	2.90	[23]
	2PG7		2.80	[11]					
	3EBS	phenacetin	2.15	[12]		3GPH	10-(1H-imidazol-1-yl) decanoic acid	2.70	[23]
	2FDY	4,4'-dipyridyl disulfide	1.95	[13]					
	2PG5	1,2-ethanediol	1.95	[11]	2R1	3C6G	vitamin D3, β-cyclodextrine†	2.80	[24]
	1Z10	coumarin	1.90	[14]					
	1Z11	methoxsalen	2.05	[14]		3DL9	1-hydroxy-vitamin D2, β-cyclodextrine†	2.72	[25]
	2FDU	N,N-dimethyl(5-(pyridin -3yl)furan-2-yl) methanamine	1.85	[13]					
						3CZH	vitamin D2, β-cyclodextrine†	2.30	[26]
	2FDV	N-methyl(5-(pyridin-3-yl) furan-2-yl) methan amine	1.65	[13]	3A4	1TQN		2.05	[27]
						1W0E		2.80	[28]
						1W0F	progesterone†	2.65	[28]
	2FDW	(5-(pyridin-3-yl)furan-2-yl) methanamine	2.05	[13]		1W0G	metyrapone	2.73	[28]
						2J0D	erythromycin A	2.75	[29]
2A13	2P85	indole	2.35	[15]		2V0M	ketoconazole	3.80	[29]
2B6	3IBD	4-(4-chlorophenyl) imidazole	2.00	[16]		3NXU	ritonavir	2.00	[30]
					7A1	3DAX		2.15	[31]
2C8	1PQ2	palmitic acid†	2.70	[17]	11A1	3NA0	20,22-dihydroxy	2.50	[32]
	2NNH	retinoic acid, palmitic acid†	2.60	[18]			cholesterol		
					19A1	3EQM	androstenedione	2.90	[33]
	2NNI	montelukast, palmitic acid†	2.80	[18]	46A1	2Q9G		2.40	[34]
						2Q9F	cholesterol-3-sulphate	1.90	[34]
	2NNJ	felodipine, palmitic acid†	2.28	[18]		3MDV	clotrimazole	2.40	[35]
						3MDT	voriconazole	2.30	[35]
	2VN0	troglitazone, palmitic acid†	2.70	[18]		3MDR	tranylcypromine	2.00	[35]
						3MDM	thioperamide	1.60	[35]
2C9	1OG2		2.60	[19]	51	3JUV		3.12	[36]
	1OG5	S-warfarin	2.55	[19]		3JUS	econazole	2.90	[36]
	1R9O	flurbiprofen	2.00	[20]		3LD6	ketoconazole	2.80	[36]
2D6	2F9Q		3.00	[21]	PGIS	2IAG		2.15	[37]
2E1	3E4E	4-methyl-1H-pyrazole	2.60	[22]		3B6H	minoxidil,	1.62	[38]
	3E6I	indazole	2.20	[22]			β-octylglucoside†		

† - outside of the active site.

important CYP forms in drug metabolism; with the rabbit enzyme CYP2B4, they are the most studied CYP enzymes [2, 83, 84]. A typical substrate of CYP2A6 is coumarin [2, 85]; CYP2A6 is also involved in the metabolism of nicotine and tobacco-specific procarcinogens. Rabbit CYP2B4 was chosen because its X-ray conformation is rather open [54]; we sought to identify and describe any closures of its open structure that might occur on the 100+ ns time scale. CYP2C9 is one of the most important CYPs in drug metabolism, being responsible for the biotransformation of a wide range of weakly acidic drugs such as warfarin and non-steroidal anti-inflammatory drugs and for the metabolism of many endogenous compounds [2, 85]. CYP2D6 is responsible for the metabolism of many psychoactive drugs and other compounds (such as beta adrenergic blocking agents) that have a basic nitrogen and a planar aromatic ring [21]. CYP2E1 was originally identified as an ethanol-oxidizing enzyme; it metabolizes organic solvents including benzene, alcohols and halogenated alkanes [86, 87]. CYP3A4 is the most abundant P450 enzyme in the human liver and intestines, being involved in the biotransformation of more than half of all commercial drugs [2, 3].

METHODS

MD simulation is a computational tool, which can be used in an attempt to model the biological system in an atomistic resolution. The available timescales of MD simulations have been extensively enlarged these days from picoseconds to hundreds-of-nanoseconds or even microseconds depending on the size of the system. However, they are still very short in comparison with biologically relevant timescales comprising seconds. Smaller systems can be simulated on a longer timescale and thus can overcome the so called "local conformational trap", which means that the structures generated by MD simulation often fluctuate around the starting structure. Larger timescale thus enables to observe extensive conformational changes such as domain motions or protein folding.

In the typical MD protocol, the biomolecular structure is subjected to the equations of motion using an empirical force field. The empirical force field is a set of equations and empirical parameters that relate molecular structure to energy. Current classical force fields have also some drawbacks because they do not explicitly account for polarization, the partial charges are conformation-insensitive, and the compensation of errors in non-covalent parameters is alarming [88]. Nevertheless, MD simulations provide unique data at atomic resolution on the limited timescales often unattainable by the experiment and can, when wisely used, complement the experimental data.

We conducted MD simulations of a set of drug-metabolizing mammalian CYPs (2A6, 2B4, 2C9, 2D6, 2E1, 3A4). The starting ligand-free structures of CYP2A6 (PDB ID: 1Z10 [14]), CYP2B4 (PDB ID: 1PO5 [54]), CYP2D6 (PDB ID: 2F9Q [21]), CYP2E1 (PDB ID: 3E6I [22]), and CYP3A4 (PDB ID: 1TQN [27]) were taken from the Protein Data Bank database [89]. The structure of the wild-type (wt) CYP2C9 enzyme was taken from our previous MD simulation [41], because the reported X-ray structure of CYP2C9 (PDB ID: 1OG2 [19]) is in fact that of a chimera with seven mutations in the F/G-loop (K206E, I215V, C216Y, S220P, P221A, I222L and I223L) that were inserted into the primary sequence to facilitate crystallization; these mutations were reversed in our earlier simulation. The positions of the backbone atoms of the missing residues (between P41 and Q52) of the CYP2D6 structure were modeled on those of the analogous residues in CYP2C8 (PDB ID: 1PQ2 [17]); the side chains of these residues were built using the Leap program of the AMBER package. The enzymes' N-terminal transmembrane domains were not included in the simulations.

The force field parameters for non-standard heme residue were developed *in situ* using the procedure of Cornell *et al*. [90]. In the studied enzymes, the one of the axial ligands of the heme iron is a cysteinate sulfur; the other is a molecule of carbon monoxide. The starting geometry of the heavy atoms of the heme residue for the quantum chemical calculation was taken from the X-ray structure of CYP2C9 (PDB ID: 1OG2 [19]); the missing CO and hydrogen atoms were added. The geometry of the system was optimized and RESP charges were added as in our previous works [41]. All enzymes were solvated by a TIP3P water box with ions added to neutralize the system containing usually around 60,000 atoms. After minimization, a 10 ns long unconstrained equilibration simulation under NpT conditions (T = 293.15 K, p = 0.1 MPa) with weak-coupling thermostat and barostat (both with 0.2 ps time constants) and 2 fs timestep was performed, followed by 100 ns production simulations run under the same conditions using the AMBER 9.0 package and the parm99 force field [91]. Cut-off for noncovalent interactions was 1 nm and electrostatics was treated with particle mesh Ewald method. Bonds involving hydrogen atoms were constrained with SHAKE algorithm. All MD trajectories were analyzed using the ptraj module of the AMBER package.

The volume and solvation of the active sites of the CYP enzymes were monitored throughout these long simulations. The volume of the active site was calculated using the CASTp server [92] with a probe radius of 1.4 Å with inclusion of the heme. Solvation was analyzed using the radial distribution functions (RDF) for the water molecules around the heme. The space-resolved distribution of water molecules was further evaluated on a grid with a resolution of 1 $Å^3$ similarly as in the analysis of waters within CYP2D6 [66]. Snapshots taken each 500 ps were superimposed on the backbone atoms and the positions of the water oxygen atoms on the grid were recorded. Under these conditions, there were approximately seven water molecules in each space element of the grid that contained bulk water. Sites within the enzyme were identified as water coordination positions if their average water occupancy was at least three times greater than this average value. Active site volumes were visualized by Hollow software [93].

To assess the flexibility of the active site, the motions of the structurally conserved residues that form the cavity were analyzed. Structural alignment was performed using ptraj over the Cα atoms, and the active site-forming residues in each crystal structure were identified using the CASTp server [92], followed by manual verification. The root-mean-square deviation (rmsd) of the positions of the main- and side-chain atoms of these residues over the course of the simulations were calculated on the basis of snapshots recorded every 500 ps. The motion of the main chain was analyzed in terms of the rmsd of the positions of the Cα atoms, while that of the side chains was analyzed in terms of the rmsd of their centroids' positions. The centroids were defined using the heavy atoms of the side-chains only; they were chosen because they facilitate comparisons between different amino acids.

Channel bottlenecks were analyzed using MOLE with the same set of snapshots [94]. The chosen starting point was the position of the oxygen of the carbon monoxide probe. The CO probe itself, along with all waters and hydrogens, was not factored into calculations of the channel diameter. The location and radius of the narrowest part of the largest of 20 potential channels was analyzed. Channels were classified as "open" if their bottleneck radius exceeded 1.52 Å (the MOLE parameter for the van der Waals radius of oxygen) and "broad" when their radius exceeded 3 Å.

RESULTS AND DISCUSSION

Overall Convergence

Six important drug-metabolizing microsomal CYPs were studied using classical MD simulations with explicit TIP3P solvation on the 100+ ns time scale. The enzymes' gross structural parameters are listed in Table **2**. The characteristic cytochrome P450 fold was conserved in all simulations. However, each structure underwent some rearrangements, as shown by their high measured flexibilities and the documented changes in their active site volume and radius

Table 2. General Structural Parameters of the CYPs Examined in this Study

CYP	PDB	Rg [Å]	RMSD [Å]	V_{init} [Å^3]	V_{sim} [Å^3]	V_{ligand} [Å^3][c]	N_{WAT}
2A6	1Z10	22.44 ± 0.05	1.69 ± 0.08	330	480 ± 430	140 ± 65	3 ± 1.1
2B4[a]	1PO5	22.75 ± 0.07	4.13 ± 0.10	12600	3900 ± 2000		14 ± 1.9
2C9	1OG2	22.79 ± 0.05	2.68 ± 0.11	1960	480 ± 440	250 ± 30	2.5 ± 1.5
2D6	2F9Q	22.70 ± 0.06	2.44 ± 0.09	1800	1500 ± 1100	265 ± 40	4 ± 1.4
2E1	3E6I	22.55 ± 0.04	1.95 ± 0.05	290 (800)[b]	440 ± 310	95 ± 30 (226)[d]	1 ± 1.2
3A4	1TQN	22.66 ± 0.05	2.33 ± 0.05	1950	2000 ± 1300	550 ± 275	14 ± 2.2

V_{init} is the volume of the active site cavity above heme plane, as calculated with the CASTp server, from the crystal structure [92]. V_{sim} is the volume of the same cavity averaged from 10 snapshots acquired over a time frame of 10 ns. N_{WAT} – number of waters identified by visual inspection inside the active site; quoted value represents an average from 10 snapshots acquired over 10 ns. The radius of gyration (Rg) and RMSD were calculated from the last 5 ns of the 100ns trajectory. [a] – values for CYP2B4 were calculated from a longer 150 ns simulation. [b] – value in parentheses is the sum including the cavity present in the CYP2E1 crystal structure bound with lipid molecules [23]. [c] – averaged on at least 8 most typical substrates for each human CYP (listed in Table ST2). [d] – van der waals volume for lauric acid.

of gyration. Most structures achieved RMSD equilibration within 50 ns, but each structure achieved equilibration on a different timescale. There were two extremes: CYP2A6 and CYP2E1 equilibrated rather quickly (within 20 ns), whereas the RMSD of CYP2C9 did not reach equilibrium even after 100 ns (see Fig. **S1** of the Supplemental Information).

Analysis of the enzymes' radius of gyration (Rg) showed that CYPs with 'closed' starting structures tended to retain their closed conformation (CYP3A4, CYP2A6, CYP2E1) or to exhibit a rather limited degree of opening (CYP2C9, CYP2D6) (see Fig. **S2** of the Supplemental Information). The very open structure of CYP2B4 closed over the course of the simulation, exhibiting substantial motion over the first 50 ns of the simulation (see Movie M1 of the Supplemental Information, or http://www.molmovdb.org/cgi-bin/morph.cgi?ID=980050-28395). The behavior of individual CYPs is discussed in the following sections.

Specific Findings Concerning Individual Members of the Mammalian CYP set

CYP2A6 is involved in the metabolism of relatively small molecules [2, 14, 85, 95]. This seems to be reflected in its overall rigidity; no remarkable structural changes were observed during the MD simulation. This finding is consistent with previously-reported data on the flexibility of this CYP [42]. The most flexible regions were the distal sides of the F/G, G/H and H/I loops, as described in our previous publications [42, 52]. CYP2A6 also has smaller B-factors than the other CYPs examined in this work, with the exception of the similarly rigid CYP2E1 (see Fig. **S3** in the Supplemental Information). The active site access paths in this enzyme are closed; the diameters of their bottlenecks fluctuate, but are generally less than 1.5 Å (cf. Fig. **7**). However, water molecules can enter and leave the active site *via* alternative routes such as channel 2b (Fig. **1**; the nomenclature of the active site access paths was chosen using the nomenclature proposed by Wade *et al.* [40]). On average, the active site contains 4 water molecules.

Rabbit **CYP2B4** crystallized as a very open homodimer in which each the F/G loop of one subunit projects into the active site of the other [54]. This open structure closed over the course of the simulation (Fig. **2**). In this closing motion, the F/G-loop moved over the active site cavity, towards the N-terminal β1-sheet. The motion of the F/G-loop is facilitated by the movement of the H-helix, in which the H/I loop acts as a hinge. Movements of the B'/C-loop also contribute to the closure of the active site, by repositioning the C-helix. The net result of these movements is a slow shrinking of the active site cavity over first 30 ns, which be comes even slower once the motion of the F/G loop starts to affect

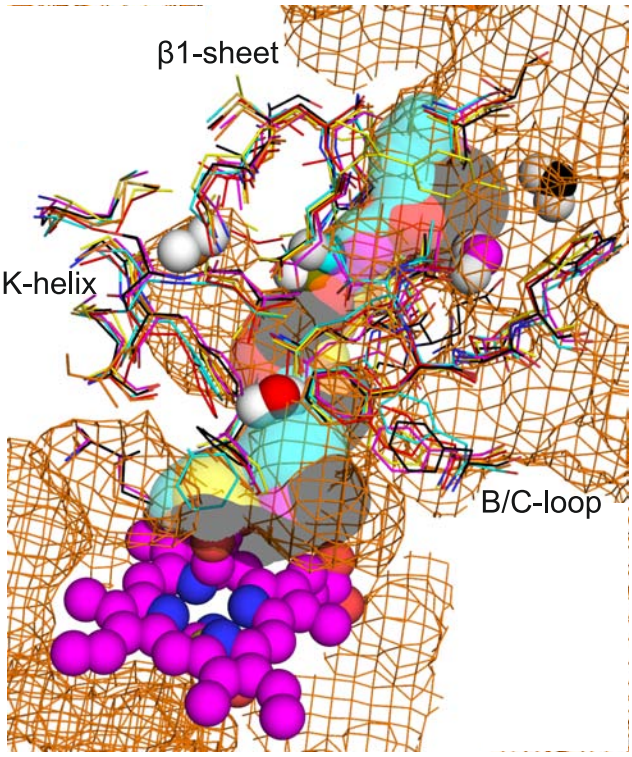


Fig. (1). The mechanism by which water molecules enter the CYP2A6 active site *via* channel 2b, as revealed by MD simulation within 5 ns. Waters are represented by spheres, with different colors indicating the time at which the water occupies the position indicated by the sphere; black shows their position at 0 ns, by the entrance to channel 2b, the violet spheres show the position after 1 ns, cyan corresponds to 2 ns, yellow to 3 ns, orange to 3.5 ns, and red to 5 ns and beyond. Channels were identified using MOLE [94] in the same frames and are colored accordingly. At a point 3.5 ns into the simulation, the water molecule was observed to occupy a side cavity, which is shown in the orange mesh used to denote the enzyme surface.

the A-helix, resulting in slow side-chain rearrangements. It is worth noting that the movable segments detected in our simulation have different positions in the open (1PO5) and closed (3MVR [96]) crystal structures of CYP2B4 shown in Fig. **2**. Overall, these data suggest that the open structure of CYP2B4 is unstable under ambi-

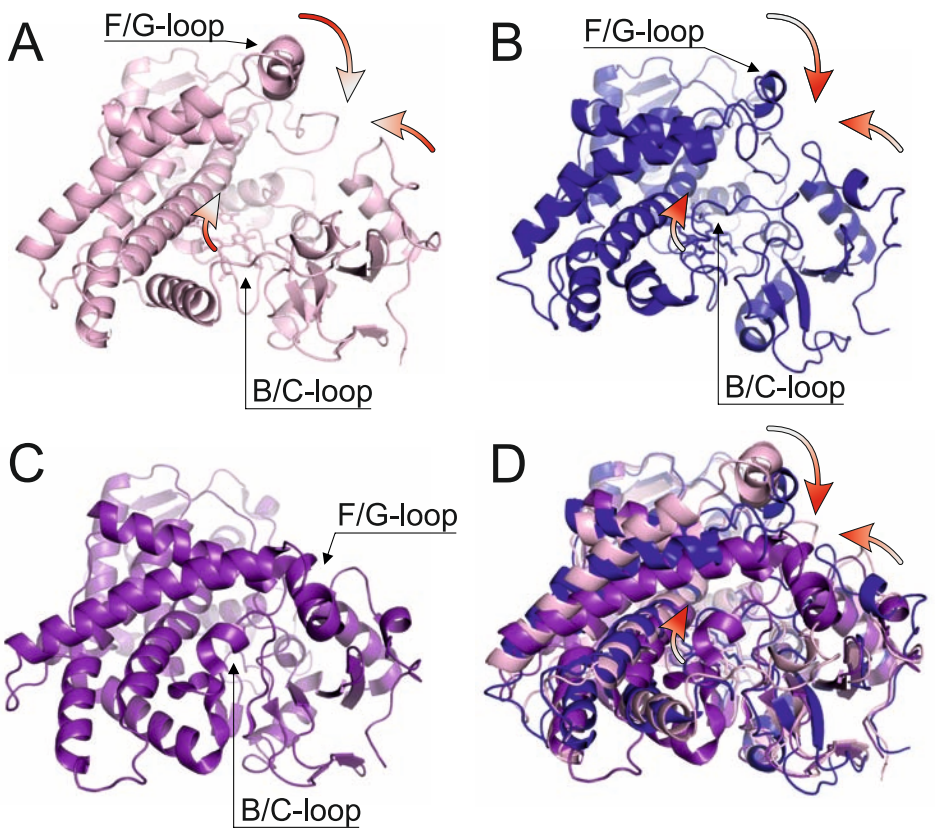


Fig. (2). The closing motion of rabbit CYP2B4. **A** – The open X-ray structure of CYP2B4 (1PO5), which was used as the starting structure for the MD simulation, **B** – a snapshot from the MD simulation taken at 100 ns; **C** – the closed X-ray structure (3MVR), **D** – a superposition of all of the CYP2B4 structures. Arrows show the major motions of the B/C-loop, the F/G-loop, and the N-terminal segment.

ent conditions, i.e. when the protein is not tightly packed in the crystal lattice.

