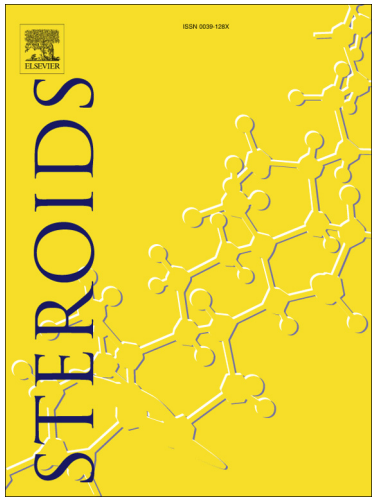


Ile351, Leu355 and Ile461 residues are essential for catalytic activity of bovine cytochrome P450scc (CYP11A1)

Anna V. Glyakina, Nicolai I. Strizhov, Mikhail V. Karpov, Nikita V. Dovidchenko, Bakhyt T. Matkarimov, Ludmila V. Isaeva, Vera S. Efimova, Mikhail A. Rubtsov, Ludmila A. Novikova, Marina V. Donova, Oxana V. Galzitskaya

PII: S0039-128X(19)30004-2
DOI: https://doi.org/10.1016/j.steroids.2019.01.002
Reference: STE 8362

To appear in: *Steroids*

Received Date: 19 October 2018
Revised Date: 28 December 2018
Accepted Date: 7 January 2019

Please cite this article as: Glyakina, A.V., Strizhov, N.I., Karpov, M.V., Dovidchenko, N.V., Matkarimov, B.T., Isaeva, L.V., Efimova, V.S., Rubtsov, M.A., Novikova, L.A., Donova, M.V., Galzitskaya, O.V., Ile351, Leu355 and Ile461 residues are essential for catalytic activity of bovine cytochrome P450scc (CYP11A1), *Steroids* (2019), doi: https://doi.org/10.1016/j.steroids.2019.01.002

Ile351, Leu355 and Ile461 residues are essential for catalytic activity of bovine cytochrome P450scc (CYP11A1)

Anna V. Glyakina[a,b], Nicolai I. Strizhov[d], Mikhail V. Karpov[c,d], Nikita V. Dovidchenko[a], Bakhyt T. Matkarimov[e], Ludmila V. Isaeva[f], Vera S. Efimova[g], Mikhail A. Rubtsov[g,h], Ludmila A. Novikova[f], Marina V. Donova[d]*, Oxana V. Galzitskaya[a]*

[a]Institute of Protein Research, Russian Academy of Sciences, Pushchino, Moscow Region 142290, Russia
[b]Institute of Mathematical Problems of Biology, Russian Academy of Sciences, Keldysh Institute of Applied Mathematics, Russian Academy of Sciences, 142290 Pushchino, Moscow Region, Russia
[c]Institute of Biochemistry and Physiology of Microorganisms, Federal Research Center "Pushchino Center for Biological Research of the Russian Academy of Sciences", 142290 Pushchino, Moscow Region, Russia;
[d]Pharmins, Ltd., R&D, 142290 Pushchino, Moscow Region, Russia
[e]National Laboratory Astana, Nazarbayev University, Astana, Kazakhstan
[f]Belozersky Institute of Physico-Chemical Biology, M.V. Lomonosov Moscow State University, Leninskie Gory, 1/40, 119234, Moscow, Russia
[g]Department of Molecular Biology, Faculty of Biology, M.V. Lomonosov Moscow State University, Leninskie Gory, 1/12, 119234, Moscow, Russia
[h]Department of Biochemistry, Institute of Translational Medicine and Biotechnology, I.M. Sechenov First Moscow State Medical University (Sechenov University), 119991, Moscow, Russia

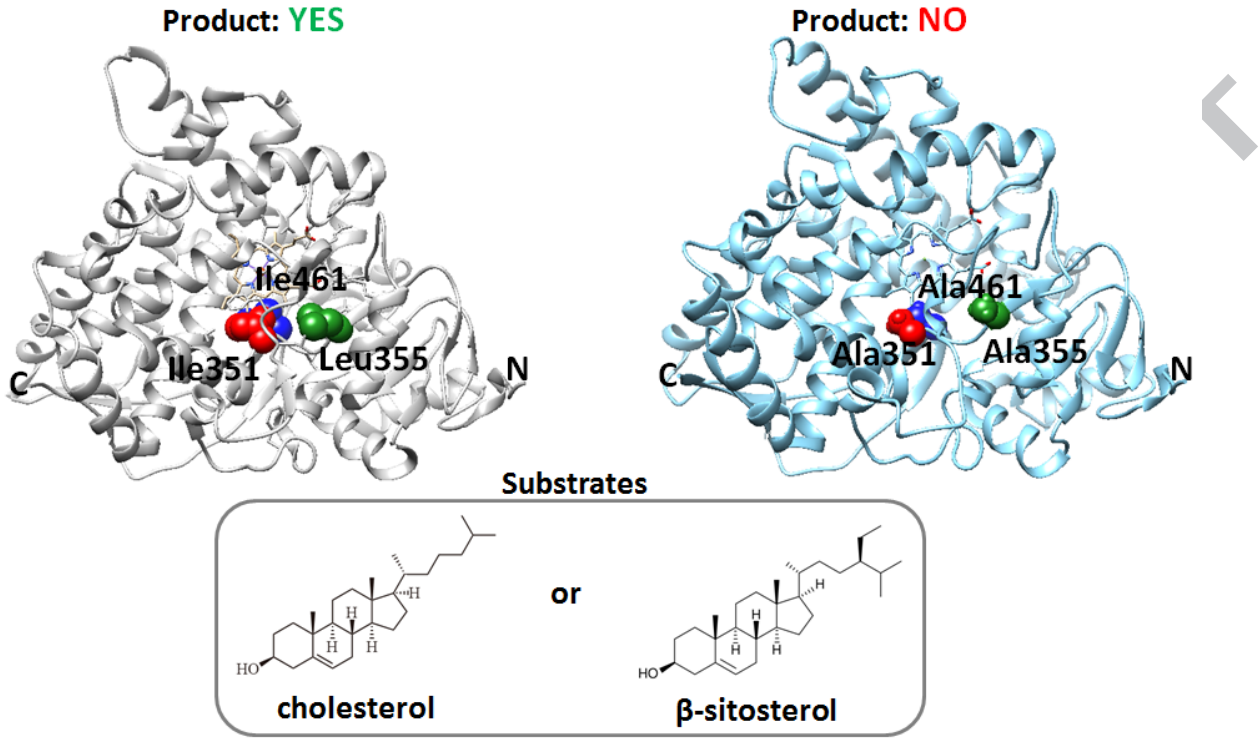
*Correspondence: Dr. Marina Donova, Institute of Biochemistry and Physiology of Microorganisms, Federal Research Center "Pushchino Center for Biological Research of the Russian Academy of Sciences", 142290 Pushchino, Moscow Region, Russia donova@ibpm.pushchino.ru
Dr. Oxana Galzitskaya, Institute of Protein Research, Russian Academy of Sciences, Pushchino, Moscow Region, 142290 Russia *E-mail*: ogalzit@vega.protres.ru

Abstract

Cytochrome P450scc (CYP11A1) is a mammalian mitochondrial enzyme which catalyzes cholesterol side chain cleavage to form pregnenolone. Along with cholesterol, some other steroids including sterols with a branched side chain like β-sitosterol are the substrates for the enzyme, but the activity towards β-sitosterol is rather low. Modification of the catalytic site conformation could provide more effective β-sitosterol bioconversion by the enzyme. This study was aimed to find out the amino acid residues substitution of which could modify the conformation of the active site providing possible higher enzyme activity towards β-sitosterol. After structural and bioinformatics analysis three amino acid residues I351, L355, I461 were chosen. Molecular dynamics simulations of P450scc evidenced the stability of the wild type, double (I351A/L355A) and triple (I351A/L355A/I461A) mutants. Mutant variants of cDNA encoding P450scc with the single, double and triple mutations were obtained by site-directed mutagenesis. However, the experimental data indicate that the introduced single mutations Ile351A, Leu355A and Ile461A dramatically decrease the target catalytic activity of CYP11A1, and no activity was observed for double and triple mutants obtained. Therefore, isoleucine residues 351 and 461, and leucine residue 355 are important for the cytochrome P450scc functioning towards sterols both with unbranched (cholesterol) and branched (sitosterol) side chains.

Key words: P450scc; catalytic activity; cholesterol; β-sitosterol; mutations; MD simulations

Graphical abstract



Highlights

The considered cytochrome P450 enzymes have low sequence identities, but quite similar three-dimensional structures.

The cytochrome P450 proteins are depleted in Ala and Gly and enriched with Met, Phe, Ile versus the proteome values.

Residues Ile351, Leu355 and Ile461 are essential for the functioning of cytochrome P450scc towards cholesterol and β-sitosterol.

Abbreviations:

P450scc, CYP11A1 – cytochrome P450scc

Adx – Adrenodoxin

AdR – Adrenodoxin reductase

MD simulations – Molecular dynamics simulations

RMSD – Root mean square deviation

3D structure – three-dimensional structure

SDS-PAGE – sodium dodecyl sulfate polyacrylamide gel electrophoresis

Introduction

The cytochrome P450 enzymes (P450s or CYPs) are known to form a large superfamily of the enzymes, which are the external monooxygenases activating molecular oxygen via their heme iron and catalyze variety of oxidations of the numerous exogenous and endogenous molecules [1]. CYPs are quite diverse in their functions, types of enzymatic activity and substrate specificity [2]. The oxidation reactions with involvement of CYPs include transformations of drugs, xenobiotics, synthesis and transformation of steroids, and other important reactions [3]. These capacities make CYPs essential for the endogenous metabolism in all living systems.

57 CYPs genes and 58 pseudogenes were identified in humans. Most part of 18 families of cytochrome P450 genes and 43 subfamilies are related to steroidogenesis, or steroid hydroxylations (e.g. CYP2, CYP7A, CYP7B, CYP8B, CYP11, CYP17, CYP19, CYP24, CYP27A, CYP27B, CYP39, CYP46, CYP51)[2]. Mutations in the CYPs genes or deficiencies of the enzymes are responsible for several human diseases [4–7]. Induction of some CYPs is considered as a risk factor for several types of cancer since these enzymes can convert procarcinogens to carcinogens [8].

Despite the low homology of the primary structures between CYPs of different families (often less than 20%), their secondary and tertiary structures are highly conservative thus determining the general mechanism of the reaction catalysis [9]. A high degree of sequence and structural homology of the heme-binding site was reported for CYPs [10]. The first three-dimensional (3D) structures were represented for soluble bacterial CYPs by Poulos et al. [11]. Unlike bacterial CYPs, mammalian ones are membrane associated proteins. Current data on the 3D structures of P450 enzymes evidence that all CYPs have similar structural elements: 12 helices (denoted as A–L) and loops [12,13]. In mammalian CYPs in addition to the common CYP fold helices, other helices were described mainly in the F/G loop region such as additional helices in CYP11A1 termed A′, G′ and K″ [14]. These helices seem to provide the enzyme binding to membrane, forming substrate access channel, and play a role in substrate recognition. All CYPs contain heme (an iron-containing protoporphyrin IX) located between the helices I and L.

Diversity of the CYPs substrate specificities is explained by the differences in their active-site structures. The most important segments/amino acid residues of the P450s proteins contacting the substrate and providing the substrate specificity and regioselective hydroxylation both in bacterial and mammalian CYPs are the loop between the B and C helices positioned over the heme [15–17], I helix, located in the active site of all CYPs [16–18] and the loop between the F and G helices [12].

Cytochrome P450scc (CYP11A1) is a mammalian mitochondrial enzyme catalyzing oxidation cholesterol to pregnenolone (see Fig. 1). This reaction takes place mainly in the male and female reproductive tissues and the adrenal gland and represents the first, rate-limiting and hormonally

regulated step in the synthesis of all steroid hormones in mammals. The CYP11A1 gene is found only in vertebrates [19].

