

Supporting Information

Molecular Dynamics Simulations Reveal Structural Differences Among Allelic Variants of Membrane-Anchored Cytochrome P450 2D6

*André Fischer, Charleen G. Don, Martin Smiesko**

Molecular Modeling, Pharmacenter of the University of Basel,
Klingelbergstrasse 50, 4056 Basel, Switzerland
martin.smiesko@unibas.ch

Table S1. The stages of Desmond MD relaxation and simulation prior to the actual MD simulation.

<i>Desmond Stage</i>	<i>Procedure</i>
Stage 1	Task (reading files, initializing parameters)
Stage 2	Simulate, Brownian Dynamics NVT, T = 10 K, small time steps, and restraints on solute heavy atoms, 100 ps
Stage 3	Simulate, NVT, T = 10 K, small time steps, and restraints on solute heavy atoms, 12 ps
Stage 4	Simulate, NPT, T = 10 K, and restraints on solute heavy atoms, 12 ps
Stage 5	Solvate pocket
Stage 6	Simulate, NPT and restraints on solute heavy atoms, 12 ps
Stage 7	Simulate, NPT and no restraints, 24 ps

The seven stages were derived from the log file of a Desmond MD simulation.

Table S2. Predictions for membrane-bound residues from web-based protocols.

<i>Protocol</i>	<i>Prediction N-terminus</i>
TMHMM v2.0	A5-H24
Amphipaseek	V7-D21
TMSEG	L6-H24

Three web based protocols were employed to get an estimate of the membrane-bound residues in the transmembrane helix of CYP2D6. The predictions were used to place the helix in a POPC membrane.

Table S3. Experimental validation of the CYP2D6 WT model adapted from Cojocaru et al.

<i>Residues in CYP2D6 (CYP2C9*)</i>	<i>Result</i>	<i>Correlation</i>	<i>Experimental result*</i>	<i>Experimental method*</i>
1-35 (1-30)	M-HG-C	+	Inaccessible	Site-directed antibody ⁴⁸
18-32 (17-28)	M-HG-C ^b	0	Accessible	Site-directed antibody ⁴⁹
27-41 (23-37)	HG-C-HG-M	+	Accessible	Site-directed antibody ⁴⁸
43-51 (39-47)	M-HG-C	+	Inaccessible	Site-directed antibody ⁴⁸
64-75 (60-71)	C-HG-M	+	Accessible	Site-directed antibody ⁴⁸
96-101 (92-97)	C	-	Inaccessible	Site-directed antibody ⁴⁸
111-122 (107-115)	HG-C	+	Accessible	Site-directed antibody ⁴⁸
129-138 (121-130)	C	+	Accessible	Site-directed antibody ⁴⁸
194-201 (185-192)	C	+	Accessible	Site-directed antibody ⁴⁸
218-229 (210-222)	HG-M	+	Inaccessible	Site-directed antibody ⁴⁸
231-238 (224-231)	M-HG	-	Accessible	Site-directed antibody ⁴⁸
322-330 (314-322)	C	+	Accessible	Site-directed antibody ⁴⁸
404-414 (397-407)	C	+	Accessible	Site-directed antibody ⁴⁸
40 (36)	M	+	M	Trp fluorescence quenching ⁴⁶
73 (69)	C ^a	0	M	Trp fluorescence quenching ⁴⁶
387 (380)	C	-	M	Trp fluorescence quenching ⁴⁶
84, 128, 198, 246, 355 (80,120, 189, 239, 347)	C	+	C (HG)	Trp fluorescence quenching ⁴⁶
232 (225)	M	-	C (HG)	Trp fluorescence quenching ⁴⁶

1. Column: Residues in CYP2D6 data based on sequence alignment to CYP2C9. The corresponding residues in CYP2C9 are shown in brackets.

2. Column: Location of the residues in our model related to the membrane. The order of the location is given from N-terminus to C-terminus. The residues can either be positioned in the membrane (M), headgroup region (HG), or cytosol (C).

3. Column: Correlation of our results with the experimental data: (+) indicates a correlation, (0) indicates inconclusive results, and (-) indicates a missing correlation.

4. Column: Experimental results on the location and accessibility of the related protein regions. Data from site-directed antibodies is divided in “accessible” and “inaccessible”, while data from tryptophan fluorescence quenching directly indicates the location of the residues.

5. Column: The experimental methods used to produce the mentioned data on the residue location. The referred references are listed in the main article.

*data directly adapted from Cojocaru et al.¹⁸

^a not inside membrane, but buried inside the protein

^b fewer than 10% of the residues inside the range match the experimental data

Table S4. Values of the control simulations.