Structurally, **CYP2C9** is one of the best-described CYP enzymes [2, 85]. The most flexible part is F/G-loop, as shown in Fig. **3A**. The movement of F/G-loop is reflected in the movement of the A helix and is connected with changes in the conformations of several phenylalanine residues (F69, F219, and F476). CYP2C9 has an open solvent (S) channel. Interestingly, the solvation of the active site is increased by a change in the conformation of the segment formed by the A and B' helices that occurs between 30 and 50 ns.

Unfortunately, the only available crystal structure of **CYP2D6 (**2F9Q) has a rather low resolution (3.00 Å) and so does not provide information on the position of the A'-segment or those of water molecules. The core fold of CYP2D6 was stable throughout the simulation; the only significant change observed was a shift in the F/G-loop region as shown in Fig. **3B**. It should be noted that the F/G-loop makes contact with the A' region, which is disordered (missing) in the crystal structure of CYP2D6 and was therefore modeled on the crystal structure of CYP2C8 (PDB ID: 1PQ2 [17]) and still this structure was stable during the simulation.

Like CYP2A6, **CYP2E1** metabolizes low polar molecules with comparatively small molecular weights [86]; this is reflected in its high rigidity, abundance of hydrophobic residues (mostly phenylalanines), and small active site relative to those of other CYPs. Its ability to biotransform also some endogenous substrates (fatty acids) is thought to be facilitated by the presence of two additional cavities that are very near to the active site: the second cavity, which is located between the I, F, G and B' helices [23], and the third cavity, which is located between the β1 sheet, the F' helix, and the B/C loop (see Table **3** for detailed information on the lining residues). Our results show that these secondary cavities can be opened selectively, causing them to 'fuse' with the active site. The third cavity was opened by a sliding motion of a phenylalanine residue, F478, at around 30 ns (see Fig. **4**). This opening closed within 5 ns. A concerted rotation of two phenylalanines, F106 and F116, opened the connection to the second cavity at around 100 ns (see Fig. **4**).

CYP3A4 is involved in the metabolism of many structurally diverse compounds [2, 3]. Its broad substrate scope probably reflects the high intrinsic flexibility of its active site [42, 52]. The biggest movements in this species involved the F'-helix, the F'/G'-loop, and the G/H-loop. The so-called "phenylalanine roof" consists of residues F108, F213, F215, F241, F304, and F220; it is located directly above the active site and has been proposed to act as a gate for the substrate [59]. While the side-chain of F215 was observed to undergo a sliding motion and there was some rotation of the side-chain ring of F220 (Fig. **5**), most of the flexibility of CYP3A4 stems from the motions of the main chain (see Table **4**).

Common Properties of the Active Sites

The active sites of the CYP enzymes are situated above the heme, surrounded by residues from several parts of the molecule. Table **4** lists all of the structurally conserved residues that were found to be part of the active site walls. Most of these residues belong to the substrate recognition sites (the process used to define the substrate recognition sites is discussed elsewhere [44]). Additionally, most of them are located on the I-helix, including the most

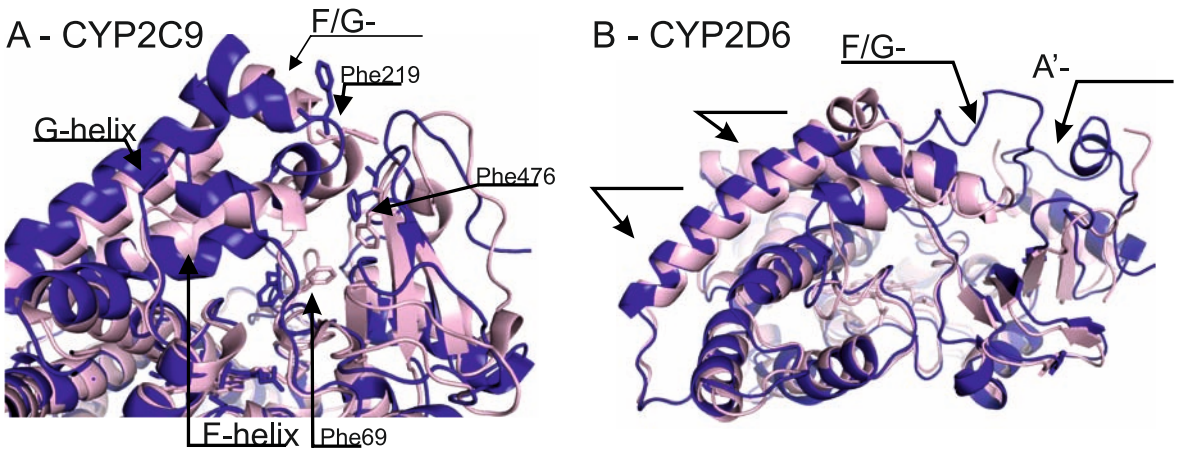


Fig. (3). Motion of F/G-loop of CYP2C9 and CYP2D6 during the simulation. The left panel shows the superimposed structures of CYP2C9 – crystal structure 1OG2 (light pink) and a snapshot taken at 100 ns (blue). The shift of F/G-loop and of A-segment is visible. The movement of the main chain causes movements of several phenylalanines (F69, F219, and F476). The right panel shows the aligned structures of CYP2D6 – crystal structure 2F9Q (light pink) and a 100 ns snapshot (blue). Here, the shift of the F/G-loop associated with the shift of the F and G helices is less pronounced.

Table 3. Fluctuations in the Volume of the CYP2E1 Active Site

Time [ns]	V [Å^3]	Residues encapsulating cavity of active site of CYP2E1
Crystal	360 (1760)	I115, F116, F207, F298, A299, E302, T303, V364, L368, F478
10	480	F106, H109, I115, F116, F207, L210, S211, D295, F298, A299, E302, T303, L363, V364, L368, F478, G479
20	250	I115, F116, F207, L210, S211, F298, A299, T303, L363, V364, L368, F478, G479
30	1310	I41, I42, L48, L50, I53, P54, F57, L70, V72, G73, S74, Q75, M77, G101, D102, L103, P104, I115, F116, F207, L210, S211, Q216, N219, N220, F221, P222, S223, F224, F298, A299, T303, L363, V364, P365, S366, N367, L368, P369, V388, V390, G477, F478, G479, I481
40	270	I115, F116, F207, L210, S211, F298, A299, T303, L363, V364, L368, F478, G479
50	300	I115, F116, F207, L210, S211, F298, A299, E302, T303, L363, V364, L368, F478, G479, I481
60	220	I115, F116, F207, L210, F298, A299, E302, T303, V364, L368, F478, G479
70	280	I115, F116, F207, L210, S211, F298, A299, E302, T303, L363, V364, L368, F478, G479
80	580	F106, I115, F116, F207, L210, F298, A299, E302, T303, V364, L368
90	410	I115, F116, F207, L210, S211, F298, A299, E302, T303, L363, V364, L368, F478, G479, I481
100	830	G101, D102, L103, F106, H107, H109, R110, R112, G113, I114, I115, F116, N117, N118, L201, F202, N205, F207, L210, S211, N238, V239, V242, V291, A294, D295, F298, A299, E302, T303, V364, L368, F478, G479, I481

The volume was calculated with the CASTp server [92], using snapshots taken from the MD simulation. The value in brackets shows the volume of cavity if the heme atoms are ignored; the heme thus blocks around ~1400 Å^3. Two large openings to the side cavities were detected after 30 ns and 100 ns, respectively; their volumes are shown in bold. Residues shown in bold encapsulated the cavity throughout the whole simulation, while grey bolded residues are those which constituted part of the active site wall for at least 75% of the simulated time.

highly conserved residue – a threonine that is essential for the enzymes' catalytic activity [5]. The conserved residues are usually hydrophobic amino acids; many are phenylalanines, which are thought to form "gate doors" governing the access of the substrate to the active site [23, 28, 59].

The biggest changes in the size and configuration of the active site were generally attributable to motions of the main chains (i.e. movements of secondary structure elements); only a few side chains exhibited significant independent motion (see Table **4** and Fig. **S5** in Supplemental Information). Of the active sites examined in this work, the most rigid was that of CYP2A6, followed by CYP2E1. The active sites of CYP3A4, CYP2D6, and CYP2C9 were significantly more flexible, and that of CYP2B4 was the most flexible of all. However, this may be primarily due to the large movements associated with the closing of the very open starting structure used for CYP2B4, as discussed above. The largest movements were mostly located in F/G-loop, which was recently found to have uncorrelated motions with other parts of the CYP119 [61]. Thus the F/G-loop is able to move independently on the rest of the protein.

Table 4. Flexibility of Common Structurally Aligned Residues in the CYP Active Sites

CYP	PDB	Residues											RMSD	
		B'	BC		F	I				K	β1	β4	Cα	SC
		SRS1			SRS2	SRS4				SRS5	SRS5	SRS6	[Å]	[Å]
2A6	1Z10	**F107**	V117	**F118**	**F209**	I300	G301	*E304*[a]	T305[b]	I366	L370	**F480**	7.2	+1.7
2B4	1PO5	V104	I114	**F115**	**F206**	**F297**	A298	E301	T302[b]	I363	V367	V477	20.0	+3.6
2C9	1OG2	A103	V113	**F114**	I205	G296	A297	E300	T301[b]	L362	L366	**F476**	13.0	+1.6
2D6	2F9Q	I106	V119	**F120**	L213	S304	A305	V308	T309[b]	V370	M374	**F483**	12.8	+3.8
2E1	3E6I	***F106***[a]	I115	**F116**	**F207**	**F298**[c]	A299	E302	T303[b]	V364	L368	**F478**	10.3	+1.1
3A4	1TQN	**F108**	S119	I120	R212	**F304**	A305	E308	T309[b]	A370	E374	L482	12.3	+1.2

Single letters denote helices (e.g. I indicates that the residue is part of the I-helix), two-letter codes denote loops (e.g. BC means the residue is in the B/C-loop), and Greek letters denote sheets. Phenylalanine residues are shown in bold. RMSD – root-mean-square deviation of the aligned residues' positions over the course of the whole simulation, calculated on the basis of the position of Cα with additional contribution from the sidechains (SC). SRS denotes amino acids from substrate recognition sites, as specified elsewhere [44]. [a] – not in the active site in this specific enzyme crystal structure, but present in other CYPs; [b] – highly conserved Thr residue important for the catalytic activity of all CYPs; [c] – gate-opening residue located between the two cavities identified in the crystal structures of CYP2E1 [23].

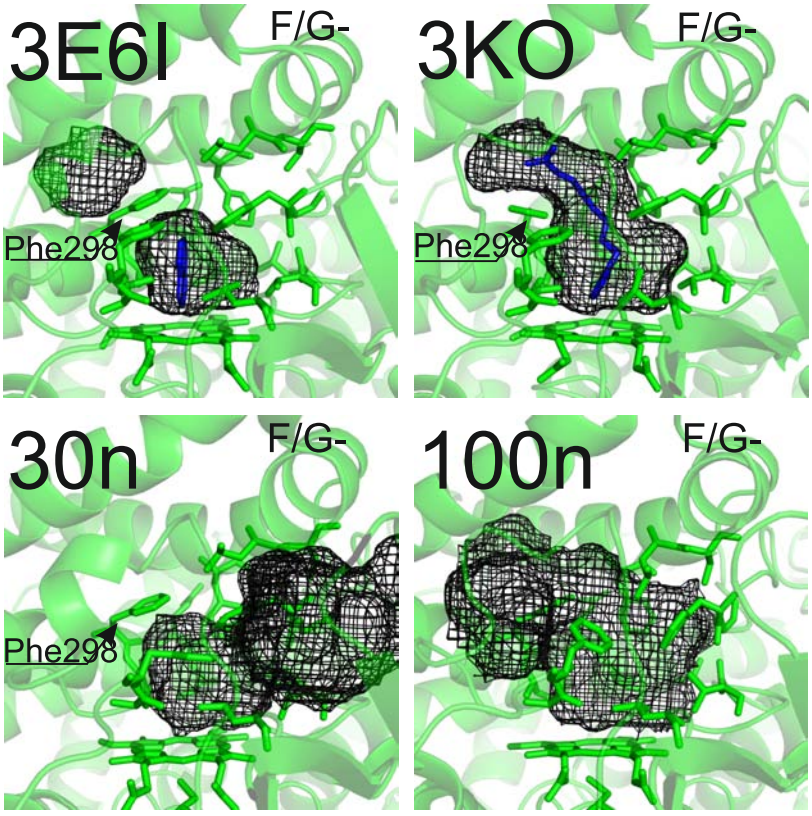


Fig. (4). Joining of the secondary cavities in CYP2E1. The upper left panel shows the active site in crystal structure of CYP2E1 with small ligand shown in blue (3E6I), where two cavities are separated by F298. The upper right panel shows active site of crystal structure of CYP2E1 with lipid-like ligand shown in blue (3KOH). Here the cavities are joined together. Same joining is shown in the lower right panel where these cavities are also connected in a snapshot from time 100 ns. Different cavity joining was detected in a snapshot from 30 ns (lower left), where active site opens to the cavity leading into 2a or 2b channel.

The closing of CYP2B4 was also accompanied by a rearrangement of the side chains (mostly those of F206, F297, V104 and E301), which is reflected in the high contribution of side-chain movements to the overall flexibility of this CYP2B4. Prominent side chain motions were also observed in CYP2D6, primarily because of the movements of the F483 and V119 side-chains. Amino acids whose side chains exhibited significant motion that was not connected to larger structural rearrangements were also identified in CYP2A6 (F209) and in CYP2C9 (F476). No side chains with significant independent motion were identified in the active sites of CYP2E1 and CYP3A4, although significant movements of the main chain were observed in the latter.

The CYPs examined in this work can be divided into three groups on the basis of the volumes of their active sites, as seen in

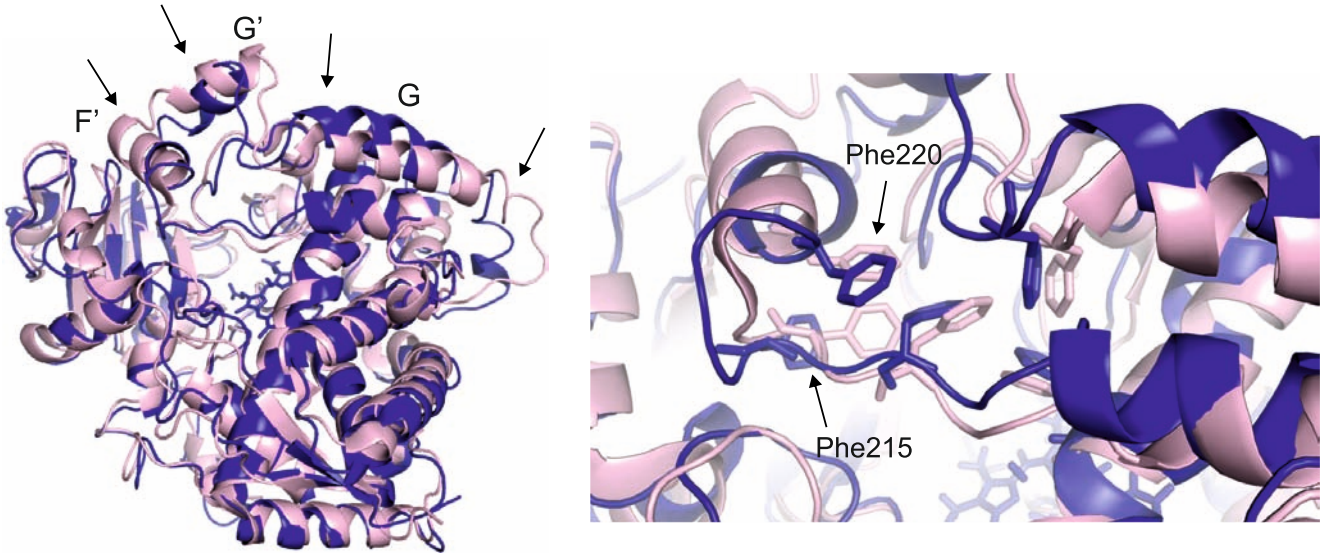


Fig. (5). Changes in the structure of CYP3A4 during the simulation. The left panel shows a superposition of the X-ray structure (pink) of CYP3A4 (1TQN) and a snapshot of the MD run at 100ns (blue). The regions that underwent the largest conformational changes are highlighted by black arrows. The right panel shows the motions of the "phenylalanine roof" residues.

their X-ray structures. The first group consists of CYP2A6, CYP2E1, which have the smallest active sites; in both, the active site has a volume of around 300 Å^3. The second group consists of CYPs with a moderate active site volume of around 1900 Å^3 (CYP2C9, CYP2D6, CYP3A4); the third contains only CYP2B4, which has an enormous active site of around 12500 Å^3. During the MD simulations, the volumes of the CYP active sites fluctuated widely (see Table **2**). However, two outliers were identified: the largest cavity of CYP2B4 closed progressively over the course of the simulation, while CYP2E1 has several cavities adjacent to the active site that opened and closed on timescales of tenths of a nanosecond over the course of the simulation, as shown in Table **3**; their opening and closing causes significant variation in the volume of the CYP2E1 active site.

The volumes of the active sites reflect also the size of the ligand, which can be accepted by the active sites. For instance, CYP2A6 and CYP2E1 with the smallest active site cavities generally metabolize smaller substrates like chloroform or ethanol, respectively. Opening of external cavities in CYP2E1 allows accommodation of larger substrates, e.g. fatty acids. CYP2C9 and CYP2D6 have intermediate active site volumes and they both interact with medium sized substrates, e.g. propranolol or tramadol. Cavities of these two CYPs have similar size especially in the crystal structures and they are also widely fluctuating in simulations with volumes between 100 Å^3 and 2000 Å^3. The largest substrates in humans (e.g. cyclosporine) are metabolized by CYP3A4 having the largest cavity. However what is true about the average ligand size is not necessarily true for individual ligands, where each CYP is able to metabolize at least one substrate with medium-sized van der Waals volume of about 250-300 Å^3 (see Table **ST2**), while extreme sized molecules (either too small or too big) are metabolized by specifically sized active sites.