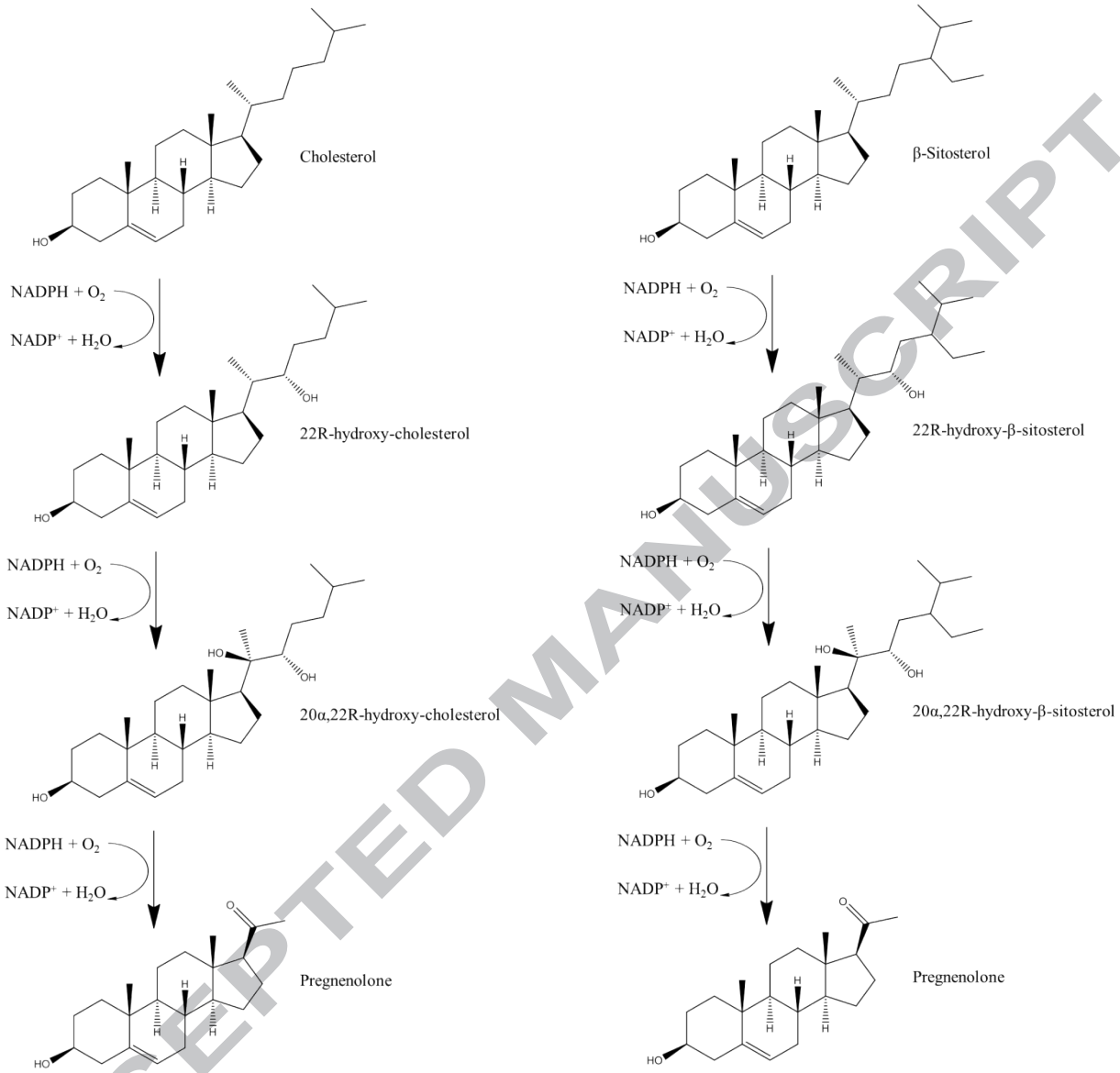


Figure 1. Chemical reactions catalysed by cytochrome P450scc (CYP11A1) for both substrates (cholesterol, β-sitosterol) with conversion to product – pregnenolone.

The natural redox partners of CYP11A1 are the [2Fe2S] ferredoxin, adrenodoxin (Adx) and NADPH-dependent ferredoxin reductase, adrenodoxin reductase (AdR). Adx transfers electrons from AdR to the heme iron in CYP11A1. The terminal electron acceptor CYP11A1 takes up water-insoluble cholesterol from the inner mitochondrial membrane and converts it to the less hydrophobic product pregnenolone, removing the unpolar side chain via the three sequential reactions: production of 22R-hydroxycholesterol (22HC), 20*α*,22R-dihydroxycholesterol and further C20-C22 cleavage, which occurs at a single active site on the CYP11A1 molecule [20].

To date, several CYP11A1 models have been designed by computer modeling based on the experimental 3D structures of cytochromes P450cam, P450-BM3, P450terp, P4502C5 and P4502B4 [21–24]. Storbek et al. [25] presented a model of the active site of P450scc which was created based on the crystalline structures of CYP102 и CYP2C5. In 2011, the crystalline 3D structure of the bovine [20] and human CYP11A1 [14] was revealed by X-ray crystallography. Besides, the structures of CYP11A1 bound with cholesterol, the reaction intermediates and Adx were resolved [14,20]. These data allowed correct determination of the corresponding protein segments involved in the enzyme binding to the mitochondrial membrane and interaction with the redox partner. Sequence alignment of different mammalian P450scc and site-directed mutagenesis in combination with biochemical methods and analysis of the created 3D-structure models of the protein allowed to obtain essential information about the organization of the functionally important protein sites, namely, the site responsible for protein interaction with the inner mitochondrial membrane, the binding site with Adx, the site of interaction with the substrate, and the catalytic center. These data showed that P450scc has three additional helices to the common CYP fold termed A′, G′ and K″. Association of protein CYP11A1 with the membrane is mediated by the F-G loop and the A′ helix [23,26]. As identified by computational analysis [20] and complying with the data obtained earlier using site-directed mutagenesis and mass spectrometry studies [23,26,27], membrane-interacting residues in P450scc are Asn19, Trp21, Leu22, Tyr25, Leu216, Leu219, and Phe220.

Cytochrome P450scc interacts electrostatically with its redox partner Adx by means of positively charged amino acid residues of P450scc with negatively charged ones of Adx [22]. P450scc amino acid residues important for this interaction were identified: according to the molecular modeling and the biochemical studies [22], Lys377 and Lys381 (Lys339 and Lys343 for 3mzs) [28] and Lys267, Lys403, Lys405 and Arg426 (these residues form four key salt bridges with Glu47, Asp76, Asp72 and Glu73 of Adx) (Lys268, Lys404, Lys406 and Arg427 for 3mzs) play a role in the P450scc interaction with adrenodoxin [22].

The size of the interacting surfaces and the number of amino acid residues involved in the formation of the P450 complex with Adx was determined using multiple sequence alignment of the mitochondrial cytochrome P450 proteins from subphylum *Vertebrata* [22]. Amino acids conserved for all mitochondrial P450s were observed in the interface thus suggesting the presence of some common mechanism for all mitochondrial CYPs. These conserved tertiary structure interactions probably underlie the specificity of mitochondrial P450s for their shared redox partner Adx.

The enzymatic center of P450 is located inside the protein globule. After binding of the substrate to the protein surface, it moves to the active site along the hydrophobic channel [29]. The region of the primary binding of the substrate is located in the N-terminal part of the polypeptide chain (amino acid residues 8-28) [29,30]. It is suggested that the interaction of Ile98 (Ile99 for 3mzs),

located in the B'-helix, with the I-helix opens a hydrophobic channel upon substrate binding to the P450scc molecule [18].

Positioning of cholesterol in the active site of P450scc is due to its interactions with many hydrophobic residues of the enzyme. The cholesterol 3β-OH group forms the H-bond with water but does not directly interact with the protein. According to Mast et al. [20], the active site cavity in P450scc consists of 39 amino acid residues from 14 different secondary structural elements. Only 12 of these residues are within 4 Å of the substrate. They are Ile351 (β1–4 strand region), Thr291 (I-helix), Phe203 (F-helix), Met202 (F-helix), Ile85 (B-B' loop), Leu460 (β4-1/4-2 loop), Trp88 (B'-helix), Gln356 (β1–4 strand region), Thr354 (β1–4 strand region), Val353 (β1–4 strand region), Ser352 (β1–4 strand region) and Phe458 (β1–4 strand). Strushkevich et al. [14] suggested that the amino acid residues Leu101 (102), Trp87 (88), Phe202 (203), Ile461, Ser352, His39 (40), Tyr61 (62), Asn210 (211), Gln377 are responsible for the preparation of the correct conformation of the substrate for further catalytic reaction, and the amino acid residues Gly287 and Thr291 are directly involved in the transfer of the electron to the catalytic center. Thr291 was assumed to play an essential role in the enzyme interaction with cholesterol, while Arg357 and Ser352 provide its interactions with the intermediates thus providing hydroxylation and further C-C-cleaving [14]. Structural studies of P450scc bound with substrates showed that during cascade reactions of cholesterol transformation in the enzyme active center the position of the steroid core rings remains fixed during changes in the side chain.

Up to now, little is known about the activity of CYP11A1 towards other steroid substrates. It was shown that the enzyme transforms 7-dehydrocholesterol (7DHC), the precursor of vitamin D3, to 7-dehydropregnenolone (7DHP), whose photo-transformed 5,7-diene derivatives exhibit an anti-proliferative effect against melanoma and leukemia cells [19]. The data on the activity of CYP11A1 towards plant sterols, phytosterols, which structurally differ from cholesterol by a branched side chain, are scarce [19]. It was shown that the enzyme activity with sitosterol is much lower as compared with cholesterol.

In this study, we have analyzed cholesterol metabolizing CYPs, identified the amino acid residues essential for the enzyme potent activity towards branched sterols and experimentally examined their significance for the enzyme activity.

2. Materials and Methods

2.1 Structural and Bioinformatics analysis and modeling

2.1.1 Search for structural homologues of P450scc

Twenty-seven structural homologues for P450scc with ability to interact with cholesterol were selected from the Uniprot database. Of these, 14 structures have resolved 3D structures. Root mean square deviation (RMSD) and sequence identity of the selected structures were calculated using the program UCSF Chimera [31]. Flexible/disordered regions in the proteins were identified using the FoldUnfold [32] and IsUnstruct [33] programs. Affinities for binding to different ligands (cholesterol and β-sitosterol) of the homologous structures were calculated using the program AutoDock Vina [34].

The amino acid composition of the selected structures was compared with that of the corresponding proteomes (human, bovine, bacterial). For the bacterial structures the comparison was carried out with the average bacterial proteomes (averaging was carried out for 18 species of *Mycobacterium tuberculosis*: ATCC 25177, F11, KZN 1435, CDC1551A, SUMu001, SUMu002, SUMu003, SUMu004, SUMu005, SUMu006, SUMu007, SUMu008, SUMu009, SUMu010, SUMu011, SUMu012, ATCC 25618, Oshkosh).

2.1.2 Molecular dynamics simulations

The objects of the simulation study were wild type proteins P450scc and CYP7A (pdb ids 3mzs and 3v8d, respectively) and double (I351A/L355A) and triple (I351A/L355A/I461A) mutants of 3mzs with two sterol substrates, - cholesterol and β-sitosterol. Replacement of the residues and insertion of the cholesterol and β-sitosterol in the structure was made using the YASARA program [35].

Molecular dynamics simulations were performed using the program PUMA [36,37]. The system of classical equations of motion for the atoms was resolved in the all-atom force field AMBER-99 [38]. To maintain a constant temperature, a collisional thermostat [39,40] was used. The mass of each virtual particle was one atomic mass unit, and the mean collision frequency of the atoms with virtual particles was 10 ps^{-1}. The simulations were performed at 300 K and 350 K. The equations of motion were integrated numerically using the velocity version of the Verlet algorithm [41] with a time step of 1 fs (10^{-15} s). For water molecules, the TIP3P [42] model was used. The fraction of secondary structure during the modeling was calculated using the YASARA program [35].

2.2 Experimental procedures

2.2.1 Chemicals, enzymes, oligonucleotides and antibodies

Isopropyl-β-D-thiogalactopyranoside (IPTG), δ-aminolevulinic acid (ALA), diaminobenzidine tetrahydrochloride hydrate (DAB), 20α-hydroxycholesterol, β-sitosterol, cholesterol oxidase and horse-radish peroxidase-conjugated anti-rabbit antibodies were supplied by Sigma (USA) and nitrocellulose filters Hybond-C extra by Amersham (USA). Growth media (LB, TB [43]) were prepared using reagents from Difco (USA).

DNA-modifying enzymes, *Pfu* DNA polymerase, RNase A and DNA Extraction Kit were purchased from MBI Fermentas (Lithuania). Oligonucleotides used in the study were synthesized by Evrogen (Russia). Primary antibodies (IgG-fraction) against cytochrome P450scc were generously provided by Prof. V. M. Shkumatov (Institute of Physical Chemical Problems, Minsk State University, Belarus).

2.2.2 Bacterial strain and plasmids

E. coli strain DH5αc (supE44 Δlac U169 (ϕ 80 lacZ ΔM15) hsdR17 recA1end A1 gyrA96 thi-1 relA1) was supplied by Gibco-BRL, Germany.

Vector pTrc99A/P450scc [28] containing a hybrid trp/lac (trc) promoter and cDNA encoding the mature bovine cytochrome P450scc was provided by M. R. Waterman (University of Texas, Dallas, USA). Plasmid pBar_Triple [44] with the cDNAs of mature bovine cytochrome P450scc, AdR [45] and Adx (4-108) [46] was constructed earlier in our laboratory.

2.2.3 Plasmids construction

The molecular cloning was performed by standard protocols [43]. The pTrc99A/CHL plasmid containing the tricistronic cassette with the cDNAs encoding the mature forms of cytochrome P450scc, AdR and Adx1-108 from bovine adrenal cortex was constructed on the base of the vector pTrc99A/P450scc [28]. The resulting pTrc99A/CHL plasmid has the cDNAs in the following sequential order: *trc* promoter, P450scc, AdR and Adx. The construction details of the plasmid will be described elsewhere soon. The plasmids pBar_Twin[47] and pBar_Triple [44] served as the source of the cDNAs encoding AdR and Adx1-108. The sequence of the constructed tricistronic cassette in the pTrc99A/CHL plasmid was confirmed by the DNA sequencing.