<i>System</i>	<i>Simulation time</i>	<i>Comment</i> ^a	<i>Heme tilt angle</i> ^b	<i>Burying depth</i> ^b	<i>Location of residues</i> ^c		<i>Overall Appearance</i>
					<i>Q27</i>	<i>R28</i>	
AnchorSim	200 ns	Used for final system	n/a	n/a	HG	HG	angled, not straight
Anchor1	129 ns	Different burying depth	n/a	n/a	HG	S	angled, straight
Anchor2	40 ns	Different burying depth	n/a	n/a	HG	S	angled, not straight
Anchor3	200 ns	Different burying depth	n/a	n/a	HG	S	angled, not straight
Anchor4	200 ns	Different burying depth	n/a	n/a	HG	HG	angled, straight
Anchor 5	200 ns	Different burying depth	n/a	n/a	HG	HG	angled, not straight
MemAs	100 ns	Used for final system	64.4°	41.4 Å	-	-	
Mem1	100 ns	Different burying depth	77.9°	37.1 Å	-	-	
Mem2	100 ns	Different burying depth	56.1°	39.9 Å	-	-	
CYP2DT WT	300 ns	Used for final system	44.4°	36.1 Å	HG	HG	
Fusion1	47 ns	Different linkage ^d	64.3°	36.3 Å	S	S	
Fusion2	65 ns	Different linkage ^d	39.5°	41.0 Å	HG	HG	
Fusion3	200 ns	Different linkage ^d	50.1°	34.7 Å	HG	HG	anchor moved away from protein

^a Background of the simulations.

^b Final value after the whole simulation.

^c The location of two selected residues at the protein-membrane interface was determined in the last frame of the respective simulation. Divided in HG (head groups) and S (solvent).

^d Structures at different time points of the AnchorSim simulation were covalently linked to the catalytic domain.

Table S5. RMSD differences of the replica simulations to assess structural convergence.

<i>Compared Simulations</i> ^a	<i>Backbone RMSD (Å)</i> ^b	<i>Heavy atom RMSD (Å)</i> ^b
Replica 1 to Replica 2	0.59	0.72
Replica 1 to CYP2D6 WT	0.41	0.51
Replica 2 to CYP2D6 WT	0.52	0.67
Average CYP2D6 WT	0.51	0.63
Replica 3 to Replica 4	0.70	0.89
Replica 3 to CYP2D6*2	1.26	1.43
Replica 4 to CYP2D6*2	1.20	1.32
Average CYP2D6*2	1.05	1.21
Replica 5 to Replica 6	0.92	1.00
Replica 5 to CYP2D6*4	1.97	2.02
Replica 6 to CYP2D6*4	1.55	1.64
Average CYP2D6*4	1.48	1.55
Replica 7 to Replica 8	0.92	1.03
Replica 7 to CYP2D6*10	0.64	0.76
Replica 8 to CYP2D6*10	0.74	0.89
Average CYP2D6*10	0.76	0.89
Replica 9 to Replica 10	0.69	0.83
Replica 9 to CYP2D6*17	0.88	1.02
Replica 10 to CYP2D6*17	0.80	0.91
Average CYP2D6*17	0.79	0.92
Replica 11 to Replica 12	0.63	0.84
Replica 11 to CYP2D6*53	0.77	0.89
Replica 12 to CYP2D6*53	0.62	0.70
Average CYP2D6*53	0.67	0.81

To assess the structural convergence of the simulations we compared average structures from the 100 ns trajectories and compared their RMSD on two different atomic levels.

^a The production phase was used in the assessment for the simulations that were no replicas.

^b Hydrogen atoms were excluded for this calculation.

Table S6. Tunnel opening frequency calculated for production simulations and replicas.