Solvation of the Active Site

The solvation of the active site was studied by analyzing the radial distribution function (RDF) for the positions of water molecules (specifically, the positions of their oxygen atoms) relative to the heme's CO ligand and by visual inspection. Examination of the radial distribution functions revealed that each CYP active site contained a different number of water molecules in the vicinity of the heme (see Table **2**). Inspection of the waters within the CYP structures confirmed that each CYP has different water solvation pattern within its active site and within the structure (see Fig. **6**). Inclusion of water molecules was shown to influence the reliability of site-of-metabolism prediction in CYP2D6 [66].

There are several common regions that exhibit dense water solvation [67], including: (A) the region between the heme and the I-helix, (B) the entrance to the water channel, (C) the inner cavity between the heme and the end of the K-helix, and (D) the B'/C-loop region. However, none of these common regions is occupied to the same extent within each CYP, and each CYP has a unique water solvation region as shown in Fig. **6** and Fig. **S4** of the Supplemental Information.

Water molecules in **CYP2A6** prefer mostly the region D, where they interact with the residues from B'/C-loop. The second most densely-solvated position is region A, where one water molecule binds in a 'kink' in the I-helix. Region B which contains the opening of the water channel near the B-helix, is also strongly solvated. All of these regions are also solvated in the crystal structures of CYP2A6 (1Z10, 1Z11, 2FDU, 2FDV, 2FDW, 2FDY, 2PG5, 2PG6, 2PG7, and 3EBS), but in the crystal structures, the water groove of the I-helix in region A is shifted relative to the heme; it is situated above the pyrrole ring in the simulation, but lies above the linker between the two pyrrole rings in the crystal structures. Region C is also occupied, but the positions of the waters are in this region less well-defined; this is probably due to the presence of several hydrophobic residues (F432, V365, I366) and a relatively big cavity, which facilitates large water motions. CYP2A6 has another strongly-solvated site, located between the I-helix and the beginning of the E-helix. This position is also solvated in the CYP crystal structures [39].

Waters in **CYP2B4** are ejected from the active site as it closes; they are most abundant in region A. As in CYP2A6, the water groove in the I-helix shifts relative to the heme over the course of the simulation, adopting a position above the pyrrole ring. The second densely solvated position is region C. Regions B and D contain no water molecules; this may be due to the big motion of the B/C-loop during the closing of the pocket from open structure of the 1PO5 structure. Another strongly solvated region is located at the end of E-helix near the R145 and S181 residues. Waters are observed in this position in both published crystal structures of CYP2B4 (1PO5, 3MVR).

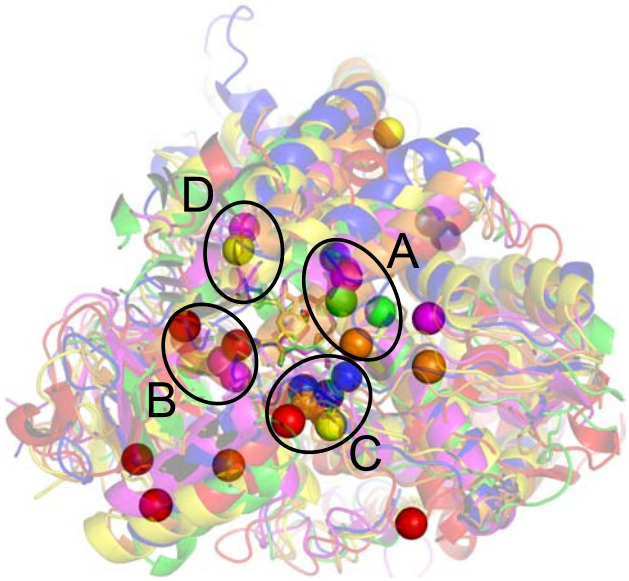


Fig. (6). Differences in active site hydration between the CYP enzymes. Spheres indicate sites occupied by water molecules for at least 35 ns of the total simulation time of 100 ns. The final backbones of the CYP structures after completion of the MD run are shown from the distal side, superimposed on one-another; they are depicted in a transparent cartoon form, with CYP2A6 shown in magenta, CYP2B4 in blue, CYP2C9 in green, CYP2D6 in yellow, CYP2E1 in orange, and CYP3A4 in red. The same color code is used for the spheres denoting the positions of the water molecules. The positions of common regions are indicated with circles: (**A**) the water between the heme and the I-helix, (**B**) the entrance to the water channel, (**C**) the inner cavity between the heme and the end of K-helix, and (**D**) the B'/C-loop region.

The water molecules in **CYP2C9** preferentially fill the water groove in the I-helix (region A), but there are also additional water positions nearby, between the I- and E-helices, forming a sort of chain of waters. The second occupied region is region B, which contains a water channel. There are additional water-occupied positions around the entrance of the solvent channel; many of these waters are also observed in the published CYP2C9 crystal structures (1R9O, 1OG2, 1OG5). Regions C and D contain fewer water molecules; in particular, D contains none.

Unusually, no water molecule was observed in the kink of the I-helix in region A in **CYP2D6**. Waters were, however, observed in other regions. Overall, the active site of this enzyme contained fewer water molecules than those of the other CYPs examined in this work, and some water molecules were observed to 'leak' from its non-polar active site over the course of the simulation. Because of its comparatively low resolution (3.0 Å), only a few water molecules can be discerned in the crystal structure of CYP2D6 (2F9Q); their positions are consistent with those identified during the MD simulation, being located in regions B and C.

Once again, no water was observed in the I-helix of region A in **CYP2E1**; in fact, the active site was not hydrated for more than half of the MD simulation. The most strongly solvated regions of this enzyme were thus region C and the solvent channel in the vicinity of the E, F, and I-helices. This picture is consistent with the CYP2E1 crystal structures (3E4E, 3E6I, 3GPH, 3KOH, 3LC4), in which active site waters are predominantly observed in region C; in fact, only one structure (3E6I) shows any water at all in region A. Regions B and D are also sporadically occupied by water molecules, albeit to a much lesser extent than in other CYP structures.

Waters were observed in all of the common regions of **CYP3A4**, but with different occupancies. The most preferred position of water molecules was region B, in which the waters fill the space within the water channel. Waters in region A seemed to be more mobile than in the other CYPs; the same was true for the waters in region C. The water in CYP3A4 also fills the positions just above the heme, in close proximity to the carbon monoxide; no water was observed in this position in the other CYPs. Water is also observed in these positions in the crystal structures of CYP3A4 (1TQN, 1W0E, 1W0F, 1W0G, 2J0D, 2V0M, 3NXU). Unlike most CYPs, the A-helix and β1-sheet of CYP3A4 are also quite strongly solvated, while these parts are thought to be immersed in the membrane of the endoplasmic reticulum membrane *in vivo* in CYPs [97] and these parts are immersed in recent atomistic membrane models of CYP2C9 [81, 82]. This might implicate weaker attachment of the CYP3A4 to the membrane.

The active sites of the individual CYP enzymes thus exhibit unique solvation patterns, as described above. The least-solvated active sites are those of CYP2D6 and especially CYP2E1, which is known to metabolize very non-polar substrates such as halogenated hydrocarbons. The differences in the solvation of the active site reflect differences in the active sites' polarities, which may in turn affect the enzymes' substrate specificities – enzymes can be expected to favor substrates whose polarity complements that of their active site. The active site itself is inside catalytic domain, which is partially immersed in the membrane. Two recent atomistic membrane models of CYP2C9 have shown that the flexibility and the overall fold are not much influenced by the presence of the membrane [81, 82]. It is therefore likely that the results about CYPs' active site based on solvated models such as those used within this review and previously should hold also for microsomal membrane-bound CYPs.

Residence times of water molecules in the active site are also affected by the opening and closing of the channels into the active site as shown in the next section. Of the enzymes studied, CYP2A6 exhibited the longest residence times for water molecules inside the active site, due to its very narrow active site access paths.

Channel Openings During the Simulation

Water molecules leaked into and out of the CYPs during the simulation *via* various channels. Because the different CYPs exhibited different solvation patterns, the width and locations of their channel openings were also different, as shown in Fig. **7** and Fig. **8**. The access path bottlenecks decreased over the course of the simulation for all of the CYPs save for CYP2C9. The only starting structure with very open channels was that of CYP2B4; like the rest of its structure, its channels closed up somewhat over the course of the simulation, but were still relatively broad by its end (see the Methods section for the definitions of open and broad channels). The channels of CYP2C9 remained relatively open throughout the simulation; this is consistent with the observed progressive increase in the enzyme's radius of gyration (see Fig. **S2** of the Supplemental Information). The channels of CYP3A4 and CYP2D6 did not remain open for the entirety of the simulation; those of CYP2A6 and CYP2E1 did not open at all. However, even the active sites of these enzymes gradually filled with water molecules; the water reached the active site by taking small steps inside the channels, which adapt to its presence as shown for CYP2A6 (Fig. **1**).

The structures of the studied CYPs contain several different access/egress channels. These were named using the devised by Wade *et al*. [40] (see Fig. **8**):

CYP2A6 – No channel was observed to be fully open at any point during the whole simulation; those that exhibited the widest openings were channels 2b and 2e. Due to the unrealistically small bottleneck radii, a lot of artifact openings were detected.

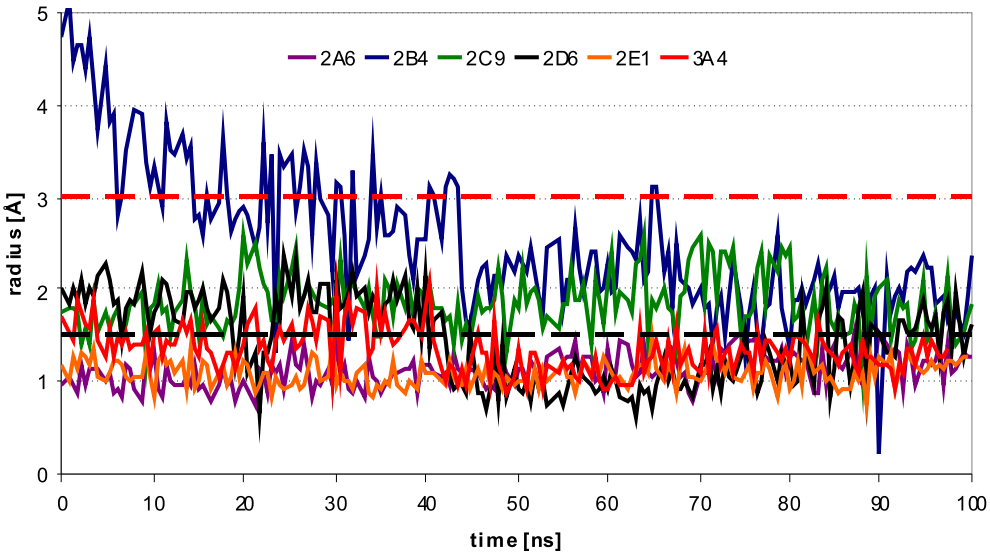


Fig. (7). The radius of CYP channel bottlenecks during the simulations. The smallest bottleneck radius of the biggest channel detected using MOLE [94] in snapshots from the simulation is shown; black dashed lines denote "open" channels, while red dashed lines show "broad" channel as defined in the Methods section. The only structure with broad channels was CYP2B4, whose channels were gradually closing over the course of the simulation. CYP2C9 was the only other structure whose channels remained open throughout the simulation, while those of CYP3A4 and CYP2D6 were only open for a part of the simulation. The channels of CYP2A6 and CYP2E1 did not open at all during the 100 ns simulation.

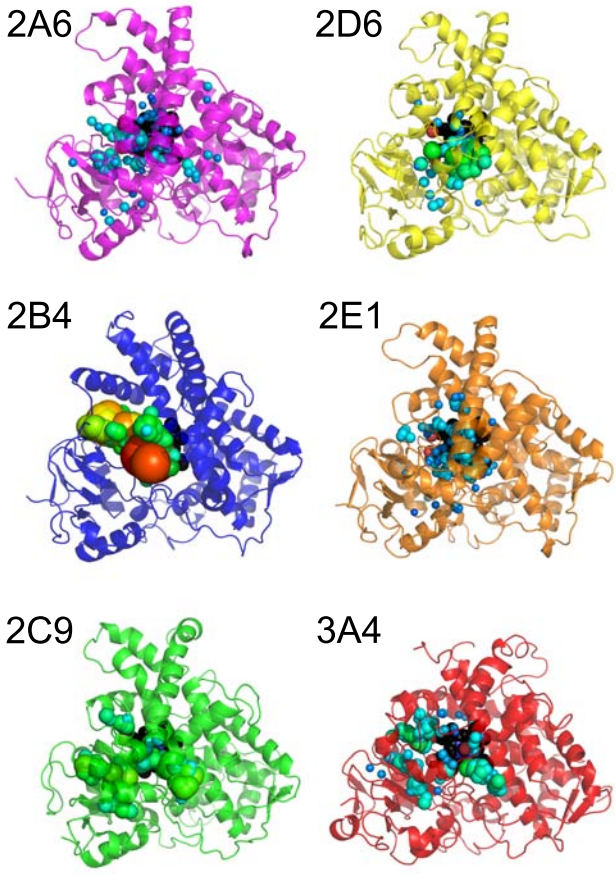


Fig. (8). Distribution of channel bottleneck positions within the CYPs during the MD simulations. The bottlenecks' positions identified by MOLE [94] are shown by spheres; their radii are shown by the sphere's colors (cyan – less than 1.5 Å; green – between 1.5 and 2 Å; yellow – between 2.and 3.0 Å; red - greater than 3.0 Å).

CYP2B4 – The widest channel was 2a, especially in the open structure. As the enzyme's structure closed up, the 2e channel became more open; occasional openings of the 2ac channel were also observed.

CYP2C9 – The 2b and solvent (S) channels were opened; the 2a and 2e channels were also partially open.

CYP2D6 – The only open channel was the solvent channel. However, because many of the bottleneck radii were less than the radius of an oxygen atom, a lot of unrealistically small artifact openings were found similarly as was observed with CYP2A6.

CYP2E1 – No channel was fully open at any point during the simulation; the most widely-opened channels were 2ac and 2e. Numerous small artifact openings were also detected.

CYP3A4 – The solvent and 2e channels were open; partial openings of channels 2b and 2a were also detected.

As listed above, the widest channels were usually 2b (CYP2A6, CYP2C9; located between B/B′ loop and the β1 and β3 sheets), 2e (CYP2A6, CYP2B4, CYP2E1, CYP3A4; running through the B/C loop) and S (CYP2C9, CYP2D6, CYP3A4; located between the E, F, and I helices and the β5 sheet). Water was able to penetrate into the active site of the closed structures such as CYP2A6, CYP2D6 or CYP2E1 by the gradual adjustment of the channel walls to enlarge certain parts of the channel while closing others. These movements of the bottleneck are illustrated in Fig. **1**, in which the position of bottleneck move to the larger parts of the channel's topology without affecting the channel's overall direction (Fig. **1**).

CONCLUSIONS

The flexibility of five metabolically active human cytochromes P450 (CYP2A6, CYP2C9, CYP2D6, CYP2E1, and CYP3A4) has been studied using all-atom molecular dynamics simulations with explicit TIP3P solvation on a 100 ns timescale. Particular attention was paid to changes in the active site volume and solvation and to the dynamics of the active site access paths. The widely open structure of rabbit CYP2B4 was also considered, on the 150 ns timescale. The overall topology of the CYP fold is preserved during MD simulations, but the structures relax in response to the simulated conditions (T = 300 K, p = 1 atm, c(enzyme) 3 mM). On aver-

age, the proteins' overall structures equilibrate on a timescale of 50+ ns, with the more rigid cytochromes P450 (CYP2A6 and CYP2E1) relaxing on shorter timescales (< 30 ns), while the more flexible ones relax rather more slowly. The very open CYP2B4 structure closes slowly over the course of the simulation, moving towards a conformation that resembles the classical fold of the microsomal CYPs.

The volumes of the enzymes' active sites fluctuated by more than 50% during the MD simulations. These fluctuations are mainly governed by the movement of main chains; independent side-chain movements were only observed in a handful of amino acid residues in CYP2B4, CYP2D6, CYP2A6 and CYP2C9. The active site volume of CYP2E1 fluctuates very significantly, from 220 to 1310 $Å^3$, because of the opening and closing of 'gates' to two nearby cavities. These movements probably enable CYP2E1 to accommodate small and relatively non-polar substrates (such as organic solvents) as well as considerably longer fatty acids. CYP2E1 also has the least solvated active site of the studied CYPs, in agreement with its preferences for relatively non-polar substrates. The most solvated active sites were those of CYP2B4 and CYP3A4. The CYP2A6 and CYP2E1 active sites are deeply buried, with access paths that are narrower than the radius of a water molecule. However, waters are still able to enter these active sites due to local adaptations of the channel width to accommodate their passage. This finding, together with the fact that access paths bottlenecks of unliganded CYPs fluctuate around 1.5 Å, may imply that the CYPs need not fully open their access paths to admit substrates to the active site; instead, the substrate may induce adaptive conformational changes as it travels down the passage. This may also explain why some substrate recognition sites are localized along the access paths.

ABBREVIATIONS

CYP	=	Cytochrome P450
Rg	=	Radius of gyration
Bf	=	B-factor
RMSD	=	Root-mean-square deviation
MD	=	Molecular dynamics

CONFLICT OF INTEREST

None declared.

ACKNOWLEDGEMENT

The authors gratefully acknowledge support from the GACR (grants 303/09/1001 and 203/09/H046) and Student Project PrF_2011_020 of Palacky University. This work was supported by the Operational Program Research and Development for Innovations - European Social Fund (CZ.1.05/2.1.00/03.0058, CZ.1.05/2.1.00/01.0030 and CZ.1.07/2.3.00/20.0017).

SUPPLEMENTARY MATERIAL

Supplementary material is available on the publishers Web site along with the published article.

REFERENCES

[1] Evans, W. E.; Relling, M. V., Pharmacogenomics: translating functional genomics into rational therapeutics. *Science,* **1999,** *286* (5439), 487-491.

[2] Anzenbacher, P.; Anzenbacherova, E., Cytochromes P450 and metabolism of xenobiotics. *Cell. Mol. Life Sci.,* **2001,** *58* (5-6), 737-747.

[3] Guengerich, F. P., Cytochrome P450 and chemical toxicology. *Chem. Res. Toxicol.* **2008,** *21* (1), 70-83.

[4] Rendic, S., Summary of information on human CYP enzymes: Human P450 metabolism data. *Drug Metab. Rev.,* **2002,** *34* (1-2), 83-448.

[5] Denisov, I. G.; Makris, T. M.; Sligar, S. G.; Schlichting, I., Structure and chemistry of cytochrome P450. *Chem. Rev.,* **2005,** *105* (6), 2253-2277.

[6] Meunier, B.; de Visser, S. P.; Shaik, S., Mechanism of oxidation reactions catalyzed by cytochrome P450 enzymes. *Chem. Rev.,* **2004,** *104* (9), 3947-3980.

[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), 2279-2328.

[8] Sono, M.; Roach, M. P.; Coulter, E. D.; Dawson, J. H., Heme-containing oxygenases. *Chem. Rev.,* **1996,** *96* (7), 2841-2887.