Site-directed mutagenesis of P450scc was performed by standard methods with the use of PCR primers carrying the GCC codon encoding the Ala amino acid residue for substitution of Ile351, Leu355 and Ile461 in the protein sequence of P450scc. All five resulting mutant DNA sequences (three single I351A, L355A, I461A, one double I351A/L355A and one triple I351A/L355A/I461A mutants) were validated by Sanger automatic DNA sequencing. These plasmids including mutated sequences, generated on the base of the parental plasmid pTrc99A/CHL were designated as pTrc99A/CHL(P450-I351A), pTrc99A/CHL(P450-L355A), pTrc99A/CHL(P450-I461A), pTrc99A/CHL(P450-I351A/L355A) and pTrc99A/CHL(P450-I351A/L355A/I461A).

2.2.4 Protein expression, preparation of bacterial homogenates

Expression of the recombinant proteins in *E. coli* cells was carried out as described earlier[48]. The cell culture was incubated at 28 °C with shaking (160 rpm) in the presence of ALA (0.5 mM) and ampicillin (100 μg/ml) for 22 h. Expression of the proteins encoded by the tricistronic cassette was induced by adding IPTG (0.5 mM) to the growth medium. *E. coli* cell homogenates were prepared as described previously[49]. Debris was removed by centrifugation (4,000 g, 10 min). The supernatant (homogenate) was used for P450scc activity measurements of the mutant/wild type proteins.

2.2.5 SDS-PAGE and Western immunoblotting

Cell lysates were analyzed by SDS-PAGE[50] followed by Western blotting [51]. Bacterial cells were pelleted by centrifugation. Pellets were re-suspended in Sample Buffer [49] and disrupted by boiling (2 min, 100 °C). Samples with 5 μg of total protein in 10 μl were loaded into each lane. The protein concentration was measured according to Lowry et al. [52] The membranes with transferred proteins separated by SDS-PAGE were incubated with the primary rabbit polyclonal antibodies against P450scc (6.0 mg/ml) at 1: 7,500 dilution (v/v) followed by treatment with the conjugate of anti-rabbit secondary antibodies with horseradish peroxidase at 1:15,000 dilution (v/v) and DAB peroxidase substrate (0.25 mg/ml). As shown previously, the primary antibodies against bovine P450scc used in this study demonstrate affinity both to natural mammalian proteins and to the recombinant P450scc proteins expressed in *E. coli* [53].

2.2.6 Cholesterol hydroxylase/lyase activity measurements

Catalytic activity of the recombinant P450scc proteins was determined *in vitro* as described previously [54]. The reaction mixture contained 1 mg of cell homogenate and 20α-hydroxycholesterol or β-sitosterol (25 nmol) in 30 mM sodium phosphate buffer (pH 7.2) containing 0.05% Tween-20 in a total volume of 0.5 ml. The reaction was induced by adding the NADPH-regenerating system and left to proceed for 30 min at 37 °C. The formed pregnenolone was oxidized to progesterone by adding cholesterol oxidase (0.5 U). The steroids were extracted with ethyl acetate; the precipitate obtained after the evaporation of the extracts on a vacuum rotor (SpeedVac Concentrator, Savant, USA) was dissolved in a sample buffer provided in the reagent kit “IFA PROGESTERONE” (Xema, Russia). The content of progesterone was determined by means of this immunoassay kit based on antiprogesterone antibodies according to the manufacturer's protocol.

3. Results and Discussion

3.1 Bioinformatics analysis of selected structures of P450scc

Fourteen structures of proteins with resolved 3D structures were selected as described above (point 2.1.1) for the further analysis. Among them, there are nine structures from human, four from bacteria, and one from yeast (Table 1). The comparison of the structures with bovine P450scc (CYP11A1) (pdb id 3mzs) showed low (about 25%) identity, however, the 3D structures of these proteins were quite similar (RMSD of about 2 Å) (Table 1).

Table 1. Identity and RMSD of the selected proteins with bovine P450scc (pdb id 3mzs), binding energy with cholesterol, β-sitosterol, living organism

Pdb id	Native ligand in pdb	Identity with 3mzs, %	RMSD with 3mzs, Å	Binding energy, kcal/mol		Organism
				Cholesterol	β-sitosterol	
2q9f	Cholest-5-en-3-yl hydrogen sulfate *Cholesterol-sulfate* $C_{27}H_{46}O_4S$	22	2.1	-11.7	-11.6	*Homo sapiens*
4wnu	Quinidine *(9S)-6'-methoxycinchonan-9-ol* $C_{20}H_{24}N_2O_2$	24	2.1	-8.4	-7.2	*Homo sapiens*
2hi4	2-phenyl-4h-benzo[h]chromen-4-one *7,8-benzoflavone; alpha-naphthoflavone* $C_{19}H_{12}O_2$	22	2.3	-6.9	-7.9	*Homo sapiens*
1r9o	Flurbiprofen $C_{15}H_{13}FO_2$	14	2.3	-9.7	-7.6	*Homo sapiens*
3n9y	Cholesterol $C_{27}H_{46}O$	73	0.9	-11.7	-11.5	*Homo sapiens*
4uhi	N-[(1R)-1-(3,4'-difluorobiphenyl-4-yl)-2-(1H-imidazol-1-yl)ethyl]-4-(5-phenyl-1,3,4-oxadiazol-2-yl)benzamide $C_{32}H_{23}F_2N_5O_2$	23	2.3	-11.3	-11.1	*Homo sapiens*
4dvq	Desoxycorticosterone *4-pregnen-21-ol-3,20-dione; doc; 21-hydroxyprogesterone* $C_{21}H_{30}O_3$	37	1.6	-7.3	-11.4	*Homo sapiens*
4d6z	Tert-butyl {6-oxo-6-[(pyridin-3-ylmethyl)amino]hexyl}carbamate $C_{17}H_{27}N_3O_3$	24	1.9	-10.5	-10.3	*Homo sapiens*
3v8d	(3beta,8alpha,9beta)-3-hydroxycholest-5-en-7-one *7-ketocholesterol* $C_{27}H_{44}O_2$	22	2.4	-13.0	-12.1	*Homo sapiens*
2wm	-	18	2.3	-9.1	-9.0	*Mycobacteriu*

5						*m tuberculosis*
2xkr	Tetraethylene glycol $C_8H_{18}O_5$	22	2.3	-7.7	-7.3	*Mycobacterium tuberculosis*
2x5l	-	17	2.4	-9.6	-9.3	*Mycobacterium tuberculosis*
4rm4	-	23	2.2	-	-	*Bacillus subtilis*
4lxj	Lanosterol $C_{30}H_{50}O$	24	2.2	-9.3	-10.9	*Saccharomyces cerevisiae*

The spatial superposition of all selected structures with each other is presented in Fig.2. Dark red regions correspond to the areas with a smaller RMSD relative to the P450scc structure. The results show that the topology is preserved in the heme region and in the areas adjacent to the catalytic center.

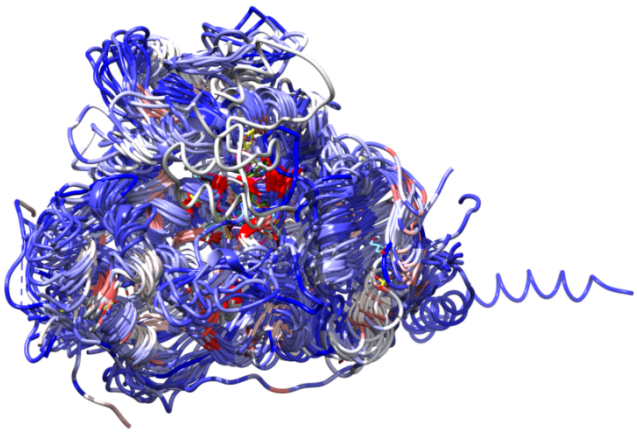


Figure 2. Spatial superposition of all 15 selected structures with each other. Dark red regions correspond to the areas with a smaller RMSD relative to the structure of bovine P450scc (pdb id 3mzs).

The spatial superposition of the structure of bovine CYP11A1 (3mzs) on the structure of human CYP7A1 (3v8d) is presented on Fig. 3. For each amino acid residue, the probability to be ordered or disordered was calculated for all structures of selected proteins using the IsUnstruct program. The results for two proteins bovine CYP11A1 (3mzs) and human CYP7A1 (3v8d) are represented in Fig. 3 (B, C). According to this program human CYP7A1 (3v8d) is the most unstructured in comparison with the selected proteins and has the lowest calculated binding energies with cholesterol and β-sitosterol (see Table 1). Analysis of the location of flexible regions for the structure of proteins showed that in general they are located around the exit site of the reaction product. Therefore, it can be assumed that these regions are important for the release of the reaction product.

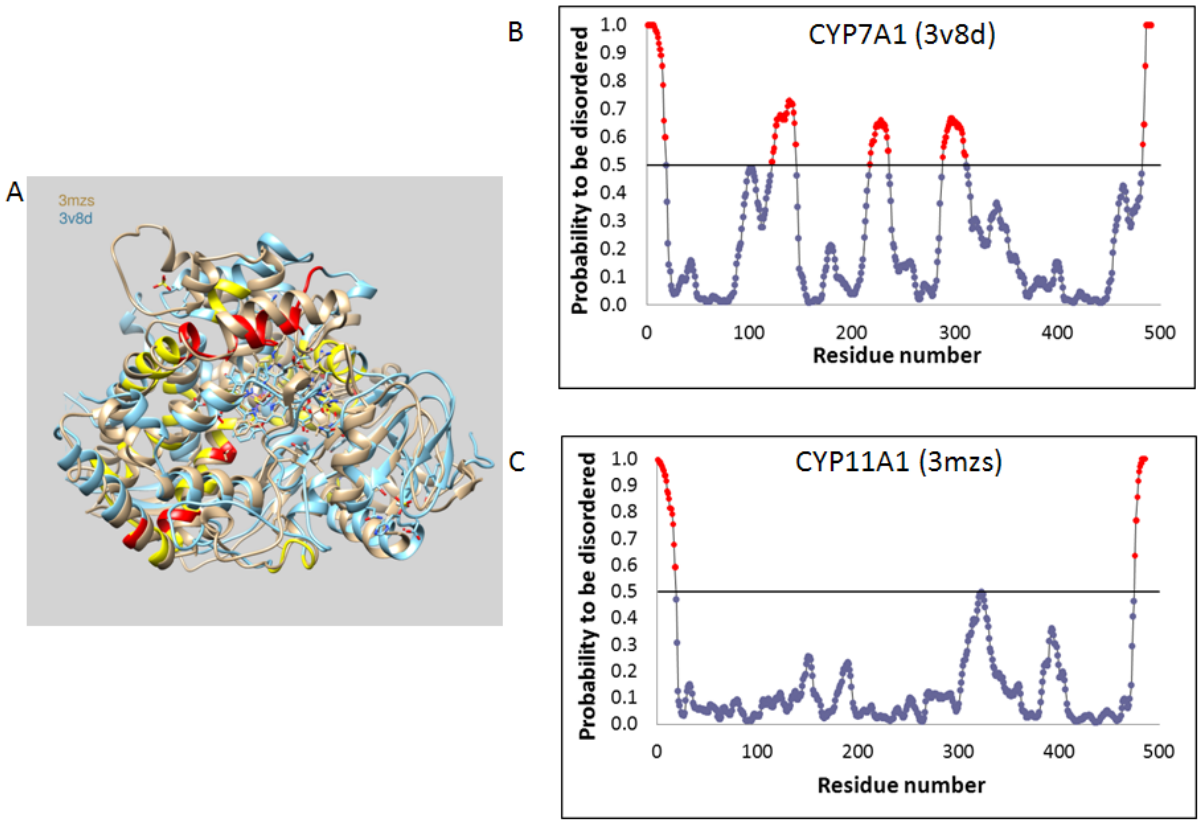


Figure 3. (A) Spatial superposition of the structure of bovine CYP11A1 (pdb id 3mzs) on the structure of human CYP7A1 (pdb id 3v8d). Unfolded (unstructured) regions, found by the FoldUnfold [32] (http://bioinfo.protres.ru/ogu) and IsUnstruct [33] (http://bioinfo.protres.ru/IsUnstruct) programs are indicated by yellow and red for CYP7A1 and CYP11A1 respectively. (B, C) Probability of each amino acid residue to be ordered or disordered in structures CYP7A1 and CYP11A1 calculated using the IsUnstruct program. Disordered regions are colored red.

The docking with cholesterol and β-sitosterol was performed for each of the selected structures (Table 1). In one case (pdb id 4rm4), the affinity value could not be determined, because there was no significant protein fragment in the structure adjacent to the catalytic pocket. In most cases no considerable difference in affinity values to the catalytic center between cholesterol and β-sitosterol was observed. In addition, for most structures, the calculated affinity values are close to those for bovine P450scc with cholesterol or β-sitosterol (about 12 kcal/mol) (Table 1). When the affinity values are small, it should be taken into account that the docking was carried out using the original structure obtained from the Protein Data Bank, and low affinity values do not exclude the possibility that for a certain protein conformation (for example, after relaxation of the protein with ligand) the values may vary.

The amino acid composition of the selected structures was analyzed and compared with the corresponding proteome values. This analysis allows us to estimate the difference between the amino acid compositions of an individual protein and a corresponding proteome. As seen from Fig. 4, the selected bacterial proteins are enriched with Met, Phe, Ile, Gln, Asn, Glu, Asp, His, Lys and depleted in Val, Ala, Gly, Thr, Arg versus the average *M. tuberculosis* proteomes (averaging was carried out for 18 species of *Mycobacterium tuberculosis,* see Methods section 2.1.1). In the selected human proteins, in general, there are more Met, Phe, Ile, Leu, His and less Cys, Ala, Gly, Ser, Glu, Gln as compared with the human proteome. The bovine P450scc protein is enriched with Met, Phe, Ile, Trp, Tyr, Asn, His, Lys and depleted in Cys, Ala, Gly, Thr, Ser, Gln versus the bovine proteome. It is worth noting that in all the selected proteins there are less Ala and Gly than in the average proteome values. Human and bacterial proteins differ in charged amino acid residues Glu, Asp and Lys. Human proteins contain less of these amino acid residues than the human proteome values, and bacterial proteins contain more than the average bacterial proteome. The same is observed for Cys and Gln amino acid residues.