<i>Compared simulations</i>	<i>Tunnel opening frequency (%)</i>					
	<i>2b</i>	<i>2c</i>	<i>2e</i>	<i>4</i>	<i>S</i>	<i>W</i>
CYP2D6 WT	73.0	20.2	13.8	9.0	5.4	0.0
Replica 1	97.6	48.6	9.6	19.2	9.2	0.8
Replica 2	82.2	73.4	16.4	16.6	4.4	0.0
Average Replicas	89.9	61.0	13.0	17.9	6.8	0.4
CYP2D6*2	94.2	78.0	23.4	35.2	0.0	0.0
Replica 3	65.2	73.6	21.8	46.2	3.8	0.2
Replica 4	92.4	80.8	6.8	n/a	14.6	0.6
Average Replicas	78.8	77.2	14.3	n/a	9.2	0.4
CYP2D6*4	45.6	37.4	29.6	37.2	3.2	0.0
Replica 5	79.0	53.0	41.9	58.4	4.6	0.2
Replica 6	63.0	10.4	21.6	34.2	3.0	0.0
Average Replicas	71.0	31.7	31.8	46.3	3.8	0.1
CYP2D6*10	98.8	0.0	4.0	39.0	0.0	0.0
Replica 7	97.6	0.0	0.6	n/a	0.0	0.0
Replica 8	n/a	n/a	n/a	26.2	0.6	0.0
Average Replicas	n/a	n/a	n/a	n/a	0.3	0.0
CYP2D6*17	100.0	1.6	1.2	54.2	0.4	0.2
Replica 9	100.0	0.8	0.2	41.6	0.2	0.2
Replica 10	100.0	1.6	1.0	41.8	0.2	0.4
Average Replicas	100.0	1.2	0.6	41.7	0.2	0.3
CYP2D6*53	98.6	1.2	15.0	65.2	11.8	0.0
Replica 11	98.8	0.0	27.0	83.3	2.8	0.0
Replica 12	98.2	0.6	20.0	25.0	0.8	0.0
Average Replicas	98.5	0.3	23.5	54.2	1.8	0.0

The tunnel opening frequencies in percent for six highest ranked tunnels. A tunnel was considered to be open if its bottleneck radius was above 1.2 Å.

n/a: The clustering performed by CAVER did not allow the correct calculation of parameters of the tunnel in the respective simulation.

Table S7. Average hydrophobic index of tunnel lining residues.

<i>Tunnel</i>	<i>Average hydrophobic index</i>
2a	0.51*
2b	0.95±0.32
2c	1.32±0.29
2d	1.39±0.20
2e	0.55±0.37
3	0.71*
4	1.00±0.28
5	0.71±0.35
6	0.61*
Water	0.88±0.19
Solvent	0.88±0.28

We calculated the average hydrophobic index of the tunnel lining residues if a tunnel cluster was present in at least 50 frames of the related simulation. The values were averaged over the variants and the wild-type.

* no standard deviation given since tunnel fitting the calculation criteria only occurred in one variant

Table S8. Average bottleneck radii of the variants for production simulations and replicas.

<i>Compared simulations</i>	<i>Average bottleneck radii of tunnels (Å)</i>			
	<i>2b</i>	<i>2c</i>	<i>2e</i>	<i>4</i>
CYP2D6 WT	1.4 ± 0.2	1.5 ± 0.2	1.1 ± 0.1	1.1 ± 0.1
Replica 1	1.7 ± 0.2	1.3 ± 0.3	1.1 ± 0.1	1.1 ± 0.2
Replica 2	1.4 ± 0.2	1.4 ± 0.3	1.1 ± 0.1	1.1 ± 0.2
CYP2D6*2	1.8 ± 0.2	1.4 ± 0.2	1.2 ± 0.2	1.4 ± 0.3
Replica 3	1.8 ± 0.3	1.4 ± 0.2	1.2 ± 0.2	1.7 ± 0.4
Replica 4	1.8 ± 0.3	1.4 ± 0.2	1.1 ± 0.2	n/a
CYP2D6*4	1.2 ± 0.2	1.2 ± 0.2	1.1 ± 0.1	1.2 ± 0.2
Replica 5	1.4 ± 0.2	1.3 ± 0.2	1.2 ± 0.2	1.3 ± 0.2
Replica 6	1.3 ± 0.2	1.1 ± 0.1	1.1 ± 0.2	1.1 ± 0.2
CYP2D6*10	1.9 ± 0.2	0.0 ± 0.0	1.1 ± 0.2	1.3 ± 0.3
Replica 7	1.8 ± 0.2	0.0 ± 0.0	1.0 ± 0.1	n/a
Replica 8	n/a	n/a	n/a	1.3 ± 0.2
CYP2D6*17	1.8 ± 0.2	1.0 ± 0.1	1.0 ± 0.1	1.4 ± 0.2
Replica 9	1.9 ± 0.2	1.0 ± 0.1	1.0 ± 0.1	1.4 ± 0.2
Replica 10	1.8 ± 0.2	1.0 ± 0.1	1.0 ± 0.1	1.4 ± 0.3
CYP2D6*53	2.6 ± 0.3	1.1 ± 0.3	1.1 ± 0.2	1.3 ± 0.2
Replica 11	2.7 ± 0.2	1.0 ± 0.1	1.2 ± 0.2	1.4 ± 0.2
Replica 12	2.5 ± 0.2	1.0 ± 0.2	1.1 ± 0.2	1.1 ± 0.2

The average bottleneck radii of enzyme tunnels in the wild-type and the five variants. We chose to compare the four highest ranked tunnels. The average values are given with standard deviation.

n/a: The clustering performed by CAVER did not allow the correct calculation of parameters of the tunnel in the respective simulation.