[9] Sansen, S.; Yano, J. K.; Reynald, R. L.; Schoch, G. A.; Griffin, K. J.; Stout, C. D.; Johnson, E. F., Adaptations for the oxidation of polycyclic aromatic hydrocarbons exhibited by the structure of human P450 1A2. *J. Biol. Chem.,* **2007,** *282* (19), 14348-14355.

[10] Wang, A.; Savas, U.; Stout, C. D.; Johnson, E. F., Structural Characterization of the Complex between alpha-Naphthoflavone and Human Cytochrome P450 1B1. *J. Biol. Chem.,* **2011,** *286* (7), 5736-5743.

[11] Sansen, S.; Hsu, M. H.; Stout, C. D.; Johnson, E. F., Structural insight into the altered substrate specificity of human cytochrome P450 2A6 mutants. *Arch. Biochem. Biophys.,* **2007,** *464* (2), 197-206.

[12] DeVore, N. M.; Smith, B. D.; Urban, M. J.; Scott, E. E., Key Residues Controlling Phenacetin Metabolism by Human Cytochrome P450 2A Enzymes. *Drug Metab. Dispos.,* **2008,** *36* (12), 2582-2590.

[13] Yano, J. K.; Denton, T. T.; Cerny, M. A.; Zhang, X. D.; Johnson, E. F.; Cashman, J. R., Synthetic inhibitors of cytochrome P-450 2A6: Inhibitory activity, difference spectra, mechanism of inhibition, and protein cocrystallization. *J. Med. Chem.,* **2006,** *49* (24), 6987-7001.

[14] Yano, J. K.; Hsu, M. H.; Griffin, K. J.; Stout, C. D.; Johnson, E. F., Structures of human microsomal cytochrome P450 2A6 complexed with coumarin and methoxsalen. *Nat. Struct. Mol. Biol.,* **2005,** *12* (9), 822-823.

[15] Smith, B. D.; Sanders, J. L.; Porubsky, P. R.; Lushington, G. H.; Stout, C. D.; Scott, E. E., Structure of the human lung cytochrome P450 2A13. *J. Biol. Chem.,* **2007,** *282* (23), 17306-17313.

[16] Gay, S. C.; Shah, M. B.; Talakad, J. C.; Maekawa, K.; Roberts, A. G.; Wilderman, P. R.; Sun, L.; Yang, J. Y.; Huelga, S. C.; Hong, W. X.; Zhang, Q.; Stout, C. D.; Halpert, J. R., Crystal structure of a cytochrome P450 2B6 genetic variant in complex with the inhibitor 4-(4-chlorophenyl)imidazole at 2.0-A resolution. *Mol. Pharmacol.,* **2010,** *77* (4), 529-538.

[17] Schoch, G. A.; Yano, J. K.; Wester, M. R.; Griffin, K. J.; Stout, C. D.; Johnson, E. F., Structure of human microsomal cytochrome P4502C8 - Evidence for a peripheral fatty acid binding site. *J. Biol. Chem.,* **2004,** *279* (10), 9497-9503.

[18] Schoch, G. A.; Yano, J. K.; Sansen, S.; Dansette, P. M.; Stout, C. D.; Johnson, E. F., Determinants of cytochrome P4502C8 substrate binding - Structures of complexes with montelukast, troglitazone, felodipine, and 9-cis-retinoic acid. *J. Biol. Chem.,* **2008,** *283* (25), 17227-17237.

[19] Williams, P. A.; Cosme, J.; Ward, A.; Angove, H. C.; Vinkovic, D. M.; Jhoti, H., Crystal structure of human cytochrome P450 2C9 with bound warfarin. *Nature,* **2003,** *424* (6947), 464-468.

[20] 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 P4502C9 complexed with flurbiprofen at 2.0-angstrom resolution. *J. Biol. Chem.,* **2004,** *279* (34), 35630-35637.

[21] 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), 7614-7622.

[22] Porubsky, P. R.; Meneely, K. M.; Scott, E. E., Structures of Human Cytochrome P-450 2E1. *J. Biol. Chem.,* **2008,** *283* (48), 33698-33707.

[23] Porubsky, P. R.; Battaile, K. P.; Scott, E. E., Human cytochrome P450 2E1 structures with fatty acid analogs reveal a previously unobserved binding mode. *J. Biol. Chem.,* **2010,** *285* (29), 22282-22290.

[24] Strushkevich, N.; Usanov, S. A.; Plotnikov, A. N.; Jones, G.; Park, H. W., Structural analysis of CYP2R1 in complex with vitamin D-3. *J. Mol. Biol.,* **2008,** *380* (1), 95-106.

[25] Strushkevich, N. V.; Tempel, W.; Gilep, A. A.; Loppnau, P.; Arrowsmith, C. H.; Edwards, A. M.; Bountra, C.; Wilkstrom, M.; Bochkarev, A.; Park, H., Crystal structure of CYP2R1 in complex with 1-alpha-hydroxy-vitamin D2. *To be published.*

[26] Strushkevich, N. V.; Tempel, W.; Gilep, A. A.; Loppnau, P.; Arrowsmith, C. H.; Edwards, A. M.; Bountra, C.; Wilkstrom, M.; Bochkarev, A.; Park, H., Crystal structure of CYP2R1 in complex with vitamin D2. *To be published.*

[27] 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-angstrom resolution. *J. Biol. Chem.*, **2004,** *279* (37), 38091-38094.

[28] 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), 683-686.

[29] Ekroos, M.; Sjogren, T., Structural basis for ligand promiscuity in cytochrome P450 3A4. *Proc. Natl. Acad. Sci. U. S. A.* **2006,** *103* (37), 13682-13687.

[30] Sevrioukova, I. F.; Poulos, T. L., Structure and mechanism of the complex between cytochrome P4503A4 and ritonavir. *Proc. Natl. Acad. Sci. U. S. A.* **2010,** *107* (43), 18422-18427.

[31] Strushkevich, N. V.; Tempel, W.; Dombrovski, L.; Dong, A.; Loppnau, P.; Arrowsmith, C. H.; Edwards, A. M.; Bountra, C.; Wilkstrom, M.; Bochkarev, A.; Park, H., Crystal structure of human CYP7A1. *To be published.*

[32] Strushkevich, N. V.; Usanov, S. A.; Park, H., Crystal structure of human CYP11A1 in complex with 20,22-dihydroxycholesterol. *To be published.*

[33] Ghosh, D.; Griswold, J.; Erman, M.; Pangborn, W., Structural basis for androgen specificity and oestrogen synthesis in human aromatase. *Nature*, **2009,** *457* (7226), 219-U119.

[34] Mast, N.; Whitet, M. A.; Bjorkhem, I.; Johnson, E. F.; Stout, C. D.; Pikuleva, I. A., Crystal structures of substrate-bound and substrate-free cytochrome P450 46A1, the principal cholesterol hydroxylase in the brain. *Proc. Natl. Acad. Sci. U. S. A.* **2008,** *105* (28), 9546-9551.

[35] Mast, N.; Charvet, C.; Pikuleva, I. A.; Stout, C. D., Structural basis of drug binding to CYP46A1, an enzyme that controls cholesterol turnover in the brain. *J. Biol. Chem.*, **2010,** *285* (41), 31783-31795.

[36] Strushkevich, N.; Usanov, S. A.; Park, H. W., Structural Basis of Human CYP51 Inhibition by Antifungal Azoles. *J. Mol. Biol.*, **2010,** *397* (4), 1067-1078.

[37] Chiang, C. W.; Yeh, H. C.; Wang, L. H.; Chan, N. L., Crystal structure of the human prostacyclin synthase. *J. Mol. Biol.*, **2006,** *364* (3), 266-274.

[38] Li, Y. C.; Chiang, C. W.; Yeh, H. C.; Hsu, P. Y.; Whitby, F. G.; Wang, L. H.; Chan, N. L., Structures of prostacyclin synthase and its complexes with substrate analog and inhibitor reveal a ligand-specific heme conformation change. *J. Biol. Chem.*, **2008,** *283* (5), 2917-2926.

[39] Otyepka, M.; Skopalik, J.; Anzenbacherova, E.; Anzenbacher, P., What common structural features and variations of mammalian P450s are known to date? *Biochim. Biophys. Acta*, **2007,** *1770* (3), 376-389.

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

[41] Skopalik, J.; Anzenbacher, P.; Otyepka, M., Flexibility of human cytochromes P450: Molecular dynamics reveals differences between CYPs 3A4, 2C9, and 2A6, which correlate with their substrate preferences. *J. Phys. Chem. B.*, **2008,** *112* (27), 8165-8173.

[42] Hendrychova, T.; Anzenbacherova, E.; Hudecek, J.; Skopalik, J.; Lange, R.; Hildebrandt, P.; Otyepka, M.; Anzenbacher, P., Flexibility of human cytochrome P450 enzymes: molecular dynamics and spectroscopy reveal important function-related variations. *Biochim. Biophys. Acta*, **2011,** *1814* (1), 58-68.

[43] Tokuriki, N.; Tawfik, D. S., Protein Dynamism and Evolvability. *Science*, **2009,** *324* (5924), 203-207.

[44] 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.

[45] Olah, J.; Mulholland, A. J.; Harvey, J. N., Understanding the determinants of selectivity in drug metabolism through modeling of dextromethorphan oxidation by cytochrome P450. *Proc. Natl. Acad. Sci. U. S. A.* **2011,** *108* (15), 6050-6055.

[46] Stjernschantz, E.; Vermeulen, N. P. E.; Oostenbrink, C., Computational prediction of drug binding and rationalisation of selectivity towards cytochromes P450. *Expert Opin. Drug Met.*, **2008,** *4* (5), 513-527.

[47] 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), 5361-5366.

[48] Otyepka, M.; Berka, K.; Anzenbacher, P., Is there a relationship between the substrate preferences and structural flexibility of cytochromes P450? *Curr. Drug Metab.*, **2012**, 130-142.

[49] Krishnamoorthy, N.; Gajendrarao, P.; Thangapandian, S.; Lee, Y.; Lee, K. W., Probing possible egress channels for multiple ligands in human CYP3A4: a molecular modeling study. *J. Mol. Model.* **2010,** *16* (4), 607-614.

[50] 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), 689-696.

[51] Rydberg, P.; Rod, T. H.; Olsen, L.; Ryde, U., Dynamics of water molecules in the active-site cavity of human cytochromes P450. *J. Phys. Chem. B.*, **2007,** *111* (19), 5445-5457.

[52] Skopalik, J.; Anzenbacher, P.; Otyepka, M., Flexibility of human cytochromes P450: molecular dynamics reveals differences between CYPs 3A4, 2C9, and 2A6, which correlate with their substrate preferences. *J. Phys. Chem. B.*, **2008,** *112* (27), 8165-8173.

[53] Seifert, A.; Tatzel, S.; Schmid, R. D.; Pleiss, J., Multiple molecular dynamics simulations of human p450 monooxygenase CYP2C9: the molecular basis of substrate binding and regioselectivity toward warfarin. *Proteins*, **2006,** *64* (1), 147-155.

[54] 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 P4502B4 at 1.6-angstrom resolution. *Proc. Natl. Acad. Sci. U. S. A.* **2003,** *100* (23), 13196-13201.

[55] 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* (5), 797-811.

[56] 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), 584-589.

[57] Li, W. H.; Shen, J.; Liu, G. X.; Tang, Y.; Hoshino, T., Exploring coumarin egress channels in human cytochrome P450 2A6 by random acceleration and steered molecular dynamics simulations. *Proteins*, **2011,** *79* (1), 271-281.

[58] 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), 813-830.

[59] Fishelovitch, D.; Shaik, S.; Wolfson, H. J.; Nussinov, R., Theoretical characterization of substrate access/exit channels in the human cytochrome P450 3A4 enzyme: involvement of phenylalanine residues in the gating mechanism. *J. Phys. Chem. B.*, **2009,** *113* (39), 13018-13025.

[60] Lampe, J. N.; Brandman, R.; Sivaramakrishnan, S.; de Montellano, P. R., Two-dimensional NMR and all-atom molecular dynamics of cytochrome P450 CYP119 reveal hidden conformational substates. *J. Biol. Chem.*, **2010,** *285* (13), 9594-9603.

[61] Brandman, R.; Lampe, J. N.; Brandman, Y.; de Montellano, P. R. O., Active-site residues move independently from the rest of the protein in a 200 ns molecular dynamics simulation of cytochrome P450 CYP119. *Arch. Biochem. Biophys.*, **2011,** *509* (2), 127-132.

[62] Markwick, P. R. L.; Pierce, L. C. T.; Goodin, D. B.; McCammon, J. A., Adaptive Accelerated Molecular Dynamics (Ad-AMD) Revealing the Molecular Plasticity of P450cam. *J. Phys. Chem. Lett.*, **2011,** *2* (3), 158-164.

[63] Park, H.; Lee, S.; Suh, J., Structural and dynamical basis of broad substrate specificity, catalytic mechanism, and inhibition of cytochrome P450 3A4. *J. Am. Chem. Soc.*, **2005,** *127* (39), 13634-13642.

[64] 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), 1602-1611.

[65] Bikiel, D. E.; Boechi, L.; Capece, L.; Crespo, A.; De Biase, P. M.; Di Lella, S.; Lebrero, M. C. G.; Marti, M. A.; Nadra, A. D.; Perissinotti, L. L.; Scherlis, D. A.; Estrin, D. A., Modeling heme proteins using atomistic simulations. *Phys. Chem. Chem. Phys.*, **2006,** *8* (48), 5611-5628.

[66] Santos, R.; Hritz, J.; Oostenbrink, C., Role of Water in Molecular Docking Simulations of Cytochrome P450 2D6. *J. Chem. Inf. Model,* **2010,** *50* (1), 146-154.

[67] Locuson, C. W.; Tracy, T. S., Identification of binding sites of non-I-helix water molecules in mammalian cytochromes P450. *Drug Metab. Dispos.*, **2006,** *34* (12), 1954-1957.

[68] de Graaf, C.; Oostenbrink, C.; Keizers, P. H. J.; van der Wijst, T.; Jongejan, A.; Vemleulen, N. P. E., Catalytic site prediction and virtual screening of cytochrome P450 2D6 substrates by consideration of water and rescoring in automated docking. *J. Med. Chem.*, **2006,** *49* (8), 2417-2430.

[69] Meharenna, Y. T.; Poulos, T. L., Using molecular dynamics to probe the structural basis for enhanced stability in thermal stable cytochromes P450. *Biochemistry*, **2010,** *49* (31), 6680-6686.

[70] Conner, K. P.; Woods, C. M.; Atkins, W. M., Interactions of cytochrome P450s with their ligands. *Arch. Biochem. Biophys.*, **2011,** *507* (1), 56-65.

[71] Li, W. H.; Tang, Y.; Hoshino, T.; Neya, S., Molecular modeling of human cytochrome P450 2W1 and its interactions with substrates. *J. Mol. Graph. Model,* **2009,** *28* (2), 170-176.

[72] Teixeira, V. H.; Ribeiro, V.; Martel, P. J., Analysis of binding modes of ligands to multiple conformations of CYP3A4. *Biochim. Biophys. Acta*, **2010,** *1804* (10), 2036-2045.

[73] Sano, E.; Li, W. H.; Yuki, H.; Liu, X. L.; Furihata, T.; Kobayashi, K.; Chiba, K.; Neya, S.; Hoshino, T., Mechanism of the Decrease in Catalytic Activity of Human Cytochrome P450 2C9 Polymorphic Variants Investigated by Computational Analysis. *J. Comput. Chem.*, **2010,** *31* (15), 2746-2758.

[74] Huang, Q. B.; Szklarz, G. D., Significant Increase in Phenacetin Oxidation on L382V Substitution in Human Cytochrome P450 1A2. *Drug Metab. Dispos.*, **2010,** *38* (7), 1039-1045.

[75] He, L.; He, F.; Bi, H. C.; Li, J. K.; Zeng, S.; Luo, H. B.; Huang, M., Isoform-selective inhibition of chrysin towards human cytochrome P450 1A2. Kinetics analysis, molecular docking, and molecular dynamics simulations. *Bioorg. Med. Chem. Lett.*, **2010,** *20* (20), 6008-6012.

[76] Johnson, E. F.; Sansen, S.; Reynald, R. L.; Stout, C. D., Plasticity of the 2c9 Active Site, Allelic Variation and Ligand Induced Changes. *Drug Metab. Rev.*, **2007,** *39*, 12-12.

[77] Li, W. H.; Ode, H.; Hoshino, T.; Liu, H.; Tang, Y.; Jiang, H. L., Reduced Catalytic Activity of P450 2A6 Mutants with Coumarin: A Computational Investigation. *J. Chem. Theory Comput.*, **2009,** *5* (5), 1411-1420.

[78] Moore, C. D.; Shahrokh, K.; Sontum, S. F.; Cheatham, T. E.; Yost, G. S., Improved Cytochrome P450 3A4 Molecular Models Accurately Predict the Phe215 Requirement for Raloxifene Dehydrogenation Selectivity. *Biochemistry*, **2010,** *49* (41), 9011-9019.

[79] Zhang, H. M.; Kenaan, C.; Hamdane, D.; Hoa, G. H. B.; Hollenberg, P. F., Effect of Conformational Dynamics on Substrate Recognition and Specificity as Probed by the Introduction of a de Novo Disulfide Bond into Cytochrome P450 2B1. *J. Biol. Chem.*, **2009,** *284* (38), 25678-25686.

[80] Hritz, J.; de Ruiter, A.; Ostenbrink, C., Impact of Plasticity and Flexibility on Docking Results for Cytochrome P450 2D6: A Combined Approach of Molecular Dynamics and Ligand Docking. *J. Med. Chem.*, **2008,** *51* (23), 7469-7477.

[81] Berka, K.; Hendrychova, T.; Anzenbacher, P.; Otyepka, M., Membrane Position of Ibuprofen Agrees with Suggested Access Path Entrance to Cytochrome P450 2C9 Active Site. *J. Phys. Chem. A.*, **2011,** *145*(41), 11248-11255.

[82] Cojocaru, V.; Balali-Mood, K.; Sansom, M. S.; Wade, R. C., Structure and dynamics of the membrane-bound cytochrome P450 2C9. **2012**, in press.

[83] Guengerich, F. P., Human Cytochrome P450 Enzymes. In *Cytochrome P450: Structure, Mechanism, and Biochemistry*, 3 ed.; Ortiz de Montellano, P. R., Ed. Kluwer Academic / Plenum Publishers: New York, **2005**; pp 377-530.

[84] Halpert, J. R., The 2010 Bernard B. Brodie Award Lecture. Structure and Function of Cytochromes P450 2B: From Mechanism-Based Inactivators to X-ray Crystal Structures and Back. *Drug Metab. Dispos.*, **2012**, in press.

[85] Ortiz de Montellano (Ed.), P. R., *Cytochrome P450: Structure, Mechanism, and Biochemistry*. 3rd. ed.; Kluwer Academic: New York, 2005.