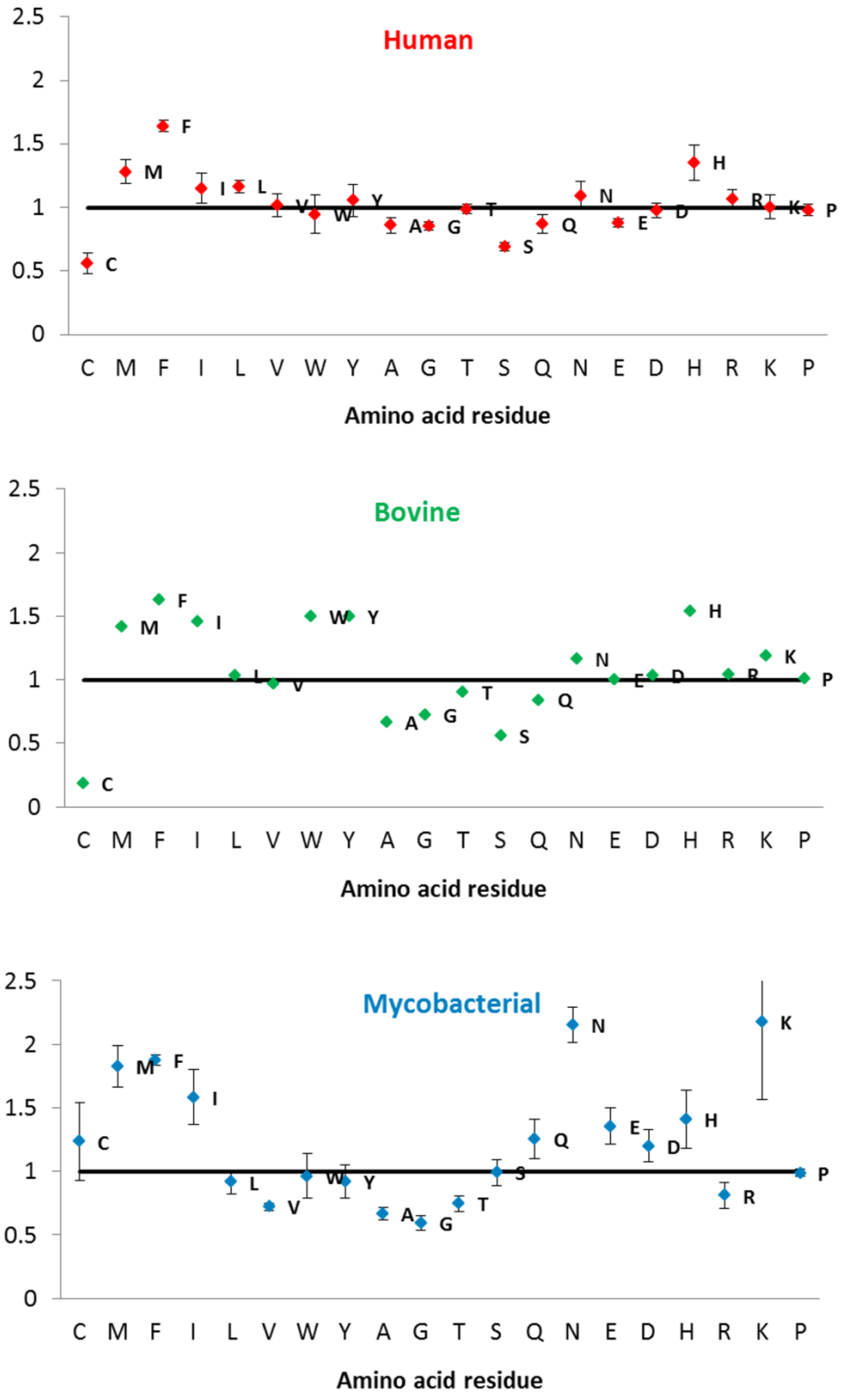


Figure 4. Comparison of the amino acid composition of the selected structures with the considered proteomes (human, bovine and mycobacterial).

3.2 Molecular dynamics simulations of bovine P450scc

3.2.1 Simulations of wild type P450scc and CYP7A1

Molecular dynamics simulations of structures 3v8d (at T = 350 K) and 3mzs (at T = 300 K) with the heme and cholesterol have been performed. According to the program IsUnstruct (http://bioinfo.protres.ru/IsUnstruct), these structures were selected for the following modeling because one of them (3v8d) has more and the other (3mzs) has less disordered residues (see Fig. 3) in comparison with other structures.

Fig. 5 represents graphs of changes in the secondary structure during the modeling. Straight lines on the graphs indicate the initial level of each type of secondary structure before the simulations. Fig. 5 shows that the increase in the fraction of irregular structure is caused by destruction of α-helices. In contrast, the β-structure is stabilized. The largest deviations from the initial structure were observed for 3v8d because its modeling was performed at a higher temperature.

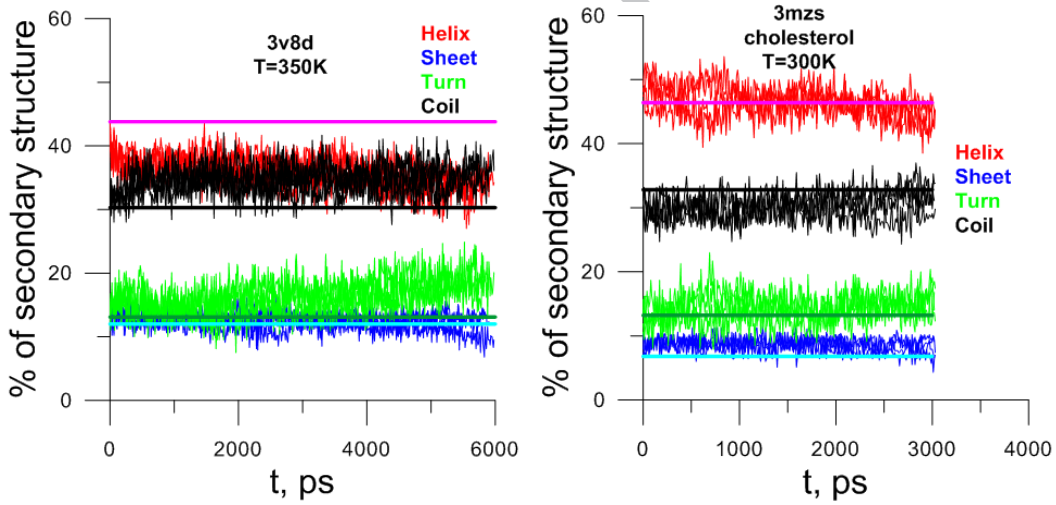


Figure 5. Different changes in secondary structure during simulations of structures 3v8d and 3mzs. Straight lines on the graphs indicate the initial level of each type of secondary structure before simulations.

Although a small change in the secondary structure during the simulations was observed, the root mean square deviation (RMSD) calculated from the C-alpha atoms between the initial structure and structures obtained after the simulation is within 3 Å (Fig. 6). Thus, these structures remain relatively stable over the simulation time.

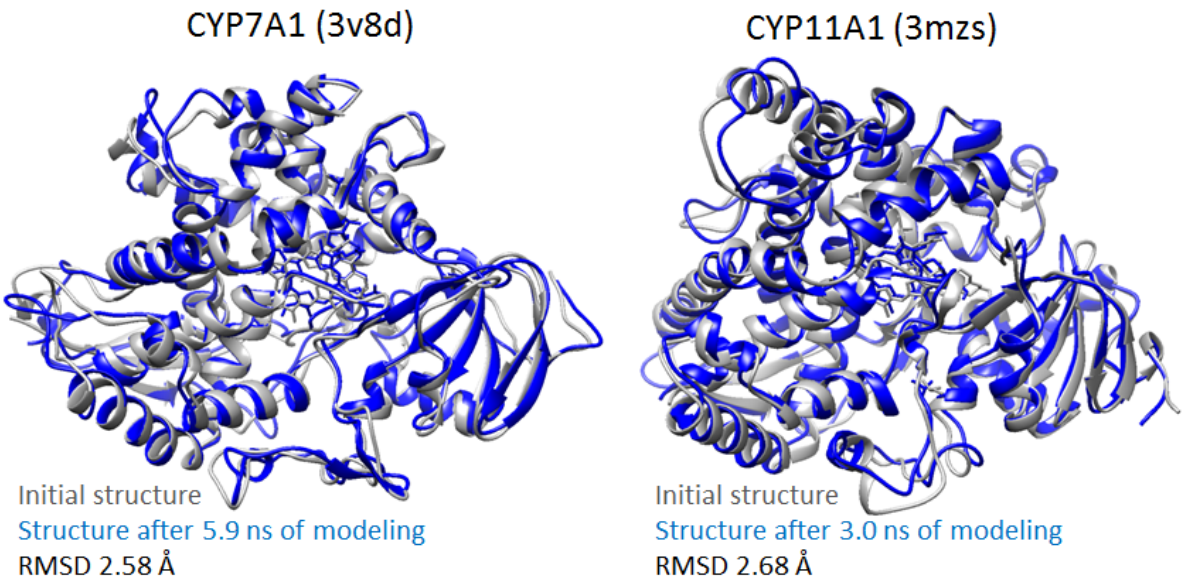


Figure 6. Spatial superposition of structures of 3v8d and 3mzs before (gray) and after modeling (blue).

3.2.2 Simulations of P450scc structures with mutated residues I351A, L355A, I461A

Further, it was necessary to select several mutations in protein 3mzs, leading to an increase in the size of the binding pocket, so that the sterol (cholesterol or β-sitosterol) molecule could fit there. To this effect, the number of atom-atom contacts at the distance of 6 Å for each amino acid residue in protein 3mzs with cholesterol was calculated. Amino acid residues involved somehow in the catalytic reaction according to Strushkevich et al. [14] have not been considered. By these criteria, the following mutations were selected: L102A, I351A, L355A and I461A. The affinities of β-sitosterol to various mutant proteins based on the structure of 3mzs were calculated by using the program for docking AutoDock Vina [34] (Table 2). According to the affinity of β-sitosterol, the best mutants were the following: single – I351A and L355A, double – I351A/L355A, and triple – I351A/L355A/I461A. In Table 2 they are represented in bold. We have found two SNPs only for Leu355 https://www.ncbi.nlm.nih.gov/projects/SNP/snp_ref.cgi?geneId=1583. SNP data confirms our result that I351, I461, and L355 replacements are lethal to the function of cytochrome P450scc. Two synonymous L355 substitutions are very interesting. It should be mentioned that leucine is encoded by 6 codons. In our case we have CTG for Leu, for SNP carriers there are CTA and TTG codons. We see them, but they are only synonymous.

Disposition of mutated I351, L355, and I461 residues is shown in Fig. 7. For the following simulations the structures with double I351A/L355A and triple I351A/L355A/I461A mutations with cholesterol and β-sitosterol were constructed using the YASARA program [35]. The molecular dynamics simulations of these mutant constructions were performed at 300 K.

Table 2. Calculated affinities for binding of β-sitosterol to different P450scc mutant proteins

Protein	Binding energy, kcal/mol
Native protein with cholesterol (oxidized form, C22)	-12.5
Native protein with cholesterol (normal form)	-12.2
Native protein with β-sitosterol	-12.1
I351A	**-11.9**
I351A/I461A	-11.7
I351A/L355A	**-11.9**
I461A	-11.9
L102A	-11.6
L102A/I351A	-11.4
L102A/I461A	-11.4
L102A/L355A	-11.5
L355	**-12.0**
L355A/I461A	-11.8
I351A/L355A/I461A	**-11.7**
L102A/I351A/I461A	-11.3
L102A/I351A/L355A	-11.4
L102A/L355A/I461A	-11.3
I351A/L355A/I461A/L102A	-11.2

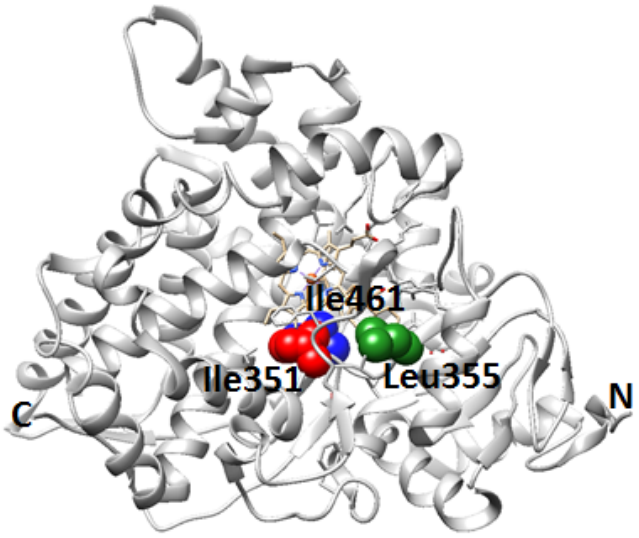
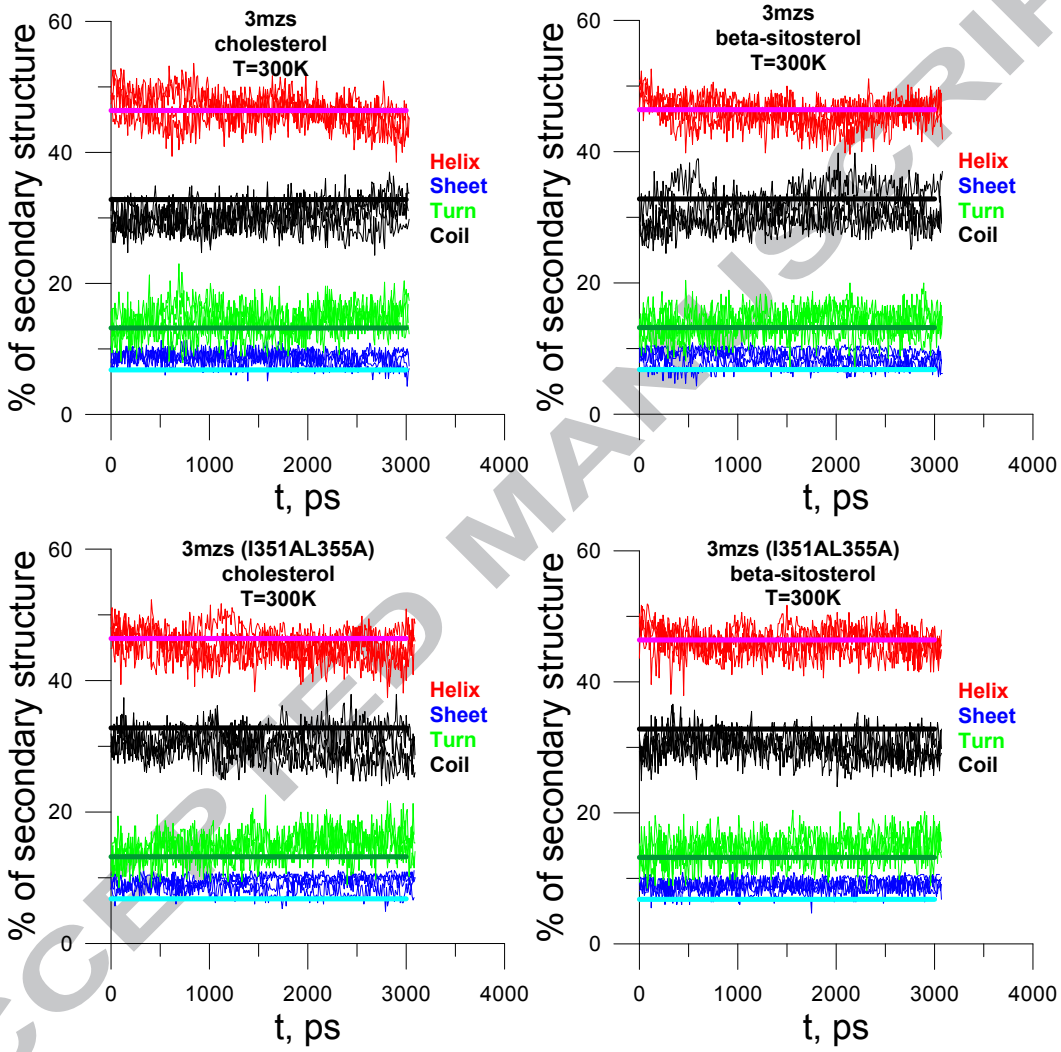


Figure 7. Spatial structure of bovine CYP11A1 (pdb id 3mzs). Mutated residues I351, L355 and I461 are shown as space-filling models.

As seen from Fig. 8, the introduced mutations do not destroy the secondary structure of the protein, and it remains stable during the simulation. The RMSD between the initial structures and structures after the simulation is about 3 Å. These mutations also do not dramatically change the size of the binding pocket.



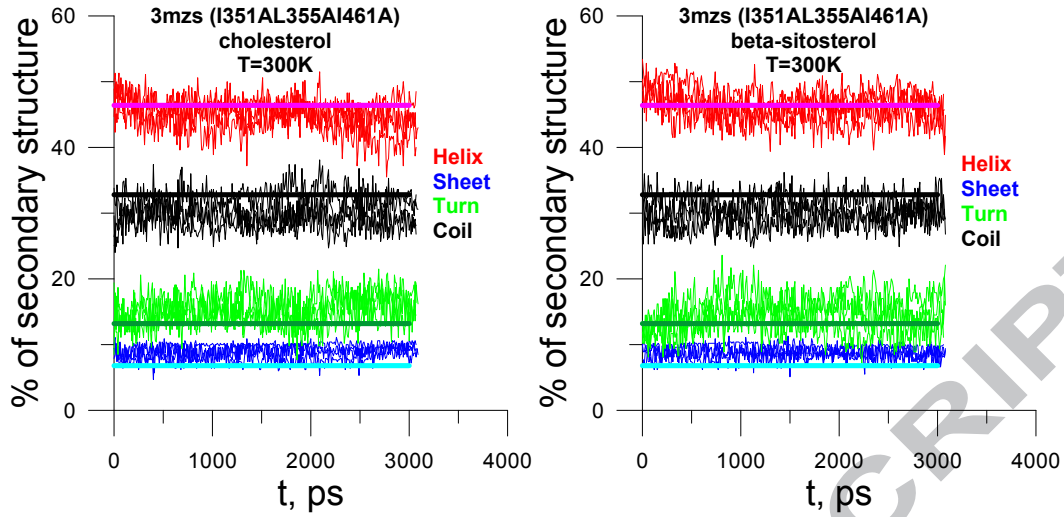


Figure 8. Different changes in the secondary structure during simulations of structure 3mzs and its double I351A/L355A and triple I351A/L355A/I461A mutant proteins with two substrates cholesterol and β-sitosterol. Straight lines on the graphs indicate the initial level of each type of secondary structure before simulations.

3.3 Effect of I351A, L355A, I461A mutations on the catalytic reaction rate of P450scc

Molecular dynamics simulations have demonstrated that the introduced mutations do not destroy the 3D structure of the cytochrome P450scc *in silico*. However, it is important to evaluate how these mutations will affect the catalytic reaction rate of the enzyme.

3.3.1 Mutagenesis of bovine P450scc

The I351A, L355A, I461A mutations were introduced into the cDNA encoding P450scc protein by site-directed mutagenesis. To this end, mutant variants of cDNA encoding P450scc with single, double and triple mutations were produced.

The generation of each of the mutant variants of the P450scc cDNA consisted of two steps. The first step was the production of PCR fragment encoding the cytochrome P450scc with the mutation using the oligonucleotide guided site-specific mutagenesis method. During the second step the resulting mutant PCR fragment was used to replace the corresponding fragment of the wild-type P450scc cDNA in the tricystronic pTrc99A/CHL plasmid.

All five resulting mutant DNA sequences (three single I351A, L355A, I461A, one double I351A/L355A and one triple I351A/L355A/I461A mutants) were validated with Sanger automatic DNA sequencing (the numbers of the amino acid residues correspond to their position in the sequence of the mature form of the protein). Based on the vector pTrc99A/CHL the generated plasmids carrying mutated cDNA of the cytochrome P450scc were designated as: pTrc99A/CHL(P450-I351A),

pTrc99A/CHL(P450-L355A), pTrc99A/CHL(P450-I461A), pTrc99A/CHL(P450-I351A/L355A) and pTrc99A/CHL(P450-I351A/L355A/I461A).

3.3.2 Expression of mutated P450scc proteins in *E. coli* cells

E. coli DH5αc cells transformed with the recombinant plasmids encoding the mutant and wild-type cytochrome P450scc were grown as described in 2.2.4 to induce expression of the heterologous cDNAs. Cell lysates were analyzed with SDS-PAGE, followed by immunoblotting for identification of the produced P450scc proteins. As a result of immunodetection with the antibodies against P450scc, we found that the proteins capable of binding to the specific antibodies have a molecular weight of ~53 kDa for all strains (Fig. 9, lanes 2-7). This molecular weight corresponds to the full-length P450scc. Moreover, the analysis did not reveal any significant differences in the expression levels of the wild-type P450scc and the proteins of P450scc with the introduced amino acid substitutions (mutations I351A, L355A, I461A, I351A/L355 and AI351A /L355A /I461A) (Fig. 9). Thus, the data obtained clearly indicate that the mutated P450scc proteins are stable and no protein degradation could be detected.

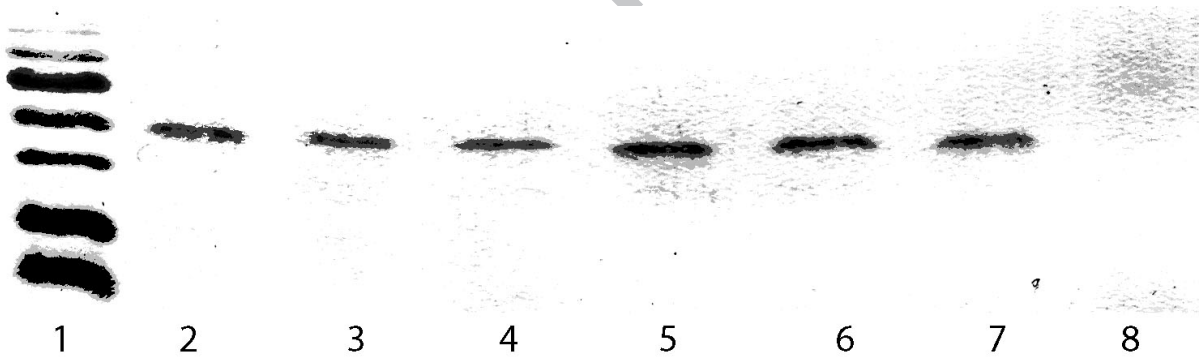


Figure 9. Western blot analysis of wild-type and mutant forms of P450scc expressed in *E. coli* cells. Protein separation of cell lysates (5 μg/lane) was performed by 10% SDS-PAAG. 1 – standard protein markers (130, 95, 72, 55, 43, 34 and 26 kDa) stained with Coomassie Brilliant Blue; 2 - 7 - lysates of *E. coli* cells transformed with pTrc99A/CHL, pTrc99A/CHL(P450-I351A), pTrc99A/CHL(P450-L355A), pTrc99A/CHL(P450-I461A), pTrc99A/CHL(P450-I351A/L355A), pTrc99A/CHL(P450I351A/L355A/I461A), respectively; 8 – lysate of control plasmid-less *E. coli* cells. The membrane (2-8 lanes) was probed with antibodies against P450scc.

3.3.2 Side-chain cleavage activity of P450scc mutant proteins

Sterol side-chain cleavage activity of the mutant and wild-type cytochrome P450scc proteins was analyzed as the rate of the side-chain cleavage from substrates: 20α-hydroxycholesterol and β-sitosterol. Substrate conversion rate to pregnenolone was recorded using the ELISA method. The results of the activity measurements (minus the background value obtained for the cell homogenate of control *E. coli* DH5αc strain) are shown in Table 3. Mean values (minus K) from 4 to 6 measurements using 20α-hydroxycholesterol and from 2 measurements conducted with β-sitosterol are represented. K, the activity measured in the homogenate of control untransformed (without P450scc) cells was 15.9 ± 6.3 (with 20α-hydroxycholesterol) and 0.57 ± 0.07 (with β-sitosterol) (pmol of pregnenolone/mg homogenate×h, respectively. As follows from the presented results, the introduction of any of the mutations leads to a sharp decrease or a complete loss of protein activity.

Table 3. Catalytic activity of wild-type and single mutant P450scc proteins (pmol of pregnenolone/mg homogenate×hour)*

Protein	20α-hydroxycholesterol	β-sitosterol
P450scc	540.00±9.40	129.00±29.50
I351A	11.45±8.03	1.24±0.36
L355A	383.00±7.92	21.40±7.70
I461A	12.09±4.50	0

*The data presented mean the values of activity measurement minus the background value obtained for the cell homogenate of control *E. coli* DH5αc strain.

The maximum recorded activity for a protein with a single amino acid replacement was found for P450scc mutant L355A, which is lower by ~ 1.4 times with 20α-hydroxycholesterol and ~ 6.5 times with β-sitosterol than the activity of native P450scc. I351A and I461A substitutions in the cytochrome P450scc molecule resulted in a dramatic decrease in activity with both substrates: 20α-hydroxycholesterol (more than 50 times) and with β-sitosterol (to almost/complete loss of activity). Moreover, the introduction of another additional mutation totally inactivates P450scc protein. All the double (I351A/L355A) and triple (I351A/L355A/I461A) mutants that we used were completely inactive.