[86] Guengerich, F. P.; Kim, D. H.; Iwasaki, M., Role of Human Cytochrome-P-450-Iie1 in the Oxidation of Many Low-Molecular-Weight Cancer Suspects. *Chem. Res. Toxicol.*, **1991,** *4* (2), 168-179.

[87] Lucas, D.; Ferrara, R.; Gonzales, E.; Albores, A.; Manno, M.; Berthou, F., Cytochrome CYP2E1 phenotyping and genotyping in the evaluation of health risks from exposure to polluted environments. *Toxicol. Lett.*, **2001,** *124* (1-3), 71-81.

[88] Zgarbova, M.; Otyepka, M.; Sponer, J.; Hobza, P.; Jurecka, P., Large-scale compensation of errors in pairwise-additive empirical force fields: comparison of AMBER intermolecular terms with rigorous DFT-SAPT calculations. *Phys. Chem. Chem. Phys.*, **2010,** *12* (35), 10476-10493.

[89] Rose, P. W.; Beran, B.; Bi, C. X.; Bluhm, W. F.; Dimitropoulos, D.; Goodsell, D. S.; Prlic, A.; Quesada, M.; Quinn, G. B.; Westbrook, J. D.; Young, J.; Yukich, B.; Zardecki, C.; Berman, H. M.; Bourne, P. E., The RCSB Protein Data Bank: redesigned web site and web services. *Nucleic Acids Res.*, **2011,** *39*, D392-D401.

[90] Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Kollman, P. A., Application of Resp Charges to Calculate Conformational Energies, Hydrogen-Bond Energies, and Free-Energies of Solvation. *J. Am. Chem. Soc.*, **1993,** *115* (21), 9620-9631.

[91] Wang, J. M.; Cieplak, P.; Kollman, P. A., How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J. Comput. Chem.*, **2000,** *21* (12), 1049-1074.

[92] Dundas, J.; Ouyang, Z.; Tseng, J.; Binkowski, A.; Turpaz, Y.; Liang, J., CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues. *Nucleic Acids Res.*, **2006,** *34* (Web Server issue), W116-8.

[93] Ho, B. K.; Gruswitz, F., HOLLOW: Generating Accurate Representations of Channel and Interior Surfaces in Molecular Structures. *BMC Struct. Biol.*, **2008,** *8*, -.

[94] Petrek, M.; Kosinova, P.; Koca, J.; Otyepka, M., MOLE: a Voronoi diagram-based explorer of molecular channels, pores, and tunnels. *Structure*, **2007,** *15* (11), 1357-1363.

[95] Malaiyandi, V.; Sellers, E. M.; Tyndale, R. F., Implications of CYP2A6 genetic variation for smoking behaviors and nicotine dependence. *Clin. Pharmacol. Ther.*, **2005,** *77* (3), 145-158.

[96] Wilderman, P. R.; Shah, M. B.; Liu, T.; Li, S.; Hsu, S.; Roberts, A. G.; Goodlett, D. R.; Zhang, Q.; Woods, V. L., Jr.; Stout, C. D.; Halpert, J. R., Plasticity of cytochrome P450 2B4 as investigated by hydrogen-deuterium exchange mass spectrometry and X-ray crystallography. *J. Biol. Chem.*, **2010,** *285* (49), 38602-38611.

[97] Black, S. D., Membrane topology of the mammalian P450 cytochromes. *FASEB J.*, **1992,** *6* (2), 680-685.

Received: June 6, 2011 **Revised: August 1, 2011** **Accepted: August 1, 2011**

Supplementary Material

Figures and tables showing the evolution of the enzymes' RMSD and radius of gyration over time, B-factors for each system studied, RMSD data for active site amino acids, solvation of each CYP structure, and data on the shapes and geometries of the active site cavities are included in the supplemental documentation. The supplemental movie shows the closing motion of CYP2B4 during the simulation.

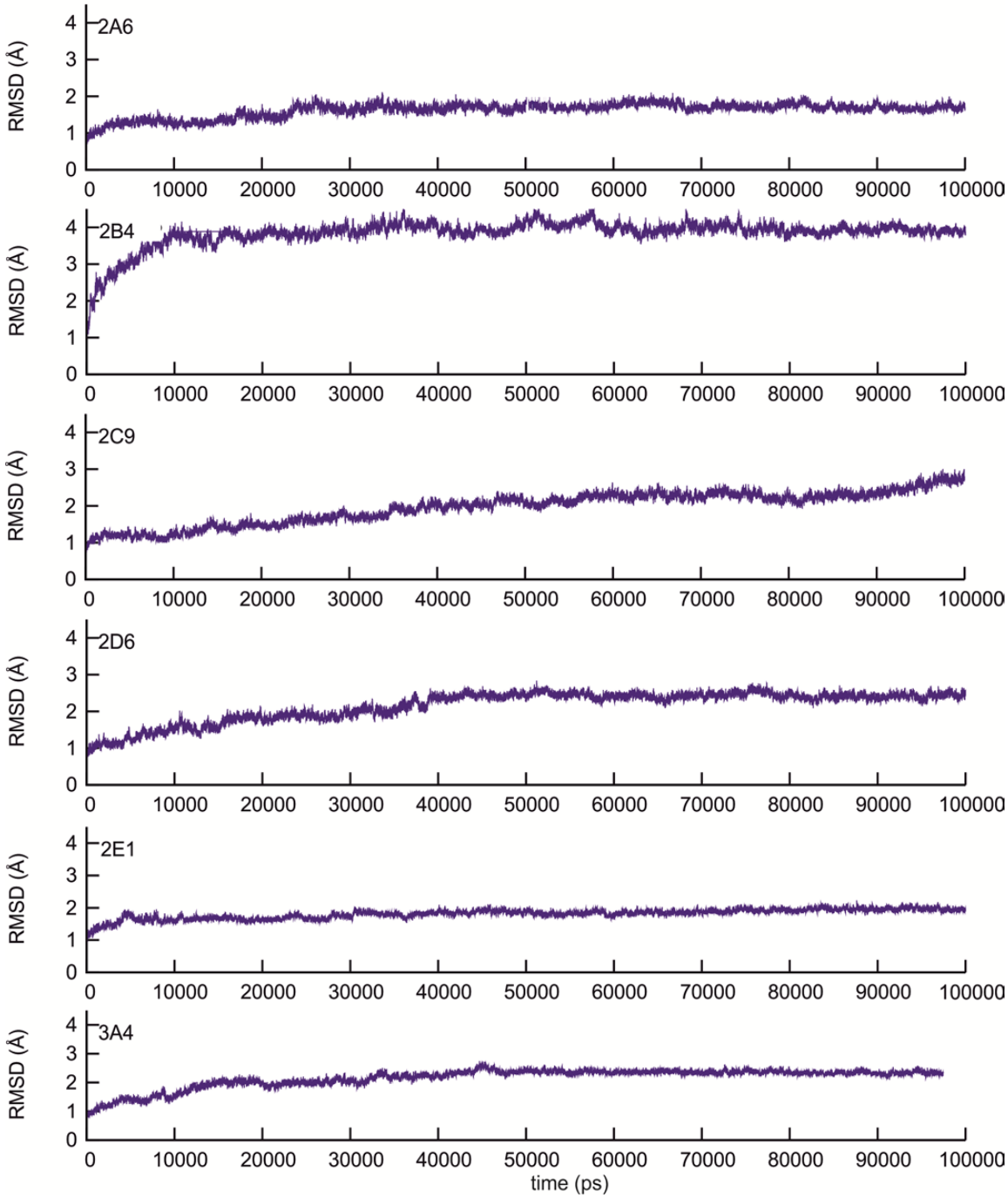


Fig. (S1). Evolution of RMSD of CYPs during the simulation. The typical overall cytochrome P450 fold was conserved in all simulations. Most structures achieved RMSD equilibration after 50 ns. However each structure achieved equilibration in a different timescale. There were two extremes: CYP2E1 equilibrated rather quickly within 10 ns, whereas RMSD of CYP2C9 did not reached equilibrium even in 100 ns.

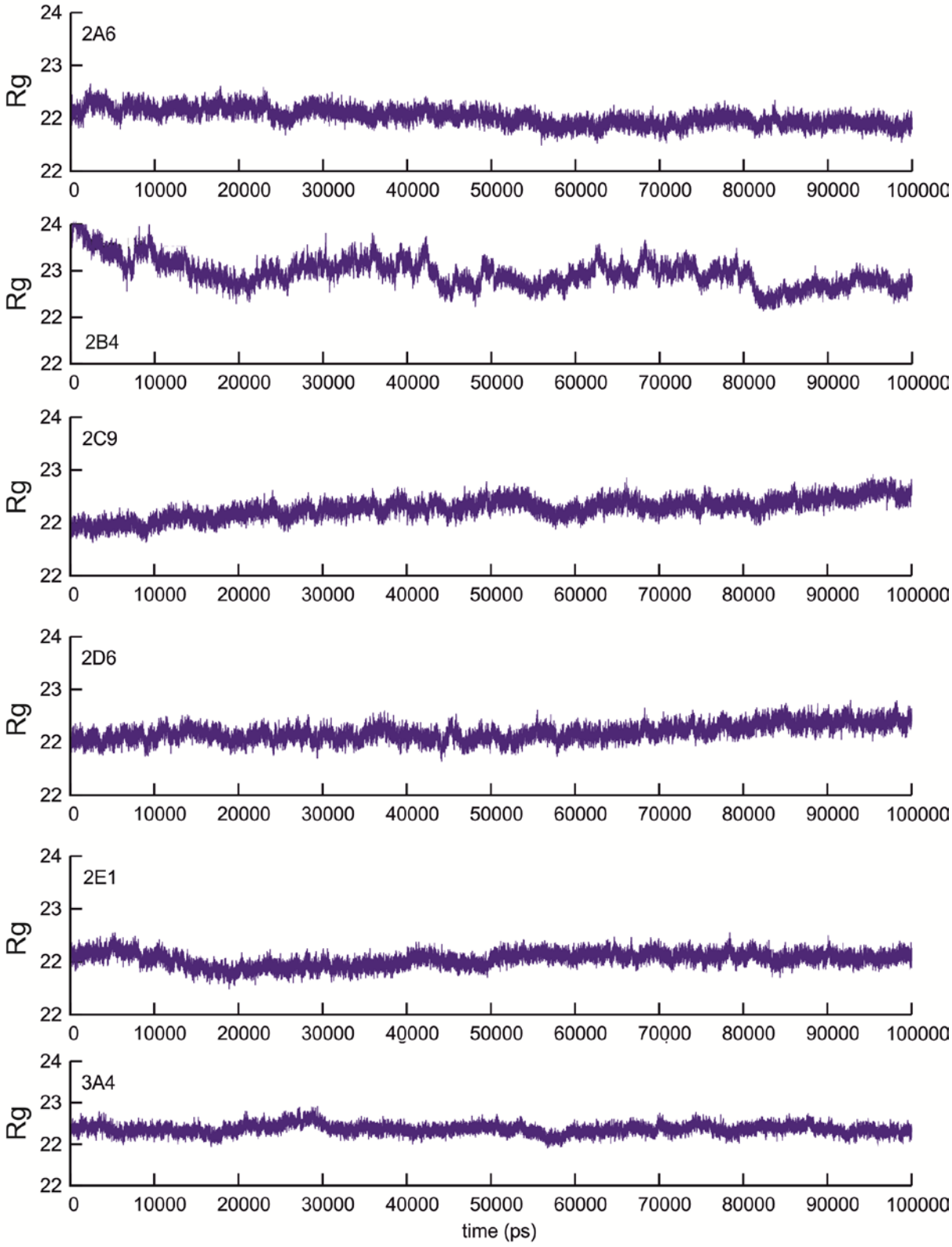


Fig. (S2). Evolution of radius of gyration of CYPs during the simulation. Radius of gyration (Rg) showed that CYPs starting from the closed structure have tendency to stay in closed conformation (3A4, 2A6, 2E1) or to slightly open (2C9, 2D6). The widely open structure of 2B4 was slowly closing during the simulation towards the closed structure. Closing of 2B4 can be seen in a Movie M1 or at http://www.molmovdb.org/cgi-bin/morph.cgi?ID=980050-28395.

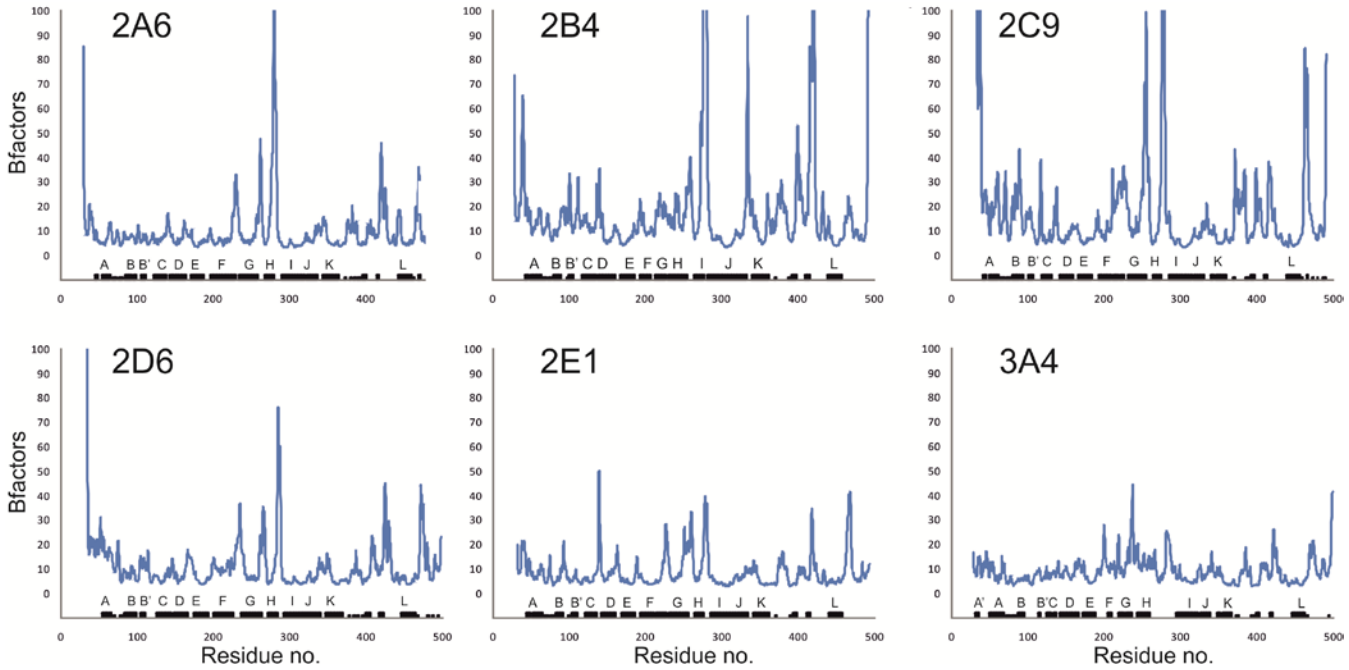


Fig. (S3). Temperature B-factors of each system counted on last 5 ns of the 100ns time scale. Common secondary structures elements, especially alpha helices, are depicted by black bricks below each Bf curves.

Table ST1. RMSD of Active Site Residues

2A6				**2B4**				**2C9**				**2D6**				**2E1**				**3A4**			
aa	Cα	SC	diff	aa	Cα	SC	diff	aa	Cα	SC	diff	aa	Cα	SC	diff	aa	Cα	SC	diff	aa	Cα	SC	diff
F107	0,7	0,8	0,1	*V104*	2,8	3,3	0,5	*A103*	2,2	2,2	0,0	*I106*	2,2	2,6	0,4	*F106*	1,4	1,4	0,0	*F108*	1,8	1,8	-0,1
V117	0,5	0,5	0,0	*I114*	3,8	3,8	0,0	*V113*	1,1	1,1	-0,1	*V119*	1,2	1,6	0,4	*I115*	1,0	1,0	0,0	*S119*	1,3	1,2	-0,1
F118	0,6	0,7	0,1	*F115*	4,3	4,5	0,2	*F114*	1,4	1,8	0,4	*F120*	1,0	1,3	0,3	*F116*	1,2	1,4	0,2	*I120*	1,3	1,4	0,1
F209	0,6	1,4	0,7	*F206*	2,0	3,1	1,2	*I205*	2,0	2,2	0,2	*L213*	2,1	2,5	0,4	*F207*	0,8	0,8	0,0	*R212*	1,4	1,5	0,1
I300	0,6	0,8	0,2	*F297*	1,0	1,8	0,7	*G296*	1,1	*1,1*	0,0	*S304*	1,1	1,2	0,2	*F298*	0,9	0,9	0,0	*F304*	0,9	1,0	0,1
G301	0,6	*0,6*	0,0	*A298*	1,0	1,1	0,1	*A297*	0,7	0,8	0,1	*A305*	0,9	1,1	0,1	*A299*	0,9	1,1	0,1	*A305*	0,9	0,9	0,1
E304	0,5	0,7	0,2	*E301*	0,8	1,3	0,5	*E300*	0,8	1,1	0,3	*V308*	0,7	0,9	0,2	*E302*	0,7	0,9	0,2	*E308*	0,9	1,2	0,3
T305	0,5	0,6	0,1	*T302*	0,7	0,7	0,0	*T301*	0,6	0,6	0,0	*T309*	0,6	0,7	0,1	*T303*	0,6	0,9	0,3	*T309*	0,7	0,8	0,1
I366	0,6	0,6	0,0	*I363*	0,9	0,9	0,0	*L362*	0,9	0,8	-0,1	*V370*	0,9	1,0	0,2	*V364*	0,7	0,8	0,0	*A370*	0,8	0,9	0,1
L370	0,7	1,0	0,2	*V367*	1,3	1,3	0,1	*L366*	1,0	1,0	0,0	*M374*	1,0	1,3	0,3	*L368*	0,8	0,8	0,0	*E374*	1,1	1,3	0,2
F480	1,3	1,4	0,2	*V477*	1,4	1,7	0,3	*F476*	1,3	1,9	0,7	*F483*	1,2	2,5	1,3	*F478*	1,2	1,5	0,3	*L482*	1,2	1,5	0,3
SUM	7,2	9,0	1,7		20,0	23,5	3,6		13,0	14,5	1,6		12,8	16,6	3,8		10,3	11,4	1,1		12,3	13,5	1,2

Columns Cα denote RMSD of main chain of residues identified in all CYP active sites over 100 ns. Columns SC show RMSD of side chain centre. Diff is the difference between the two. Side chain motions are thought to be independent when diff is greater than 0.4 Å. Such residues are shown in bold.

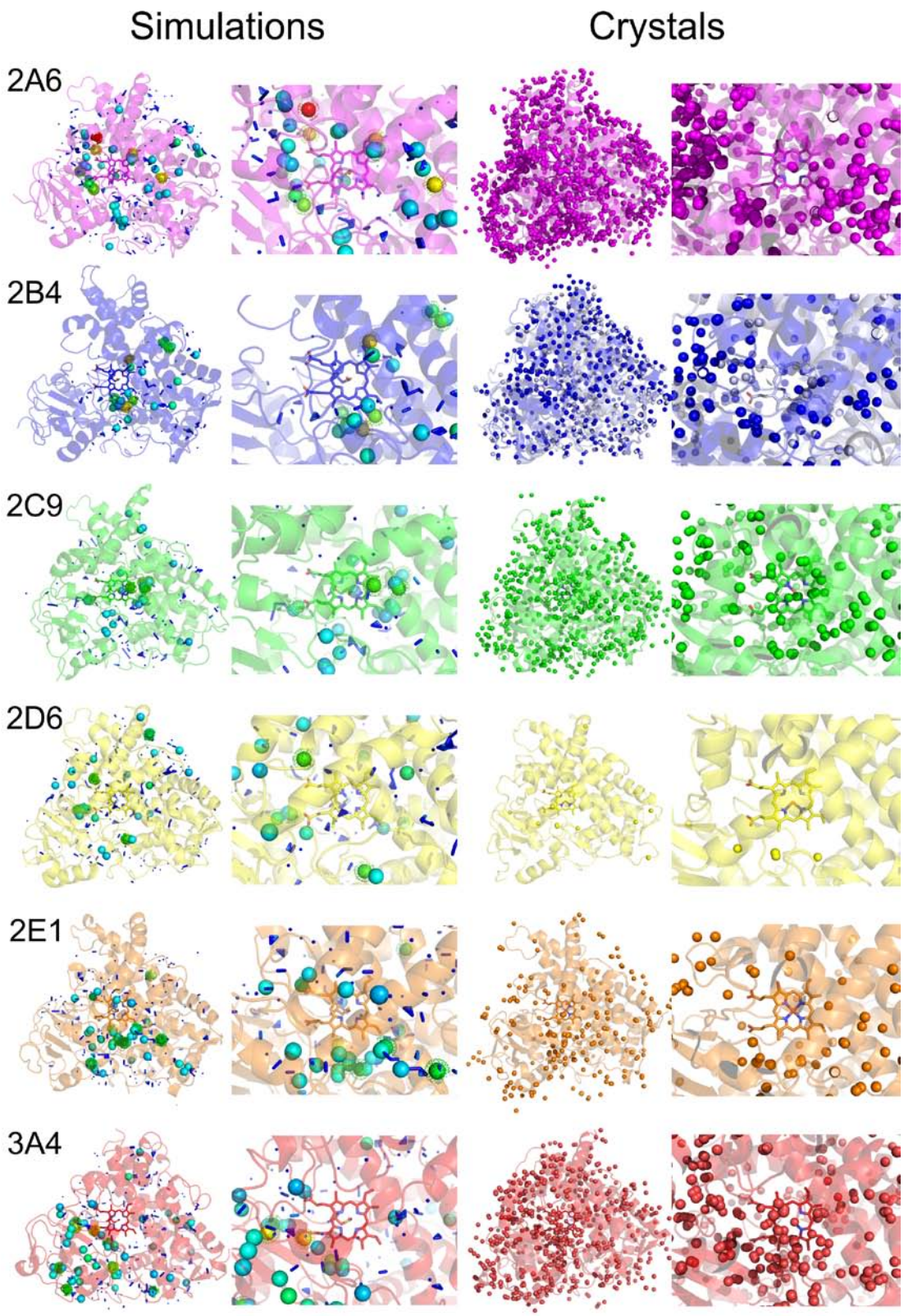


Fig. (S4). Details of solvation of CYPs in simulation and in crystals.

Left panel shows water positions as detected by simulations. Small blue dots and sticks represents positions of slightly enlarged water occupation (10 – 25 ns), bigger spheres represents positions with enlarged water presence (>25 ns), while warmer colors (yellow, red) are used to represent for more occupied water positions. The most occupied water positions (>35 ns) are further emphasized by dots. Waters in the crystal structures are shown for all known CYP structures for specified CYP aligned onto structures used for the start of the simulations. Waters in closed (3MVR) and open (1PO5) crystal structures of CYP2B4 are shown in blue/light blue respectively.

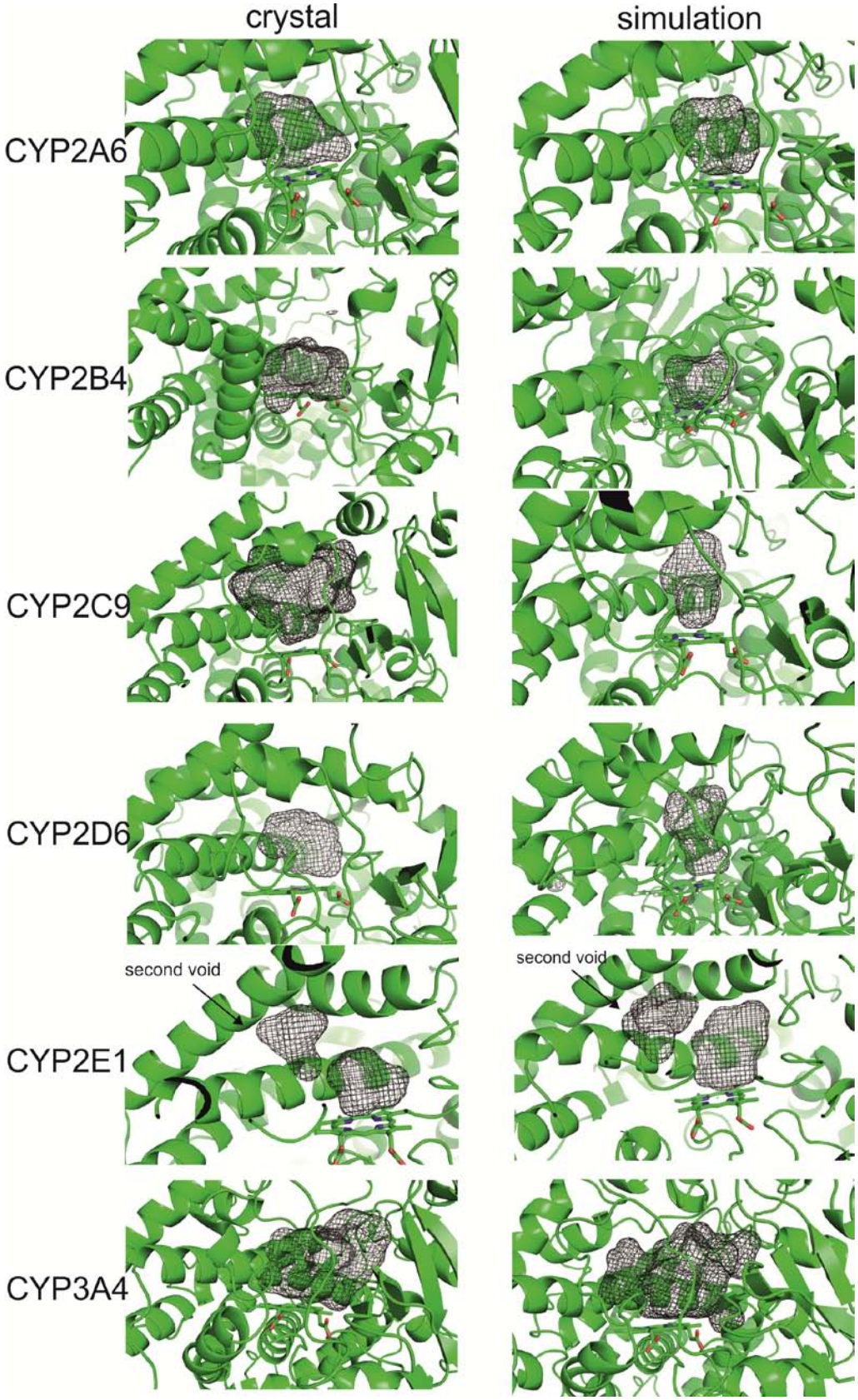


Fig. (S5). The active site cavities of CYPs. Cavities are depicted in black mesh in each system. Left panel shows cavities in the X-ray structures and right panel shows cavities in the structure from where in the B column there are 100ns snaps.

Table ST2. Van Der Waals Volumes of Selected Ligands of Human CYPs.

CYP	Ligand	V [Å^3]	CYP	Ligand	V [Å^3]
2A6	1,3-butadiene	64	**2E1**	ethanol	54
	chloroform	70		acetone	64
	tertbutyl methyl ether	106		1,1,1-trichlorethane	87
	nitrosodiethylamine	107		halothane	93
	coumarin	126		indole	109
	tegafur	160		chlorzoxazone	124
	nicotine	164		acetaminophen	138
	clozapine	287		lauric acid	226
2C9	ibuprofen	212	**3A4**	etinylestradiol	292
	flurbiprofen	219		nifedipine	300
	phenytoin	220		clotrimazole	307
	diclofenac	236		lovastatin	402
	propranolol	258		sildenafil	421
	piroxicam	268		ketoconazole	451
	(R)-warfarin	276		indinavir	592
	linoleic acid	312		erythromycin	727
2D6	debrisoquine	166		paclitaxel	764
	bufuralol	263		cyclosporin	1218
	tramadol	272			
	fluoxetin	274			
	metoprolol	274			
	dextromethorphan	275			
	fluvoxamin	285			
	quinidine	310			

Van der Waals volumes calculated with www.chemicalize.org

Biochimica et Biophysica Acta 1803 (2010) 858–864

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbamcr

A cancer-derived mutation in the PSTAIRE helix of cyclin-dependent kinase 2 alters the stability of cyclin binding

Emma S. Child [a], Tereza Hendrychová [b], Karen McCague [c], Andy Futreal [d], Michal Otyepka [b], David J. Mann [a,c,*]

[a] Department of Life Sciences, Imperial College, South Kensington, London SW7 2AZ, UK
[b] Department of Physical Chemistry, Faculty of Science, Palacky University, Olomouc, Czech Republic
[c] Chemical Biology Centre and Department of Chemistry, Imperial College, South Kensington, London SW7 2AZ, UK
[d] Cancer Genome Project, Wellcome Trust Sanger Institute, Hinxton, CB10 1SA, UK

ARTICLE INFO

Article history:
Received 23 November 2009
Received in revised form 25 March 2010
Accepted 7 April 2010
Available online 24 April 2010

Keywords:
Cyclin-dependent kinase
cdk
PSTAIRE
p21Cip1
p27Kip1
cdc28

ABSTRACT

Cyclin-dependent kinase 2 (cdk2) is a central regulator of the mammalian cell cycle. Here we describe the properties of a mutant form of cdk2 identified during large-scale sequencing of protein kinases from cancerous tissue. The mutation substituted a leucine for a proline in the PSTAIRE helix, the central motif in the interaction of the cdk with its regulatory cyclin subunit. We demonstrate that whilst the mutant cdk2 is considerably impaired in stable cyclin association, it is still able to generate an active kinase that can functionally complement defective cdks in vivo. Molecular dynamic simulations and biophysical measurements indicate that the observed biochemical properties likely stem from increased flexibility within the cyclin-binding helix.

© 2010 Elsevier B.V. All rights reserved.

1. Introduction

The central regulators of mammalian cell proliferation are the cyclin-dependent kinases (cdks) that phosphorylate key substrates in a sequential manner to facilitate orderly progression through the cell cycle [1,2]. Commitment to division is determined in the G1 phase of the cell cycle, a time when the predominantly active cdks are cdk2 and cdk4 [2]. These cdks are activated sequentially in part through interaction with their regulatory subunits, the D-type cyclins binding to cdk4 early in G1 phase, with E-type cyclins associating with cdk2 in mid-G1, and with A-type cyclins combining with cdk2 in S phase [2]. Genetic studies have largely validated this view although they have also uncovered considerable functional redundancy amongst the cdks [1].

Association of cyclins with their cognate cdk subunit typically involves interaction over a large surface area that is centred on the so-called PSTAIRE helix of the cdk [3,4], this being equivalent to the C helix in protein kinase A [5]. In associating with cdk2, E-type cyclins cause an essential rearrangement of the PSTAIRE helix to a conformation favouring the full activation of the cdk (the so-called 'helix in' arrangement) [6]. Together with the phosphorylation of a key threonine in the activation segment (also known as the T loop) adjacent to the catalytic site, the PSTAIRE helix rearrangements upon cyclin binding lead to the realignment of essential ATP binding residues into an active conformation.

Because of their central regulatory role in proliferative control, cdks (along with other signal transducing protein kinases) have become the focus of much interest in the study of diseases affecting proliferation, such as cancer. In order to gain insight into mutations that may affect kinase function, large-scale sequencing projects have been directed at identifying abnormalities in the fraction of the genome encoding the protein kinases in cancerous tissue [7]. Such studies have the power to uncover many mutations in focussed screens of interesting candidate genes although they do not in themselves address the functional consequences of detected mutations. Here we describe the characterisation of the properties of one missense mutation in cdk2 detected by the Cancer Genome Project that alters the proline of the PSTAIRE helix to a leucine in a glioblastoma cell line.

2. Materials and methods

2.1. Antibodies

The following antibodies were used: cyclin E1 (HE12), p21 (C-19), p27 (C-19) and cdk2 (sc-163) from Santa Cruz and anti-Flag (M2)

* Corresponding author. Department of Life Sciences, Imperial College, South Kensington, London SW7 2AZ, UK. Tel.: +44 20 7594 5302.
E-mail address: d.mann@imperial.ac.uk (D.J. Mann).

0167-4889/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.bbamcr.2010.04.004

antibody from Sigma. Anti-cyclin A2 (E65.1) was from Cancer Research UK. HRP-conjugated secondary antibodies (mouse, rabbit and goat) were from Jackson ImmunoResearch.

2.2. Plasmids and viruses

The P45L mutant cdk2 was generated by site-directed mutagenesis and, after sequence verification, subcloned into pBlueBacHis A (Invitrogen). This was transfected into Sf9 cells with linearised viral DNA (BaculoGold, BD Biosciences) and recombinant virus amplified to generate high titre baculovirus stock. The P45L mutant cdk2 was also cloned into pcDNA3 for mammalian expression. In addition, the P45L mutant cdk2 was cloned into the yeast two-hybrid vector pGBDU-C1 [8]. Control constructs/viruses were created with wild type cdk2. Other recombinant baculoviruses encoding full-length cyclins E1 and A2 have been described [9]. Coding sequences for p21 and p27 were cloned into pRSET and expressed in bacteria. Proteins were purified by metal immunoaffinity purification following treatment of sonicated lysates at 65 °C for 10 min and removal of the precipitated material.

2.3. Cell culture

Sf9 cells were grown in Graces insect media (PAA) supplemented with 10% foetal bovine serum (PAA) and maintained at 27 °C. The Sf9 cells were infected with baculoviruses directing the expression of wild type or P45L mutant cdk2, cyclin A2, cyclin E1 and/or wild type virus at multiplicities of infection greater than 10:1. U2OS cells were grown in DMEM supplemented with 10% foetal bovine serum in a humidified atmosphere containing 10% carbon dioxide in air at 37 °C. The cells were transfected with pcDNA3-cdk2 and cdk2 P45L by calcium phosphate precipitation [9]. The transfected cells were washed 24 h later and after a further 24 h the cells were harvested.

2.4. Yeast methods

Saccharomyces cerevisiae strain Y246a (*MATa cdc28-4 trp1 ura3-52 tyr1*; a gift from John Diffley) was transformed with plasmids using lithium acetate. Transformants were plated on selective media and incubated at permissive (30 °C) or restrictive (35 °C) temperatures for three days.

2.5. Kinase assays

Sf9 cells infected with the appropriate recombinant baculoviruses were lysed 72 h post infection by hypotonic lysis using kinase buffer (25 mM Hepes pH7.9/5 mM MgCl/0.1% 2-mercaptoethanol/0.1 mM EDTA) for 10 min on ice. Cell debris was removed by centrifugation. The cell lysate was used in kinase assays as described below. In immunoprecipitation kinase assays, immunoprecipitations were performed at 4 °C for 2 h using protein A-Sepharose beads, cdk2 antibody (0.4μg) in PBS/0.1% Tween 20. The beads were washed three times in PBS/0.1% Tween 20, followed by two further washes in kinase buffer. Kinase assays were performed in kinase buffer containing 0.5–1 μg Histone H1 or GST-pRb, 100 μM ATP and 2.5 μCi [γ^{32}P]-ATP and incubated for 30 min at 30 °C. Reactions were terminated by the addition of an equal volume of 2× SDS-PAGE loading buffer and resolved by SDS-PAGE and then subject to autoradiography.

2.6. Immunoprecipitations

Cells were washed in PBS and then lysed in 0.1% Tween 20/150 mM NaCl/1 mM DTT/50 mM Tris–HCl pH8.0/1 mM EDTA/2.5 mM EGTA/50 mM NaF at 4 °C by passing the lysate five times through a 21 gauge needle. Debris was removed by centrifugation. Lysates were then immunoprecipitated as described above. The precipitates were washed three times in lysis buffer, resuspended in 2× SDS-PAGE loading buffer and immunoblotted with detection via chemiluminescence visualised on Fuji Image Reader LAS-3000.

2.7. Circular dichroism

A Chirascan spectropolarimeter (Applied Photphysics) was used to measure the CD spectra and temperature-dependent protein unfolding profiles [10]. A 0.1 cm path length was used with a protein concentration of 0.1 mg/ml in buffer containing 10 mM Tris–HCl pH 8.0/150 mM NaCl/10% glycerol/1 mM DTT. T_m was determined at 230 nm with temperature range from 20 to 80 °C (recorded every 0.5 °C).

2.8. Thermal shift assays

Sypro Orange™ (Invitrogen) was added to a solution of cdk2 (5–10 μg) in 20 or 40 mM Tris–HCl pH8.0/150 mM NaCl/1 mM DTT/5% glycerol (total volume 95 μl) in a 96-well plate. The samples were heated from 20 to 90 °C at intervals of 0.5 °C, each held for 10 s on an IQ5 Real-Time PCR Detection System (BioRad). The fluorescence intensity was measured at excitation/emission wavelengths of 485/530 nm and the T_m was calculated as the temperature of maximum inflection of the melting curve.