4. Conclusions

We examined 14 structural homologues of P450scc CYP11A1 capable of interacting with cholesterol. These structures have low homology in the primary structure (about 25%), but are very similar in the 3D structure (RMSD is about 2 Å). It turned out that for these structures, only one amino

acid residue from those involved in the preparation of the correct conformation of cholesterol is conservative (Ile461). However, the amino acid residues that are involved in electron transfer are strictly conservative (Gly287 and Thr291). For these 14 proteins the heme binding site and the areas in the vicinity of the catalytic center are structurally very similar. The results of docking of cholesterol and β-sitosterol suggest that for half of the structures (7 of 14) the substrate molecules have the free energy similar to that of P450scc in the catalytic pocket. Molecular dynamics simulations of P450scc (pdb 3mzs) and its double (I351A/L355A) and triple (I351A/L355A/I461A) mutants with cholesterol and β-sitosterol suggest that these structures are stable during 3 ns simulations. However, the experimental data have proved that the introduced single mutations, either I351A, L355A and I461A dramatically decrease the catalytic activity of P450scc with both substrates: cholesterol and β-sitosterol. Thus, we can conclude that the catalytic activity of the enzyme is affected by the replacement of amino acid residues not directly interacting with the substrate and proved to be essential for the function

The results contribute to the better understanding of structure-function relationship of the unique and key enzyme of steroidogenesis – cytochrome P450scc.

Acknowledgments: We are grateful to T.B. Kuvshinkina for assistance in preparation of the manuscript. This work was supported by the Russian Science Foundation (No. 18-14-00321) for OVG, AVG and NVD. MD, MK and NS are greatly acknowledged the Russian Science Foundation (Grant No. 18-14-00361) for the support of their work; LN, LI, VE and MR are greatly acknowledged the Russian Foundation for Basic Research (Grant No. 16-54-00139). BM is greatly acknowledged the Nazarbayev University ORAU. The work was supported by the Ministry of Science and Higher Education of the Russian Federation (agreement No. 075-11-2018-096 / 14.588.21.0008) within the framework of implementing the Federal Target Program (for MVD, MVK and NIS).

5. References

[1] R. Bernhardt, Cytochromes P450 as versatile biocatalysts, J. Biotechnol. 124 (2006) 128–145. doi:10.1016/j.jbiotec.2006.01.026.

[2] P.R.O. de Montellano, ed., Cytochrome P450: Structure, Mechanism, and Biochemistry, 4th ed., Springer International Publishing, 2015. //www.springer.com/us/book/9783319121079 (accessed June 28, 2018).

[3] R. Bernhardt, V.B. Urlacher, Cytochromes P450 as promising catalysts for biotechnological application: chances and limitations, Appl. Microbiol. Biotechnol. 98 (2014) 6185–6203. doi:10.1007/s00253-014-5767-7.

[4] H. al Kandari, N. Katsumata, S. Alexander, M.A. Rasoul, Homozygous mutation of P450 side-chain cleavage enzyme gene (CYP11A1) in 46, XY patient with adrenal insufficiency, complete sex reversal, and agenesis of corpus callosum, J. Clin. Endocrinol. Metab. 91 (2006) 2821–2826. doi:10.1210/jc.2005-2230.

[5] S. Parajes, A.O.K. Chan, W.M. But, I.T. Rose, A.E. Taylor, V. Dhir, W. Arlt, N. Krone, Delayed diagnosis of adrenal insufficiency in a patient with severe penoscrotal hypospadias due to two novel P450 side-

change cleavage enzyme (CYP11A1) mutations (p.R360W; p.R405X), Eur. J. Endocrinol. 167 (2012) 881–885. doi:10.1530/EJE-12-0450.

[6] M.K. Tee, M. Abramsohn, N. Loewenthal, M. Harris, S. Siwach, A. Kaplinsky, B. Markus, O. Birk, V.C. Sheffield, R. Parvari, R. Pavari, E. Hershkovitz, W.L. Miller, Varied clinical presentations of seven patients with mutations in CYP11A1 encoding the cholesterol side-chain cleavage enzyme, P450scc, J. Clin. Endocrinol. Metab. 98 (2013) 713–720. doi:10.1210/jc.2012-2828.

[7] T. Sahakitrungruang, M.K. Tee, P.R. Blackett, W.L. Miller, Partial defect in the cholesterol side-chain cleavage enzyme P450scc (CYP11A1) resembling nonclassic congenital lipoid adrenal hyperplasia, J. Clin. Endocrinol. Metab. 96 (2011) 792–798. doi:10.1210/jc.2010-1828.

[8] B. Wahlang, K.C. Falkner, M.C. Cave, R.A. Prough, Role of Cytochrome P450 Monooxygenase in Carcinogen and Chemotherapeutic Drug Metabolism, Adv. Pharmacol. San Diego Calif. 74 (2015) 1–33. doi:10.1016/bs.apha.2015.04.004.

[9] D. Werck-Reichhart, R. Feyereisen, Cytochromes P450: a success story, Genome Biol. 1 (2000) REVIEWS3003. doi:10.1186/gb-2000-1-6-reviews3003.

[10] F. Hannemann, A. Bichet, K.M. Ewen, R. Bernhardt, Cytochrome P450 systems--biological variations of electron transport chains, Biochim. Biophys. Acta. 1770 (2007) 330–344. doi:10.1016/j.bbagen.2006.07.017.

[11] T.L. Poulos, B.C. Finzel, I.C. Gunsalus, G.C. Wagner, J. Kraut, The 2.6-A crystal structure of Pseudomonas putida cytochrome P-450, J. Biol. Chem. 260 (1985) 16122–16130.

[12] H. Sugimoto, Y. Shiro, Diversity and substrate specificity in the structures of steroidogenic cytochrome P450 enzymes, Biol. Pharm. Bull. 35 (2012) 818–823.

[13] F.P. Guengerich, M.R. Waterman, M. Egli, Recent Structural Insights into Cytochrome P450 Function, Trends Pharmacol. Sci. 37 (2016) 625–640. doi:10.1016/j.tips.2016.05.006.

[14] N. Strushkevich, F. MacKenzie, T. Cherkesova, I. Grabovec, S. Usanov, H.-W. Park, Structural basis for pregnenolone biosynthesis by the mitochondrial monooxygenase system, Proc. Natl. Acad. Sci. U. S. A. 108 (2011) 10139–10143. doi:10.1073/pnas.1019441108.

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

[16] S. Graham-Lorence, J.A. Peterson, P450s: structural similarities and functional differences, FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol. 10 (1996) 206–214.

[17] S.E. Graham-Lorence, J.A. Peterson, Structural alignments of P450s and extrapolations to the unknown, Methods Enzymol. 272 (1996) 315–326.

[18] K.H. Storbeck, P. Swart, S. Graham, A.C. Swart, The influence of the amino acid substitution I98K on the catalytic activity of baboon cytochrome P450 side-chain cleavage (CYP11A1), Endocr. Res. 30 (2004) 761–767.

[19] A.T. Slominski, W. Li, T.-K. Kim, I. Semak, J. Wang, J.K. Zjawiony, R.C. Tuckey, Novel activities of CYP11A1 and their potential physiological significance, J. Steroid Biochem. Mol. Biol. 151 (2015) 25–37. doi:10.1016/j.jsbmb.2014.11.010.

[20] N. Mast, A.J. Annalora, D.T. Lodowski, K. Palczewski, C.D. Stout, I.A. Pikuleva, Structural basis for three-step sequential catalysis by the cholesterol side chain cleavage enzyme CYP11A1, J. Biol. Chem. 286 (2011) 5607–5613. doi:10.1074/jbc.M110.188433.

[21] S. Vijayakumar, J.C. Salerno, Molecular modeling of the 3-D structure of cytochrome P-450scc, Biochim. Biophys. Acta. 1160 (1992) 281–286.

[22] S.A. Usanov, S.E. Graham, G.I. Lepesheva, T.N. Azeva, N.V. Strushkevich, A.A. Gilep, R.W. Estabrook, J.A. Peterson, Probing the interaction of bovine cytochrome P450scc (CYP11A1) with adrenodoxin: evaluating site-directed mutations by molecular modeling, Biochemistry. 41 (2002) 8310–8320.

[23] M.J. Headlam, M.C.J. Wilce, R.C. Tuckey, The F-G loop region of cytochrome P450scc (CYP11A1) interacts with the phospholipid membrane, Biochim. Biophys. Acta. 1617 (2003) 96–108.

[24] V. Sivozhelezov, C. Nicolini, Homology modeling of cytochrome P450scc and the mutations for optimal amperometric sensor, J. Theor. Biol. 234 (2005) 479–485. doi:10.1016/j.jtbi.2004.12.008.

[25] K.-H. Storbeck, P. Swart, A.C. Swart, Cytochrome P450 side-chain cleavage: insights gained from homology modeling, Mol. Cell. Endocrinol. 265–266 (2007) 65–70. doi:10.1016/j.mce.2006.12.005.

[26] I.A. Pikuleva, Putative F-G loop is involved in association with the membrane in P450scc (P450 11A1), Mol. Cell. Endocrinol. 215 (2004) 161–164. doi:10.1016/j.mce.2003.11.005.
[27] I.A. Pikuleva, N. Mast, W.-L. Liao, I.V. Turko, Studies of membrane topology of mitochondrial cholesterol hydroxylases CYPs 27A1 and 11A1, Lipids. 43 (2008) 1127–1132. doi:10.1007/s11745-008-3234-x.
[28] A. Wada, M.R. Waterman, Identification by site-directed mutagenesis of two lysine residues in cholesterol side chain cleavage cytochrome P450 that are essential for adrenodoxin binding, J. Biol. Chem. 267 (1992) 22877–22882.
[29] M. Tsujita, Y. Ichikawa, Substrate-binding region of cytochrome P-450SCC (P-450 XIA1). Identification and primary structure of the cholesterol binding region in cytochrome P-450SCC, Biochim. Biophys. Acta. 1161 (1993) 124–130.
[30] D.F. Lewis, P. Lee-Robichaud, Molecular modelling of steroidogenic cytochromes P450 from families CYP11, CYP17, CYP19 and CYP21 based on the CYP102 crystal structure, J. Steroid Biochem. Mol. Biol. 66 (1998) 217–233.
[31] E.F. Pettersen, T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng, T.E. Ferrin, UCSF Chimera--a visualization system for exploratory research and analysis, J. Comput. Chem. 25 (2004) 1605–1612. doi:10.1002/jcc.20084.
[32] O.V. Galzitskaya, S.O. Garbuzynskiy, M.Y. Lobanov, FoldUnfold: web server for the prediction of disordered regions in protein chain, Bioinforma. Oxf. Engl. 22 (2006) 2948–2949. doi:10.1093/bioinformatics/btl504.
[33] M.Y. Lobanov, I.V. Sokolovskiy, O.V. Galzitskaya, IsUnstruct: prediction of the residue status to be ordered or disordered in the protein chain by a method based on the Ising model, J. Biomol. Struct. Dyn. 31 (2013) 1034–1043. doi:10.1080/07391102.2012.718529.
[34] O. Trott, A.J. Olson, AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading, J. Comput. Chem. 31 (2010) 455–461. doi:10.1002/jcc.21334.
[35] E. Krieger, G. Koraimann, G. Vriend, Increasing the precision of comparative models with YASARA NOVA--a self-parameterizing force field, Proteins. 47 (2002) 393–402.
[36] M.A. Mazo, L.I. Manevitch, E.B. Gusarova, M.Y. Shamaev, A.A. Berlin, N.K. Balabaev, G.C. Rutledge, Molecular dynamics simulation of thermomechanical properties of montmorillonite crystal. 3. montmorillonite crystals with PEO oligomer intercalates, J. Phys. Chem. B. 112 (2008) 3597–3604. doi:10.1021/jp076028f.
[37] A.V. Glyakina, N.K. Balabaev, O.V. Galzitskaya, Two-, three-, and four-state events occur in the mechanical unfolding of small protein L using molecular dynamics simulation, Protein Pept. Lett. 17 (2010) 92–103.
[38] J. Wang, P. Cieplak, P.A. Kollman, How well a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules?, J. Comp. Chem. 21 (2000) 1049–1074.
[39] A.S. Lemak, N.K. Balabaev, A comparison between collisional dynamics and brownian dynamics., Mol. Simul. 15 (1995) 223–231.
[40] A.S. Lemak, N.K. Balabaev, Molecular dynamics simulation of polymer chain in solution by collisional dynamics method., J. Comp. Chem. 17 (1996) 1685–1695.
[41] M.P. Allen, D.J. Tildesley, Computer Simulation of Liquids, Clarendon, Oxford, 1987.
[42] W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, Comparison of simple potential functions for simulating liquid water., J. Chem. Phys. 79 (1983) 926–935.
[43] G.A. Evans, Molecular cloning: A laboratory manual. Second edition. Volumes 1, 2, and 3. Current protocols in molecular biology. Volumes 1 and 2: By J. Sambrook, E. F. Fritsch, and T. Maniatis. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press. (1989). 1626 pp. $115.00. Edited by F. M. Ausubel, R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl. New York: Greene Publishing Associates and John Wiley & Sons. (1989). 1120 pp. $255.00, Cell. 61 (1990) 17–18. doi:10.1016/0092-8674(90)90210-6.
[44] D.S. Makeeva, D.V. Dovbnya, M.V. Donova, L.A. Novikova, Functional reconstruction of bovine P450scc steroidogenic system in Escherichia coli, Am. J. Mol. Biol. 3 (2013) 73–182.