2.9. Computational methods

Molecular dynamics simulations were carried out using the program PMEMD (from AMBER 9.0 package) [11] with the *parm99* and *ff03* force fields [12,13]. The *ff03* force field is the most recent version of AMBER force field with revised ψ/φ torsion parameters and recalculated atomic partial charges from the electrostatic potential in the continuum ($\varepsilon_r = 4$) produced by the B3LYP/cc-pVTZ method. The *parm99* force field does not accurately represent glycine behavior and that is prone to other inaccuracies including over-stabilisation of α-helical peptide conformations and over-estimation of β-bend propensity. Nonetheless, both force fields are believed to perform well on compact and folded protein structures. The cdk2 starting structure (PDB code: 2CLX, resolution 1.8 Å with the ligand removed from the active site [14]) and the starting structure for pT160-cdk2/ATP complexed with Cyclin E was taken from PDB (PDB code: 1W98, resolution 2.15 Å [6], where ATP and Mg^{2+} was added from 1JST structure [15]). The applied MD simulation protocol, which has been repeatedly successfully used to study cdk2 systems [16–19], was as follows. First, the protonation states of all histidines were checked using WHATIF [20] to create an optimal H-bond network. All hydrogens were added using the Xleap program from the AMBER 9.0 package. The structures were neutralized by adding 4 Cl^- counter ions to the monomeric cdk2 systems and 7 Cl^- counter ions to the systems complexed with cyclin E. Each system was solvated in a rectangular water box with a layer of water molecules 9 Å thick. The energy of each system was then minimized as follows prior to the main molecular dynamics simulation run. The protein was frozen and the solvent molecules and counterions were allowed to move during a 1000 step minimization and a 10 ps long molecular dynamics run under NpT conditions. The side chains were then relaxed using several sequential minimizations with the force constants applied to the backbone atoms being decreased in each run. Following this relaxation, the system was 20 ps heated from 10 to 50 K, then 70 ps from 50 to 298.16 K and the thermalization was ended by 10 ps warming at 298.15 K. The production phases were run for 10 ns for all systems. Free cdk2 system studied comprised ~35 000 atoms and cdk2-cyclin E systems comprised ~70 000 atoms. The simulation period was chosen as a compromise between the quality of conformation space sampling and the calculation length. Time integration steps of 2 fs were used, together with particle-mesh Ewald (PME) methods for electrostatic interactions. All simulations were run under periodic boundary conditions in the NpT ensemble at 298.16 K and at a constant pressure of 1 atm. The SHAKE algorithm

with a tolerance of 10^{-5} Å was applied to fix all bonds containing hydrogen atoms. Non-bonding interactions were subject to a 9.0 cut-off. Coordinates were stored every 2 ps. All analyses of the MD simulations were carried out using the PTRAJ modules of AMBER 9.0.

3. Results

During large-scale sequencing to identify mutations in protein kinases in human cancer cells and primary tissue, we identified a heterozygous missense mutation in the cdk2 gene in a short-term culture of a glioblastoma (data not shown). The mutation was in the cdk2 kinase at position 45, altering the conserved proline of the cyclin-binding PSTAIRE motif to a leucine (P45L, hereafter). This motif is characteristic of many cdks and is central to the binding of the activating cyclin subunit to the kinase [4]. Inspection of the crystal structures of cdk2/cyclin E1 [6] revealed that the proline side chain was oriented away from the cyclin subunit (not shown) making direct side chain effects on cyclin binding unlikely. However, given the importance of this region in regulating cdk activity, the presence of this mutation in a cancer cell line, the constraints imposed by proline in terms of peptide backbone angles and the non-conservative nature of the substitution involved, we undertook to characterise the properties of this P45L cdk2 to determine if they could be contributory to deregulated cell cycle control.

Initially, we constructed a recombinant baculovirus directing expression of the mutant form of cdk2. This virus produced somewhat lower levels of recombinant protein than the wild type cdk2-encoding virus and the protein displayed a slightly enhanced mobility on SDS-PAGE (Fig. 1A) perhaps indicative of differences in retained secondary structure during electrophoresis. Co-infection of Sf9 cells with either recombinant baculovirus directing expression of wild type or P45L cdk2 and either an A- or E-type cyclin was performed and the kinase activity of the cell lysates assessed. Fig. 1B demonstrates that both forms of cdk2 generated kinase activity towards both Histone H1 and the Retinoblastoma protein, Rb. The mutant enzyme produced slightly less kinase activity than the wild type but this parallelled its reduced expression. The activities of both wild type and P45L cdk2 in complex with cyclin E1 (or cyclin A2) were equally sensitive to the cdk inhibitors p21 and p27 when titrated into the in vitro assays (Fig. 1C and D and Fig. S1).

To further demonstrate the functionality of the P45L mutant cdk2, we expressed it in *S. cerevisiae* containing a temperature sensitive allele of *cdc28* and tested for complementation at the restrictive temperature. As shown in Fig. 1E, P45L cdk2 enabled proliferation at the non-permissive temperature although proliferation was generally less robust than that of yeast expressing wild type cdk2 in agreement with previous observations [21]. Taken together, these experiments demonstrate that the P45L mutation in cdk2 is seemingly well tolerated and the kinase produced had gross properties that were similar to its wild type counterpart being functional both in vitro and in vivo.

Whilst the data suggested that the P45L cdk2 was functional, this mutant version of the enzyme appeared somewhat less efficient that the wild type equivalent in terms of expression in Sf9 cells, kinase

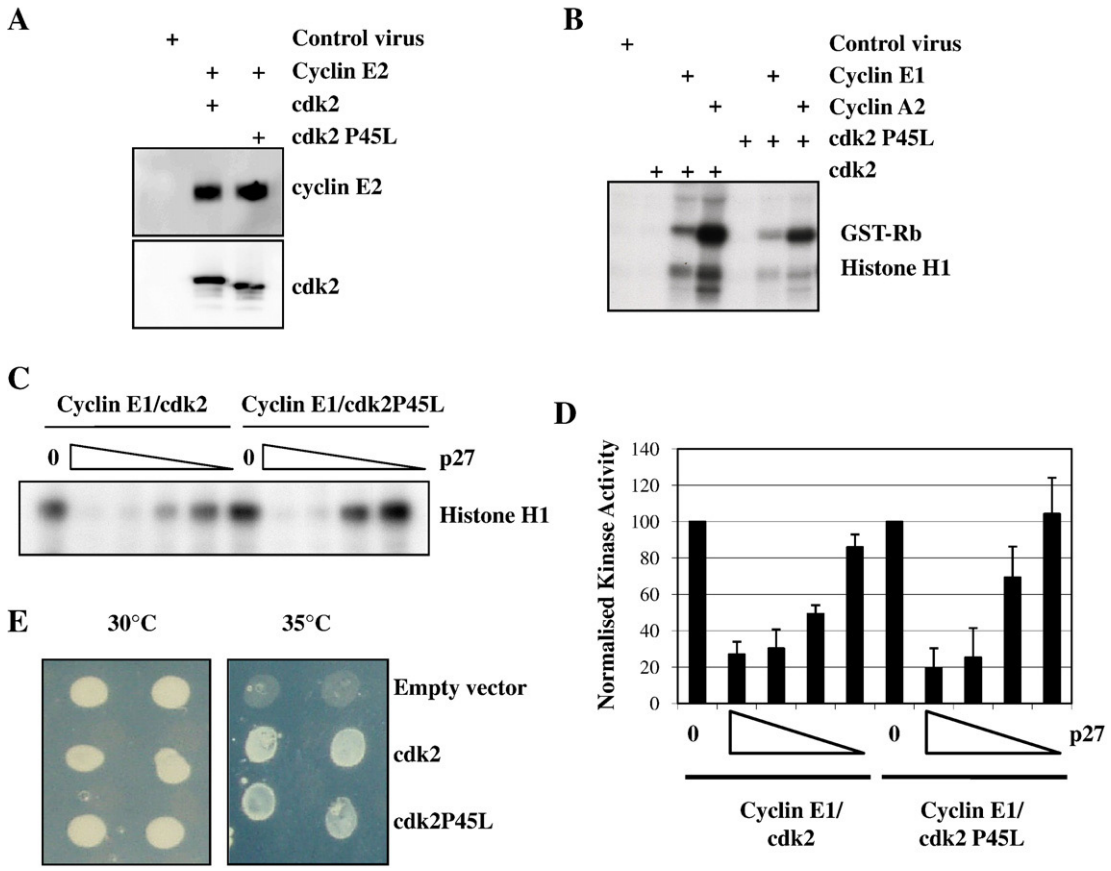


Fig. 1. P45L cdk2 is functional as a protein kinase. A Sf9 cells were co-infected with recombinant baculoviruses encoding wild type cdk2 or P45L cdk2 plus cyclin E1 or were infected with non-recombinant baculovirus. Cell lysates were immunoblotted as indicated. B Sf9 cells were infected with the indicated recombinant baculoviruses, lysed after 72 h and assayed for kinase activity against both GST-Rb and Histone H1 using [γ-^{32}P]ATP followed by autoradiography. C and D Lysates as described in A were incubated with increasing concentrations of p27 purified from recombinant bacteria for 10 min at room temperature and then incubated for 30 min at 30 °C in the presence of Histone H1 and [γ-^{32}P]ATP. 0 denotes assays performed in the absence of p27. Assays were performed in triplicate and quantified by Phosphorimage analysis. The bar chart shows the means and standard deviation with activities normalised to the uninhibited control. E *S. cerevisiae* with a temperature sensitive allele of *cdc28* were transformed with plasmids expressing either wild type or P45L cdk2 or the parental vector and tested for their ability to proliferate at 30 °C and the restrictive 35 °C. Two individual transformants are shown for each plasmid.

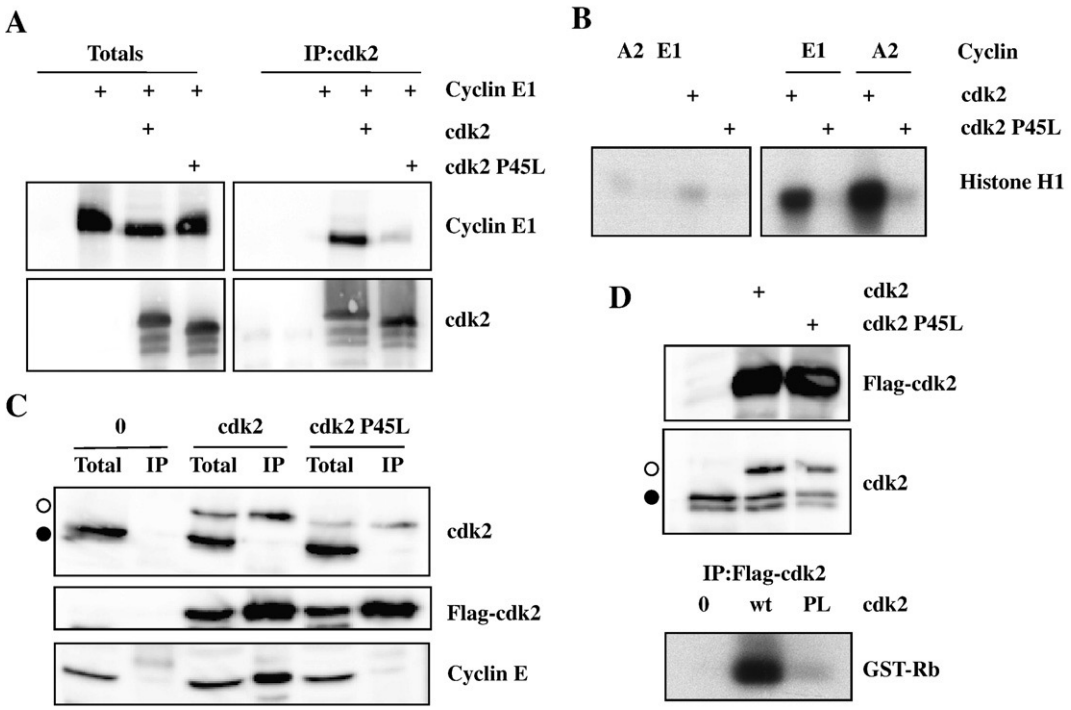


Fig. 2. P45L cdk2 fails to stably associate with cyclins. A and B Sf9 cells were infected with the indicated recombinant baculoviruses, lysed after 72 h and immunoprecipitated through cdk2. Total cell lysates and immunoprecipitates were immunoblotted as indicated (A) or immunoprecipitates were used in kinase assays with Histone H1 as substrate (B). C and D U2OS cells were transfected with Flag-tagged wild type or P45L cdk2 and cell lysates were harvested and immunoprecipitations were performed through the Flag tag of the cdk2. Total lysates (Total) and immunoprecipitates (IP) were immunoblotted for the Flag epitope and total cdk2 showing exogenous (white circle) and endogenous (black circle) expression and cyclin E. Kinase assays against GST-Rb were also performed (D, lower panel).

activity and rescue of the phenotype in cdc28^{ts} yeast. In order to understand these observations further, we investigated the interaction of cyclins with this mutant cdk. Initially we used the baculovirus system to generate cell lysates expressing both cdk2 (wild type or P45L mutant) and cyclin E1. Lysates were immunoprecipitated through cdk2 and immunoblotted for the associated cyclin. As shown in Fig. 2A, cyclin E1 was readily detected in association with wild type cdk2 but only present at very low levels in the P45L cdk2 immunoprecipitates despite being present in the initial lysate at levels similar to the wild type protein. Further, immunoprecipitation kinase assays also demonstrated a lack of activity associated with the mutant kinase (Fig. 2B) when compared to its wild type counterpart.

To validate this data, we performed similar experiments in human cells. U2OS cells were transfected with plasmids directing expression of Flag-tagged cdk2 (wild type or P45L mutant) and immunoprecipitated through the epitope tag. Immunoblotting the precipitates again demonstrated that the P45L mutant cdk2 failed to stably associate with cyclin E whereas this cyclin was readily detectable in the precipitates with the wild type cdk2 construct (Fig. 2C). Similar results were obtained when the immunoprecipitation was performed through either cyclin E1 or A2 (Fig. S2). This difference was also observed in immunoprecipitation kinase assays (Fig. 2D).

This biochemical evidence suggested that the P45L cdk2 was impaired in stable interaction with the cyclin subunit. We hypothesized that the mutation of the proline in the PSTAIRE helix may increase the flexibility in that region of the protein allowing the cdk to sample a greater ensemble of structures not all of which were permissive for cyclin binding. We reasoned that such an effect may be observed in a change in the thermal stability of cdk2. To test this idea we isolated poly(His)-tagged cdk2 (wild type and P45L mutant) from recombinant bacteria and performed circular dichroism (CD) and thermal denaturation assays. CD revealed wild type and P45L cdk2 proteins had very similar spectra indicating no gross differences in protein secondary structure (Fig. 3A). We next assessed cdk2 thermal stability by CD at 230 nm between 20 °C and 80 °C. From three separate determinations, wild type cdk2 melted at a lower temperature than the P45L mutant cdk2 (T_m 53.7 °C vs 58.1 °C, mean of two

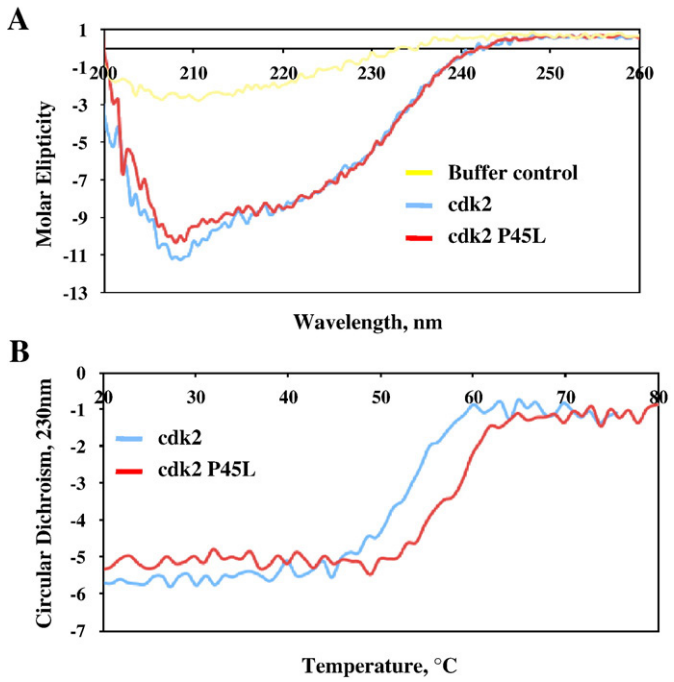


Fig. 3. P45L cdk2 is more stable than its wild type counterpart. A CD spectra of cdk2 (blue) and P45L cdk2 (red) at 20 °C from 200 to 260 nm. The yellow line delineates the buffer control. B Denaturation curves for cdk2 (blue) and P45L cdk2 (red) at 230 nm from 20 to 80 °C. For all CD data, representative plots are shown from three independent repeats.

Table 1
Summary of molecular dynamics simulations.

System	Force field	Duration (ns)	Mean RMSD (Å)	B_f PSTAIRE Helix (aa 45–59) ($Å^2$)	B_f T Loop (aa 157–163) ($Å^2$)
Wild type cdk2	*parm99*	10	1.9 ± 0.1	14.4	17.4
	ff03	30	1.7 ± 0.1	14.3	16.7
P45L-cdk2	*parm99*	30	1.9 ± 0.1	16.4	30.1
	ff03	30	1.7 ± 0.1	23.7	33.3
pT160-cdk2/ATP/ cyclin E1	*parm99*	10	3.7 ± 0.2	6.1	11.9
	ff03	30	3.2 ± 0.2	4.2	6.2
pT160-P45L-cdk2/ ATP/cyclin E1	*parm99*	10	1.8 ± 0.2	6.8	5.1
	ff03	30	3.2 ± 0.1	6.0	5.5

separate determinations; Fig. 3B). We verified this difference in a second thermal stability test, this time with a fluorescence-based assay using the environmentally sensitive dye, Sypro Orange. Again, the mutant protein was significantly more thermostable than the wild type (T_m 51.15 ± 0.13 °C vs 52.03 ± 0.35 °C, mean ± SD, $P = 0.0151$ by *t* test; Fig. S3).

To further elucidate the role of the P45L mutation on cdk2 and cdk2/cyclin E complex structure and dynamics, we performed molecular dynamics (MD) simulations of the following systems (see Table 1 for overview): monomeric wild type or P45L mutant cdk2 and both kinases (the threonine 160 phosphorylated form) in complex with cyclin E1 (and ATP). The convergence and the stability of the MD trajectories were documented by time evolution of root-mean-square deviation (RMSD) of heavy atoms from the X-ray structures (Fig. S4). The RMSD of all systems become balanced after ~3 ns then systems reached their equilibrated states and remained fluctuating around the average values (Table 1). These RMSDs show that MD simulations were convergent and ready for further structural analyses.

The typical bilobal kinase fold is retained in all structures (Fig. 4). The largest differences between wild type cdk2 and P45L cdk2 occurred in the PSTAIRE helix, which contacts the regulatory cyclins and β3-PSTAIRE helix connecting loop. (A number of exposed loops also showed substantial structural deviations but given the inherent flexibility of such regions [18] we did not consider them further.) Residue P45 at the N-terminus of PSTAIRE helix in wild type cdk2 is partially exposed to the solvent. In the crystal structure of the cdk2/cyclin E1, the PSTAIRE helix contacts cyclin E1 and adopts its rotated inner conformation, which allows hydrogen bond between E51 from PSTAIRE and K31 and thus the P45 residue is rotated into the cdk2 protein too. The mutant structures of monomeric cdk2 differed in both PSTAIRE helix position and β3-PSTAIRE loop conformation from wild type structures (Fig. 5). The PSTAIRE helix of the mutant structures is rotated from its native position by 16 degrees (Fig. 5). While β3-PSTAIRE loop conformations of the wild type align well, the loop conformations of the P45L mutant differed from the wild type significantly, one being localized below and one above the conformation of the wild type loop, indicating greater flexibility of this region in agreement with analysis of temperature *B*-factors.