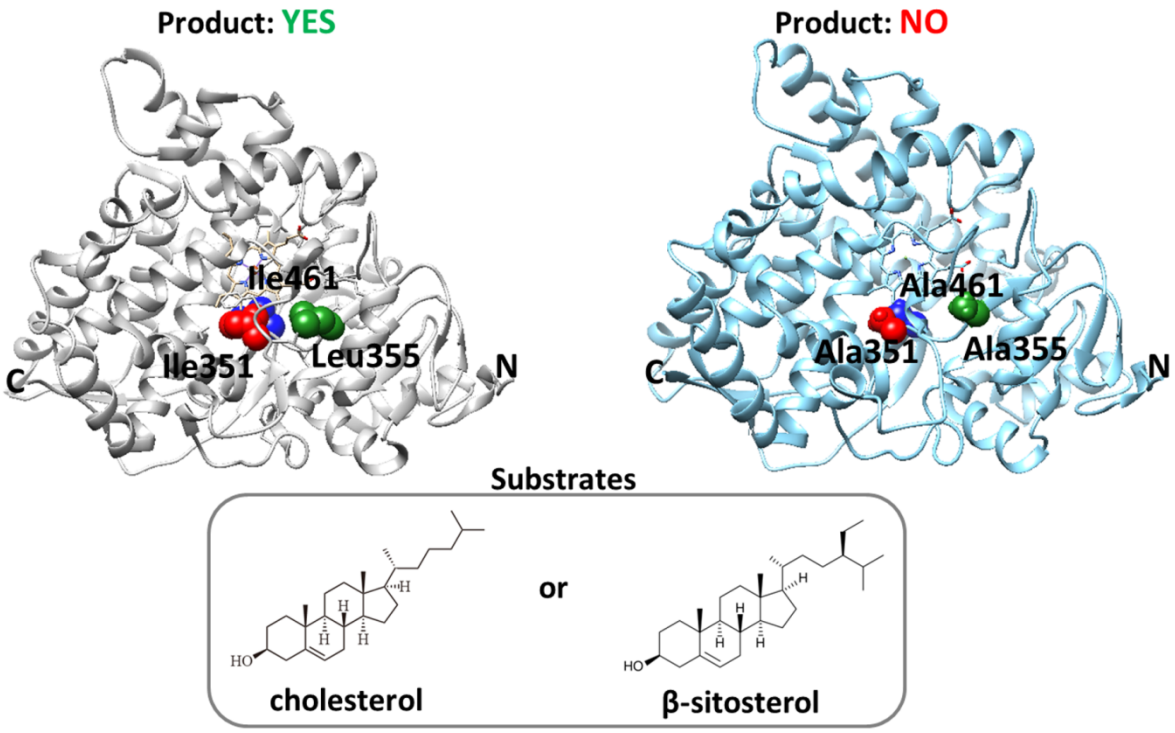
[45] Y. Sagara, Y. Takata, T. Miyata, T. Hara, T. Horiuchi, Cloning and sequence analysis of adrenodoxin reductase cDNA from bovine adrenal cortex, J. Biochem. (Tokyo). 102 (1987) 1333–1336.
[46] H. Uhlmann, R. Kraft, R. Bernhardt, C-terminal region of adrenodoxin affects its structural integrity and determines differences in its electron transfer function to cytochrome P-450, J. Biol. Chem. 269 (1994) 22557–22564.
[47] F. Hannemann, C. Virus, R. Bernhardt, Design of an Escherichia coli system for whole cell mediated steroid synthesis and molecular evolution of steroid hydroxylases, J. Biotechnol. 124 (2006) 172–181. doi:10.1016/j.jbiotec.2006.01.009.
[48] P.A. Nazarov, V.L. Drutsa, W.L. Miller, V.M. Shkumatov, V.N. Luzikov, L.A. Novikova, Formation and functioning of fused cholesterol side-chain cleavage enzymes, DNA Cell Biol. 22 (2003) 243–252. doi:10.1089/104454903321908638.
[49] V.S. Efimova, L.V. Isaeva, D.S. Makeeva, M.A. Rubtsov, L.A. Novikova, Expression of Cholesterol Hydroxylase/Lyase System Proteins in Yeast S. cerevisiae Cells as a Self-Processing Polyprotein, Mol. Biotechnol. 59 (2017) 394–406. doi:10.1007/s12033-017-0028-5.
[50] U.K. Laemmli, Cleavage of structural proteins during the assembly of the head of bacteriophage T4, Nature. 227 (1970) 680–685.
[51] H. Towbin, T. Staehelin, J. Gordon, Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications, Proc. Natl. Acad. Sci. U. S. A. 76 (1979) 4350–4354.
[52] O.H. Lowry, N.J. Rosebrough, A.L. Farr, R.J. Randall, Protein measurement with the Folin phenol reagent, J. Biol. Chem. 193 (1951) 265–275.
[53] T.V. Shashkova, V.N. Luzikov, L.A. Novikova, Coexpression of all constituents of the cholesterol hydroxylase/lyase system in Escherichia coli cells, Biochem. Biokhimiia. 71 (2006) 810–814.
[54] A.A. Vinogradova, V.N. Luzikov, L.A. Novikova, Comparative study of topogenesis of cytochrome P450scc (CYP11A1) and its hybrids with adrenodoxin expressed in Escherichia coli cells, Biochem. Biokhimiia. 72 (2007) 208–214.

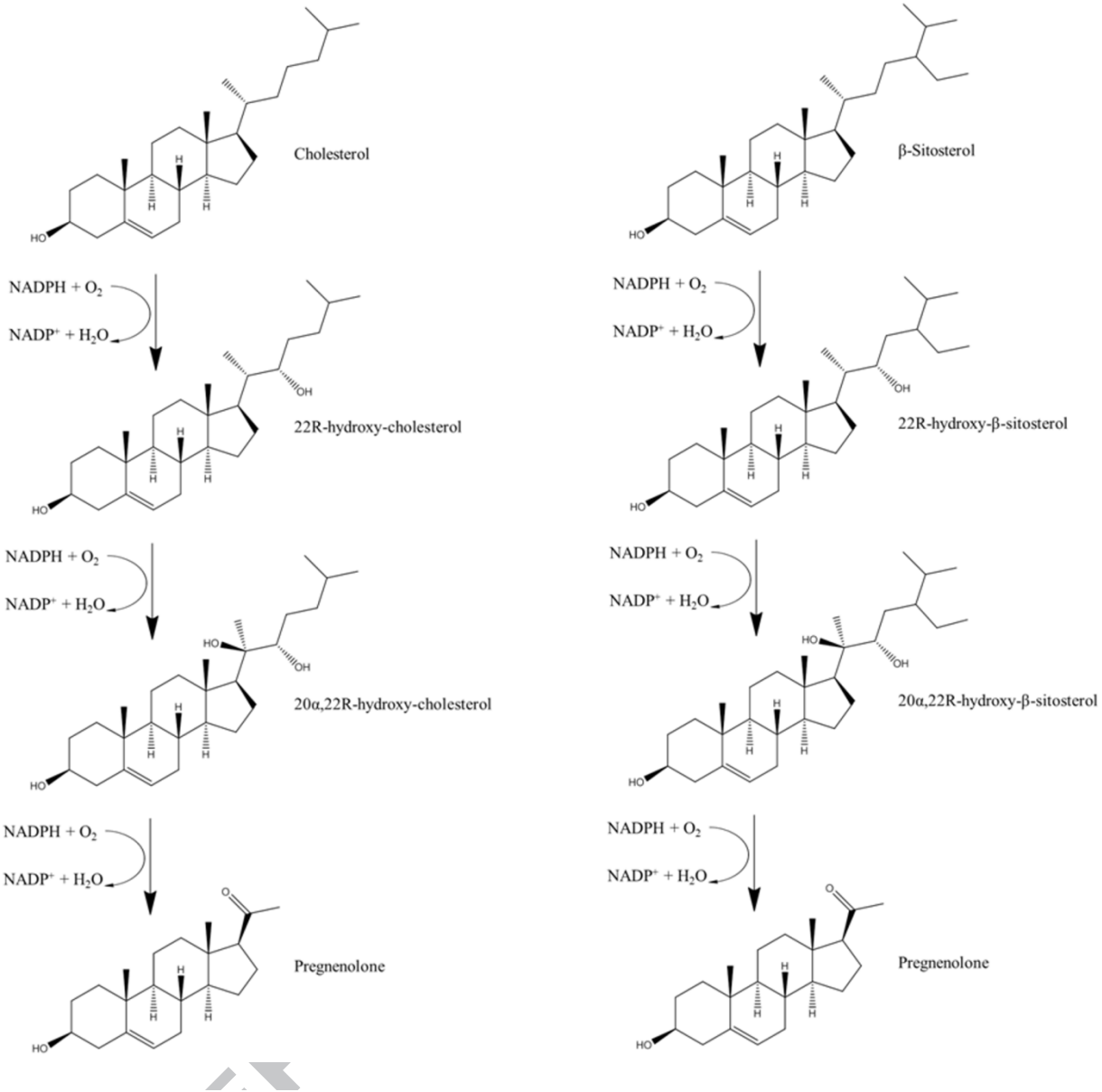
Highlights

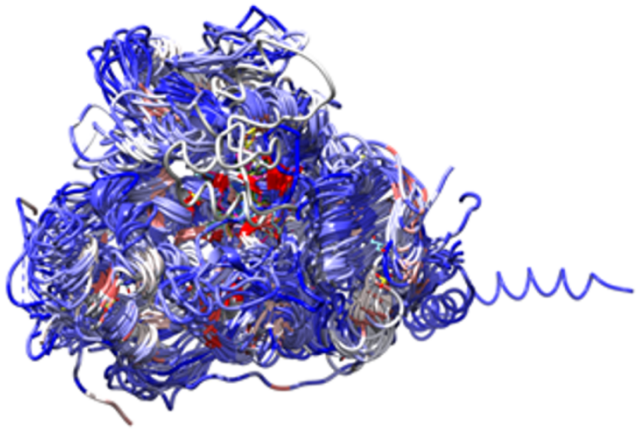
The considered cytochrome P450 enzymes have low sequence identities, but quite similar three-dimensional structures.

The cytochrome P450 proteins are depleted in Ala and Gly and enriched with Met, Phe, Ile versus the proteome values.

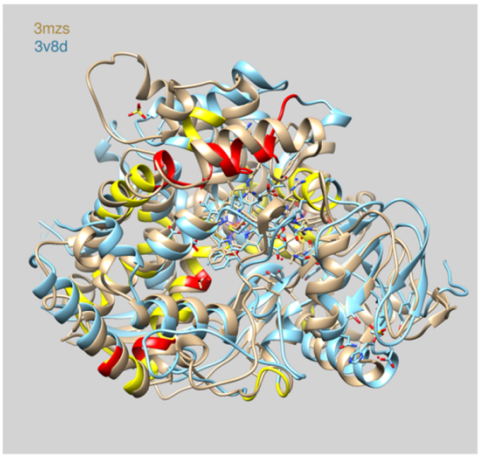
Residues Ile351, Leu355 and Ile461 are essential for the functioning of cytochrome P450scc towards cholesterol and β-sitosterol.



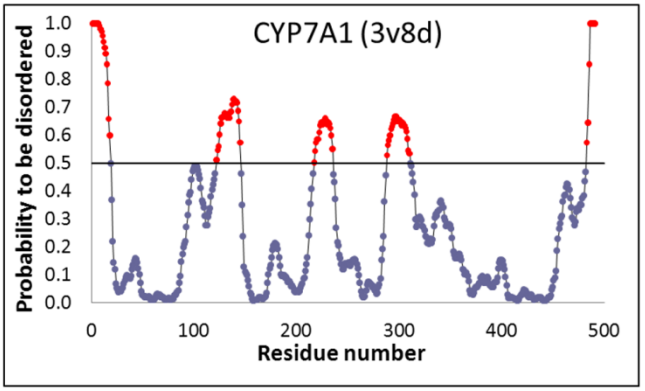

Cholesterol
β-Sitosterol
NADPH + O2
NADP+ + H2O
22R-hydroxy-cholesterol
22R-hydroxy-β-sitosterol
20α,22R-hydroxy-cholesterol
20α,22R-hydroxy-β-sitosterol
Pregnenolone
Pregnenolone