Temperature *B*-factors (B_f) provide information about thermal fluctuations of amino acids. The B_f values were calculated from the last 1 ns of the MD simulations (Fig. S5). In all cases, B_f values of the PSTAIRE helix (residues 45–59) were higher for P45L mutants than for the wild types in agreement with the increased flexibility of the P45L cdk2 in this region. The B_f values of both wild type cdk2/cyclin E1 and P45L cdk2/cyclin E1 systems were similar to one another differing in flexibility within the activation segment (*parm99* simulation, maximum at residue 156) of cdk2 (Table 1). More variation was observed

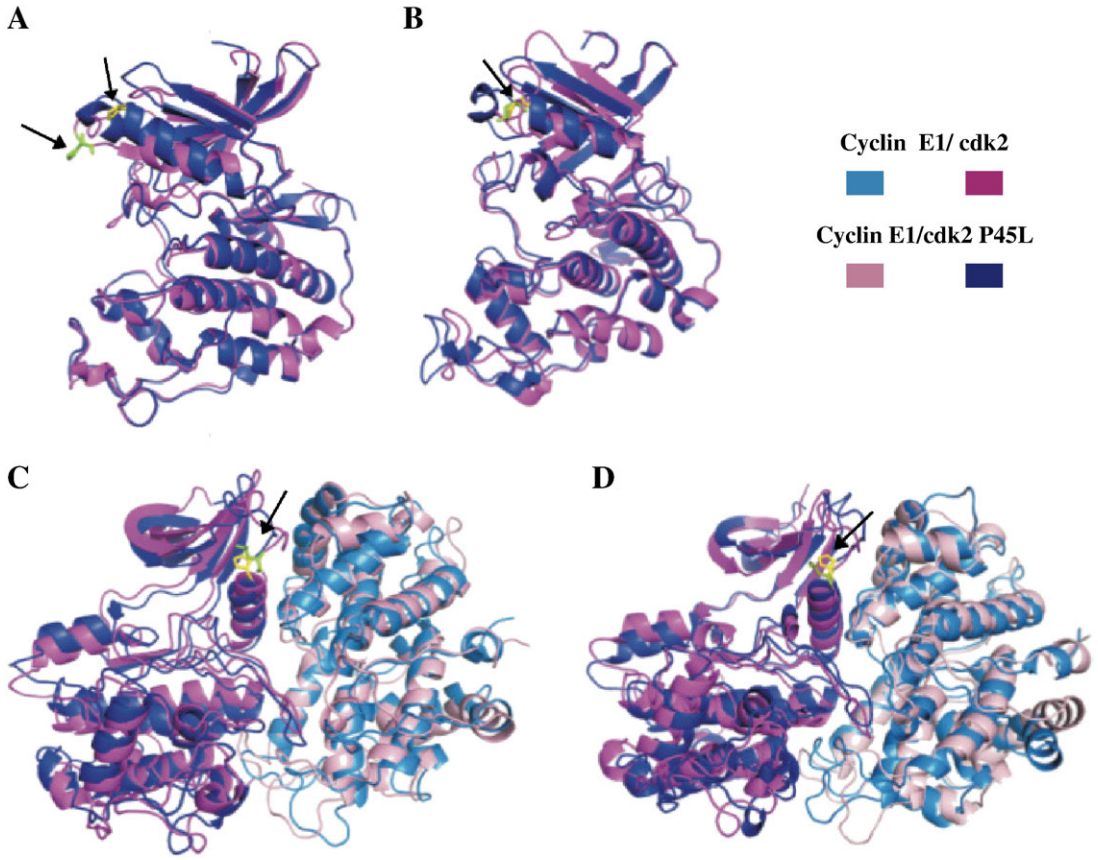


Fig. 4. Molecular dynamics simulations of cdk2 variants. Images showing the average structures over the last 1 ns of simulations with proteins coloured according to the key. Amino acid 45 of the cdk2 is indicated by arrows in each panel, with the proline coloured yellow and the leucine green. A *parm99* and B *ff03* depict monomeric wild type and P45L cdk2. C *parm99* and D *ff03* depict cyclin E1/cdk2 complexes.

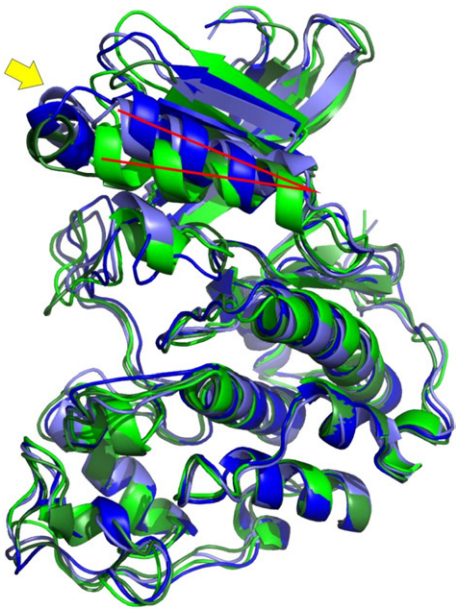


Fig. 5. Structural differences predicted between wild type and P45L cdk2. Superposition of monomeric cdk2 structures averaged over the last ns of MD simulations, cdk2 (*parm99*) in dark blue, cdk2 (*ff03*) in light blue, P45L cdk2 (*parm99*) in dark green and P45L cdk2 (*ff03*) light green. The yellow arrow indicates the β3-PSTAIRE loop and an angle between the axes of PSTAIRE and LSTAIRE helices (red lines) is 16°.

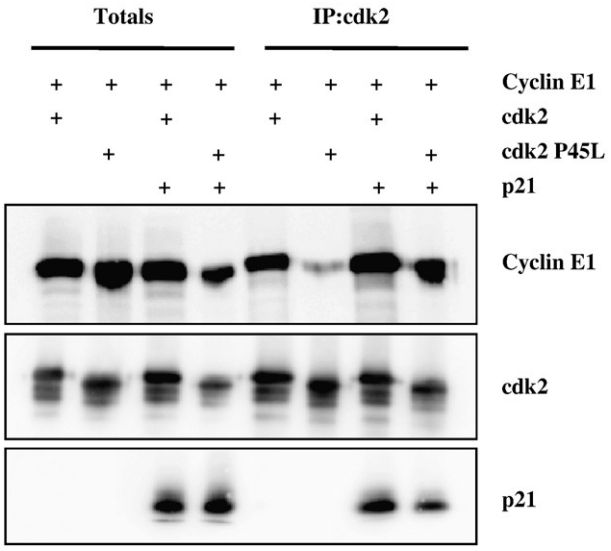


Fig. 6. CIP/KIP inhibitors can stabilise the P45L cdk/cyclin E1 complex. Sf9 cells were co-infected with recombinant baculovirus encoding cyclin E1, wild type cdk2 or P45L cdk2 in the absence and presence of p21-encoding virus. Cell lysates were immunoprecipitated through cdk2 and immunoprecipitations were immunoblotted as indicated.

in the flexibility of the cyclin E1 associated with wild type cdk2: residues 252–256 which are oriented toward the activation segment of cdk2 and residues 339–347, a solvent exposed flexible loop lying adjacent to residues 252–256 (Fig. S5). The B_f of P45 and L45 does not substantially differ in cyclin-bound cdk2 with both being lower than B_fs of the respective residues of monomeric cdk2.

The rotated position of PSTAIRE helix and increased flexibility around the PSTAIRE helix in P45L cdk2 suggested by the MD simulations could provide an explanation for the reduced binding of this cdk with cyclins observed biochemically. The cdk inhibitors p21 and p27 have been shown to help in the assembly and stabilisation of D-type cyclin/cdk complexes [22,23]. Given the lack of stable interaction between the P45L cdk2 and cyclins, we asked if p21 could stabilise this cdk/cyclin interaction. Again, Sf9 cells were infected with baculovirus directing expression of cdk2 (wild type or mutant) and cyclin E1 in the absence or presence of p21-encoding virus. Cells were lysed and immunoprecipitated through the cdk and immunoblotted for the cyclin and p21. Fig. 6 clearly demonstrates that p21 was able to substantially stabilise the mutant cdk/cyclin interaction. Similar results were obtained with cyclin A2 (not shown).

4. Discussion

Activating mutations in protein kinases are increasingly recognised as driver mutations in various forms of cancer [7]. We have analysed the P45L mutation in the cell cycle regulator cdk2 found through targeted genomic sequencing of a glioblastoma and affecting the PSTAIRE cyclin-binding helix characteristic of cdks. The mutation does not appear to cause gross alterations to the cdk2 structure as indicated by the retention of biological activity and the similarity of the wild type and mutant CD spectra. Our analysis reveals interesting features of cdk2 biochemistry but does not suggest that this mutation is likely to be causal for cancer.

The interaction of cyclins and cdks relies upon a large interface (3252 Å^2 for cyclin E1/cdk2 [6]) centred on the PSTAIRE helix of the cdk subunit. In the crystal structure of cyclin E1/cdk2 [6] and cyclin A2/cdk2 [4,15], the PSTAIRE helix lies almost parallel to the α5 helix of the cyclin and this cyclin helix is also adjacent to the loop connecting β3 and the PSTAIRE helix. These closely aligned regions make numerous hydrophobic interactions and hydrogen bonds in the cyclin E1/cdk2 and cyclin A2/cdk2 crystal structures [4,6]. The major difference observed biochemically between the wild type and P45L cdk2 proteins was the reduced ability of the mutant to maintain association with its cognate cyclin over prolonged periods. MD simulations indicate that the P45L mutation leads to a greater degree of flexibility in the β3-PSTAIRE helix loop as well as a displacement of the PSTAIRE helix in the monomeric kinase subunit and these analyses provide a potential explanation for the reduced stable association of this mutant kinase with its regulatory cyclin subunits observed experimentally.

Despite the lack of stable cyclin binding, P45L cdk2 displayed enzymatic activity very similar to the wild type kinase when assayed in lysates of recombinant baculovirus-infected Sf9 cells (the P45L mutant cdk2 retained ~62% of the activity of wild type cdk2 in an Sf9 cell lysate although this activity is almost entirely lost during immunoprecipitation). One explanation for this observation could be a weaker, transient association between cyclin and cdk favoured in lysates over-expressing cyclins and cdks allowing occasional activation. Alternatively, it may be possible that the kinase assay conditions (high levels of ATP, protein substrate and kinase components) may stabilise the active cyclin/cdk complex permitting substrate phosphorylation whereas the immunoprecipitation assay conditions (no added ATP, dilute components) are not favourable for complex stability. The in vivo rescue of yeast expressing a temperature sensitive variant of CDC28 also provides compelling evidence that P45L cdk2 is functional in a cellular setting, delivering sufficient catalytic activity to allow cell cycle progression.

The ability of the cdk inhibitor p21 to facilitate stable complex assembly between the P45L cdk2 and a cyclin is reminiscent of the dual role of p21 in the activation of D-type cyclin/cdks [22,23]. However, to our knowledge, there are no reports of cdk2 kinase activity being sustained in the face of cdk inhibitor binding binding making this is an unlikely mechanism to account for the kinase activity we observed in cell lysates or the rescue of proliferation in the cdc28^{ts} yeast. However, it is possible that other cellular chaperones account for the assembly activity required for these events; it is known, for example, that D-type cyclin/cdk complexes have assembly factors other than p21 and p27 [24,25] and such mechanisms may

account for P45L cdk2/cyclin complexation, with chaperones perhaps facilitating short-term interactions that are not maintained during the extended periods of immunoprecipitation experiments.

Isolation of the P45L mutant of cdk2 from a glioblastoma raised the possibility that the mutation could be a contributory factor to the aetiology of the disease. Our findings on the properties of the protein suggest that this is most likely not to be the case. Although the mutation produces a kinase that is functional in vitro and in vivo, the enzyme is impaired in its ability to form a stable association with its activating cyclin subunits. Thus, this mutation alone is unlikely to be a driver of oncogenesis, although in the context of the genomic instability associated with cancer, it is possible that other genetic alterations may influence its function.

Acknowledgements

MO would like to acknowledge grants from the GACR (203/09/H046) and Ministry of Education, Youth and Sports of the Czech Republic (MSM6198959216, LC512). PAF would like to acknowledge the Wellcome Trust for support under grant reference 077012/Z/05/Z. DJM would like to acknowledge grants from the BBSRC BB/D524208/1 and BB/C510859/1 and the Chemical Biology Centre at Imperial College for funding (to KM). We thank David Morgan, John Diffley and Gordon Peters for the generous provision of reagents and Kurt Drikamer for CD assistance.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bbamcr.2010.04.004.

References

[1] M. Malumbres, M. Barbacid, Cell cycle, CDKs and cancer: a changing paradigm, Nat. Rev. Cancer 9 (2009) 153–166.
[2] M. Malumbres, M. Barbacid, To cycle or not to cycle: a critical decision in cancer, Nat. Rev. Cancer 1 (2001) 222–231.
[3] D.O. Morgan, H.L. De Bondt, Protein kinase regulation: insights from crystal structure analysis, Curr. Opin. Cell Biol. 6 (1994) 239–246.
[4] P.D. Jeffrey, A.A. Russo, K. Polyak, E. Gibbs, J. Hurwitz, J. Massague, N.P. Pavletich, Mechanism of cdk activation revealed by the structure of a cyclin A–cdk2 complex, Nature 376 (1995) 313–320.
[5] C. Kim, N.H. Xuong, S.S. Taylor, Crystal structure of a complex between the catalytic and regulatory subunits of PKA, Science 307 (2005) 690–696.
[6] R. Honda, E.D. Lowe, E. Dubinina, V. Skamnaki, A. Cook, N.R. Brown, L.N. Johnson, The structure of cyclin E1/CDK2: implications for CDK2 activation and CDK2-independent roles, EMBO J. 24 (2005) 452–463.
[7] M.R. Stratton, P.J. Campbell, P.A. Futreal, The cancer genome, Nature 458 (2009) 719–724.
[8] P. James, J. Halladay, E.A. Craig, Genomic libraries and a host strain designed for highly efficient two-hybrid selection in yeast, Genetics 144 (1996) 1425–1436.
[9] D.J. Mann, E.S. Child, C. Swanton, H. Laman, N.C. Jones, Modulation of $p27^{Kip1}$ levels by the cyclin encoded by Kaposi's sarcoma-associated herpesvirus, EMBO J. 18 (1999) 654–663.
[10] T.W. Mayhood, W.T. Windsor, Ligand binding affinity determined by temperature-dependent circular dichroism: cyclin-dependent kinase 2 inhibitors, Anal. Biochem. 345 (2005) 187–197.
[11] D.A. Case, T.E. Cheatham, C.L. Simmerling, J. Wang, R.E. Duke, K.M.M. Luo, D.A. Pearlman, M. Crowley, R.C. Walker, W. Zhang, B. Wang, S. Hayik, G. Seabra, K.F. Wong, F. Paesani, X. Wu, S. Brozell, V. Tsui, H. Gohlke, C. Tan, J. Mongan, V. Hornak, G. Cui, P. Beroza, D.H. Mathews, C. Schafmeister, P.A. Kollman. AMBER 9. University of California: San Francisco, 2006.
[12] W.D. Cornell, P. Cieplak, C.I. Bayly, I.R. Gould, K.M. Merz, D.M. Ferguson, D.C. Spellmeyer, T. Fox, J.W. Caldwell, P.A. Kollman, A second generation force field for the simulation of proteins, nucleic acids, and organic molecules, J. Am. Chem. Soc. 117 (1995) 5179–5197.
[13] Y. Duan, C. Wu, S. Chowdhury, M.C. Lee, G. Xiong, W. Zhang, R. Yang, P. Cieplak, R. Luo, T. Lee, A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations, J. Comput. Chem. 24 (2003) 1999–2012.
[14] V. Krystof, P. Cankar, I. Frysová, J. Slouka, G. Kontopidis, P. Dzubák, M. Hajdúch, J. Srovnal, W.F. de Azevedo, M. Orság, M. Paprskárová, J. Rolcík, A. Látr, P.M. Fischer, M. Strnad, 4-arylazo-3, 5-diamino-1H-pyrazole CDK inhibitors: SAR study, crystal structure in complex with CDK2, selectivity and cellular effects, J. Med. Chem. 49 (2006) 6500–6509.
[15] A.A. Russo, P.D. Jeffrey, N.P. Pavletich, Structural basis of cyclin-dependent kinase activation by phosphorylation, Nat. Struct. Biol. 3 (1996) 696–700.
[16] I. Bártová, M. Otyepka, Z. Kříž, J. Koča, Activation and inhibition of cyclin-dependent kinase-2 by phosphorylation; a molecular dynamics study reveals the functional importance of the glycine-rich loop, Protein Sci. 13 (2004) 1449–1457.
[17] I. Bártová, M. Otyepka, Z. Kříž, J. Koča, The mechanism of inhibition of the cyclin-dependent kinase-2 as revealed by the molecular dynamics study on the complex CDK2 with the peptide substrate HHASPRK, Protein Sci. 14 (2005) 445–451.
[18] M. Otyepka, I. Bartova, Z. Kříž, J. Koča, Different mechanisms of CDK5 and CDK2 activation as revealed by CDK5/p25 and CDK2/cyclin A dynamics, J. Biol. Chem. 281 (2006) 7271–7281.
[19] I. Bartova, J. Koča, M. Otyepka, Regulatory phosphorylation of cyclin-dependent kinase 2: insights from molecular dynamics simulations, J. Mol. Model. 14 (2008) 761–768.
[20] G. Vriend, WHAT IF, 5.0, EMBL, Heidelberg, 1997.
[21] K. Levine, L. Kiang, M.D. Jacobson, R.P. Fisher, F.R. Cross, Directed evolution to bypass cyclin requirements for the Cdc28p cyclin-dependent kinase, Mol. Cell 4 (1999) 353–363.
[22] E.S. Child, D.J. Mann, The intricacies of p21 phosphorylation: protein/protein interactions, subcellular localisation and stability, Cell Cycle 5 (2006) 1313–1319.
[23] S.W. Blain, Switching cyclin D-Cdk4 kinase activity on and off, Cell Cycle 7 (2008) 892–898.
[24] M. Sugimoto, N. Martin, D.P. Wilks, K. Tamai, T.J. Huot, C. Pantoja, K. Okumura, M. Serrano, E. Hara, Activation of cyclin D1-kinase in murine fibroblasts lacking both p21(Cip1) and p27(Kip1), Oncogene 21 (2002) 8067–8074.
[25] T.K. Bagui, S. Mohapatra, E. Haura, W.J. Pledger, p27Kip1 and p21Cip1 are not required for the formation of active D cyclin-cdk4 complexes, Mol. Cell. Biol. 23 (2003) 7285–7290.