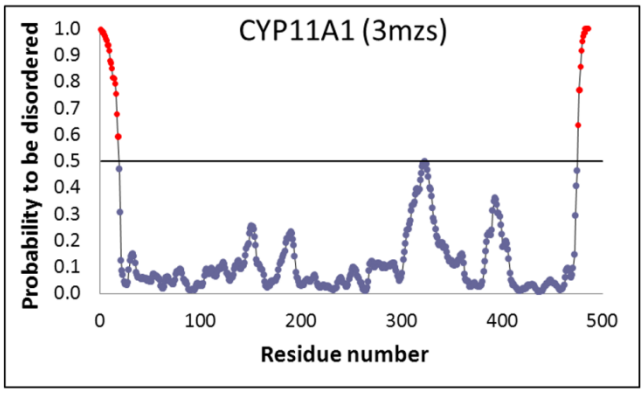
A

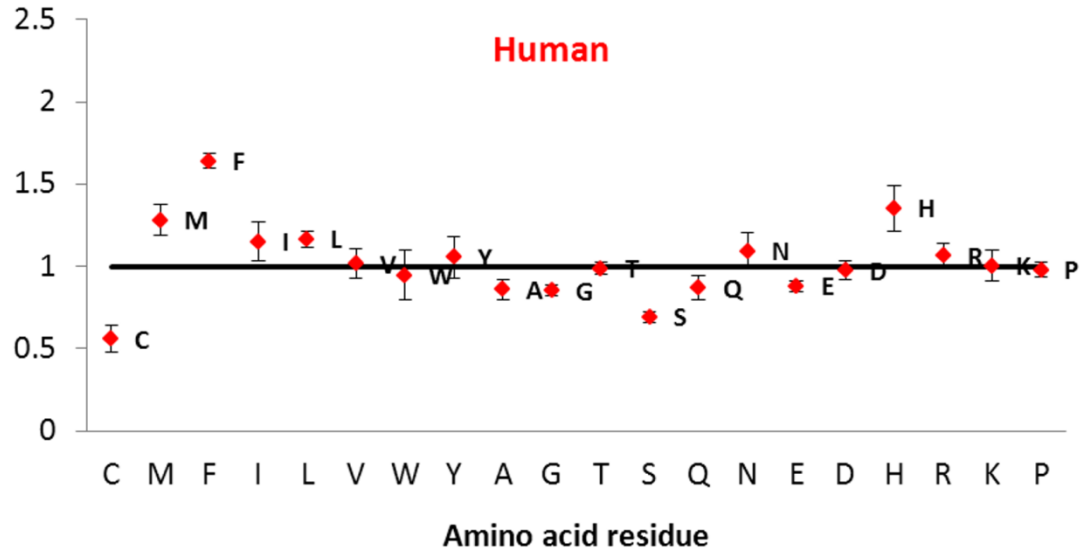


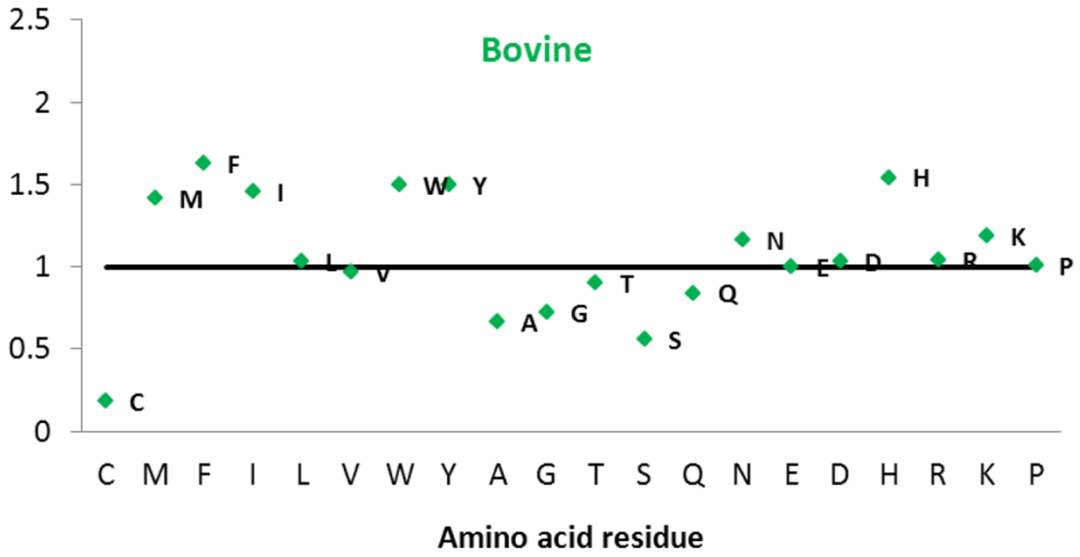
B

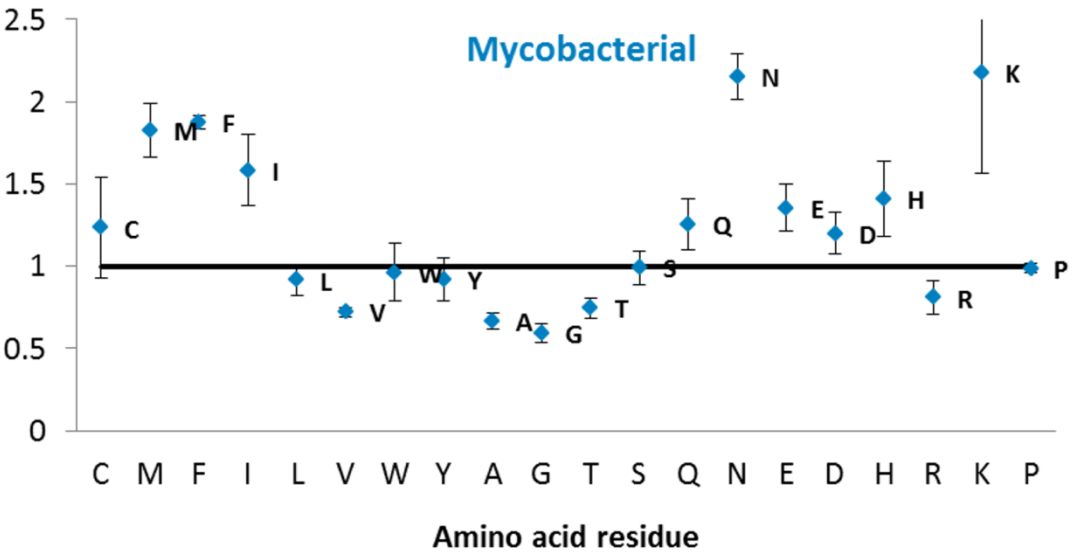


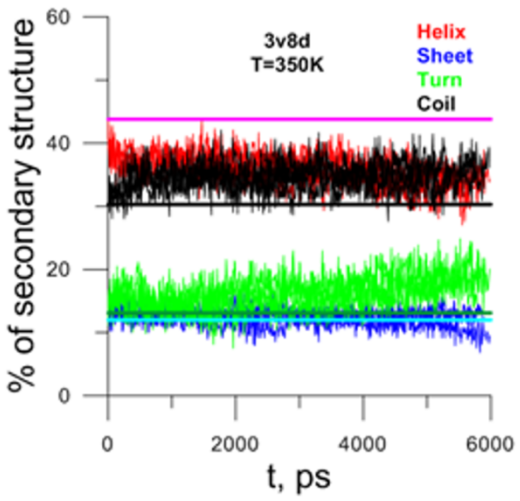
C

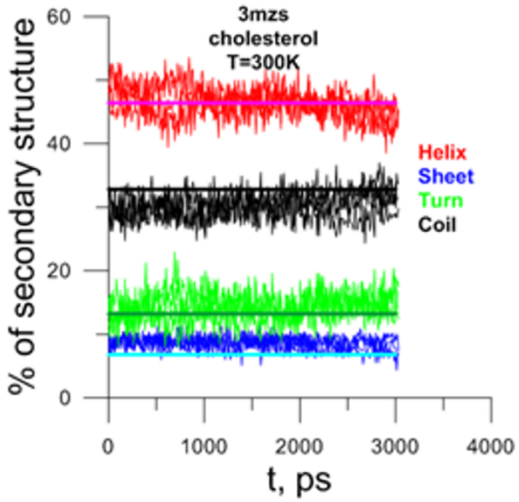




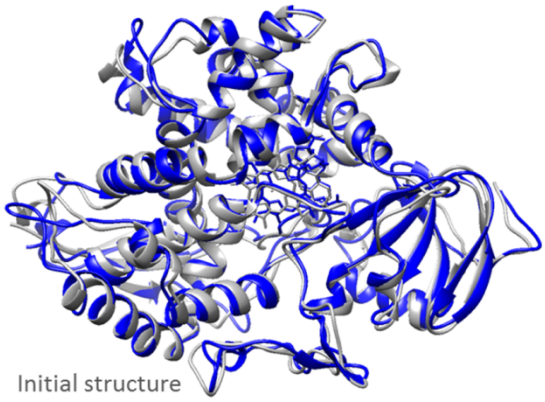






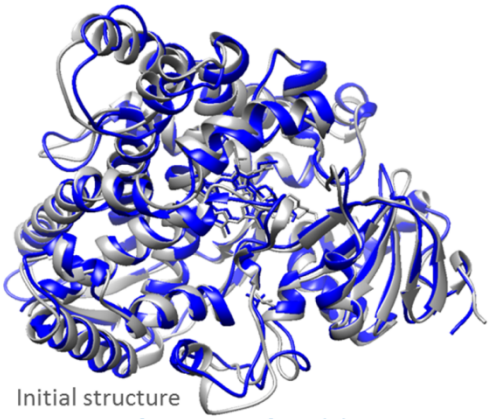


CYP7A1 (3v8d)

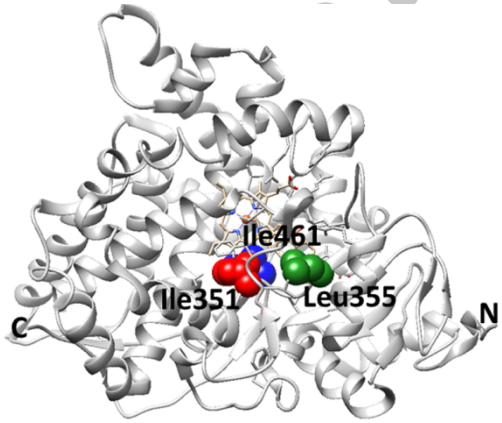


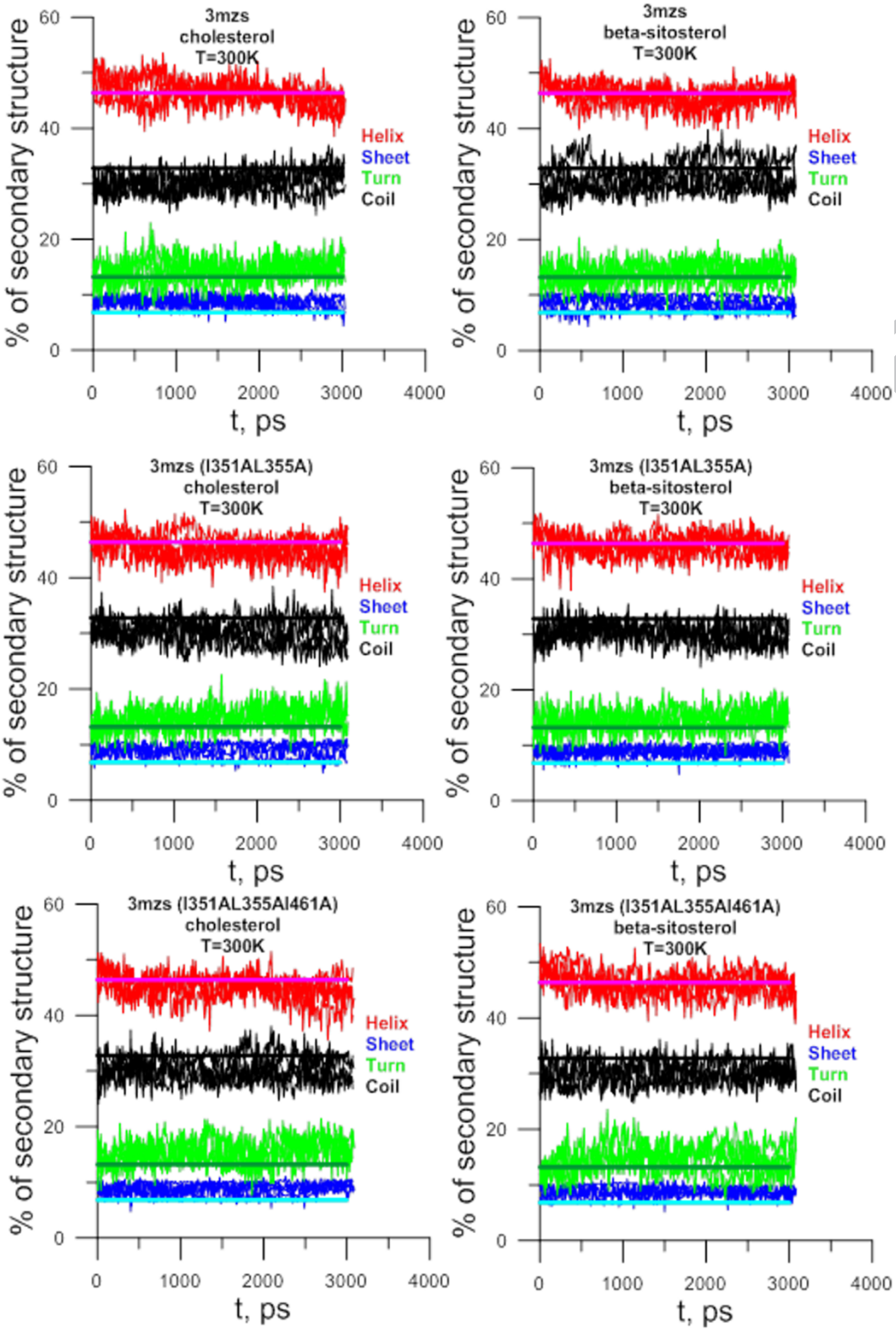
Initial structure
Structure after 5.9 ns of modeling
RMSD 2.58 Å

CYP11A1 (3mzs)



Initial structure
Structure after 3.0 ns of modeling
RMSD 2.68 Å




3mzs
cholesterol
T=300K
3mzs
beta-sitosterol
T=300K
3mzs (I351AL355A)
cholesterol
T=300K
3mzs (I351AL355A)
beta-sitosterol
T=300K
3mzs (I351AL355AI461A)
cholesterol
T=300K
3mzs (I351AL355AI461A)
beta-sitosterol
T=300K
% of secondary structure
t, ps
Helix
Sheet
Turn
Coil
0
20
40
60
0
1000
2000
3000
4000