Table S9. Major bottleneck residues calculated for production phases and replicas.

<i>Compared runs</i>	<i>Number of tunnels per residue</i>					
	<i>120</i>	<i>305</i>	<i>308</i>	<i>370</i>	<i>374</i>	<i>483</i>
Production phases	5	3	3	4	5	3
Replica 1	5	2	3	2	4	2
Replica 2	4	3	1	4	5	4

Bottleneck residues involved in the regulation of multiple tunnels are shown together with the number of tunnels they are involved in. Note that F120 was involved as a bottleneck residue in tunnel 2b, 2c, 2e, 4, and W.

Table S10. Average number of hydrogen bonds for production phases and replicas.

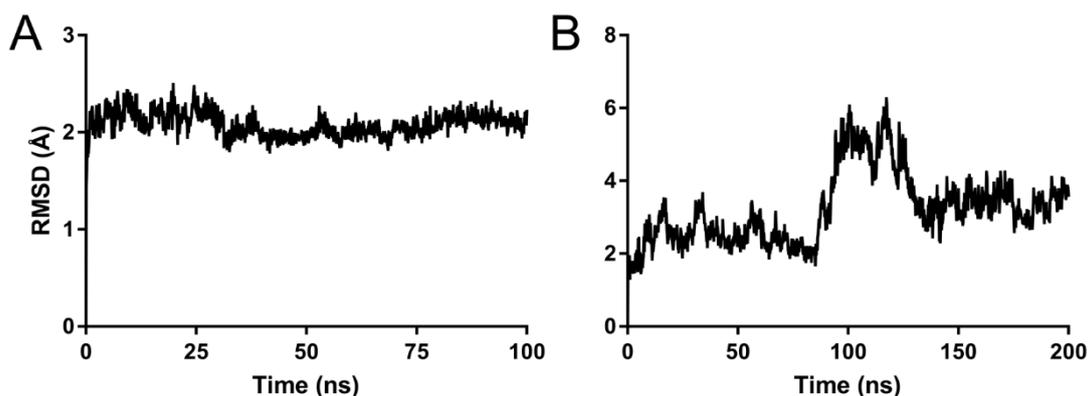
<i>Compared simulations</i>	<i>Average number of hydrogen bonds per residue</i>							
	<i>34</i>	<i>91</i>	<i>94</i>	<i>107</i>	<i>120</i>	<i>122</i>	<i>296</i>	<i>486</i>
CYP2D6 WT	0.6 ± 0.5	1.2 ± 0.8	3.4 ± 0.9	2.1 ± 0.5	2.5 ± 0.6	1.3 ± 0.5	2.5 ± 0.7	1.9 ± 0.5
Replica 1	0.7 ± 0.5	1.2 ± 0.8	3.6 ± 0.9	2.9 ± 0.8	2.0 ± 0.1	1.6 ± 0.6	2.8 ± 0.7	2.8 ± 0.7
Replica 2	0.7 ± 0.5	1.7 ± 0.5	4.1 ± 0.9	2.1 ± 0.3	2.0 ± 0.1	1.6 ± 0.6	2.9 ± 0.7	2.1 ± 0.4
CYP2D6*2	-	-	-	-	-	-	0.7 ± 0.9	2.1 ± 0.4
Replica 3	-	-	-	-	-	-	1.6 ± 1.0	2.5 ± 0.5
Replica 4	-	-	-	-	-	-	1.3 ± 1.0	2.4 ± 0.5
CYP2D6*4	0.4 ± 0.5	0.9 ± 0.5	3.2 ± 0.6	-	-	-	-	1.8 ± 0.5
Replica 5	0.0 ± 0.2	1.3 ± 0.5	3.6 ± 0.7	-	-	-	-	2.4 ± 0.5
Replica 6	0.4 ± 0.6	1.3 ± 0.5	3.5 ± 0.7	-	-	-	-	2.3 ± 0.5
CYP2D6*10	0.9 ± 0.4	-	-	-	-	-	-	2.1 ± 0.5
Replica 7	0.9 ± 0.2	-	-	-	-	-	-	2.3 ± 0.6
Replica 8	0.9 ± 0.3	-	-	-	-	-	-	2.0 ± 0.2
CYP2D6*17	-	-	-	1.1 ± 0.7	-	-	1.1 ± 0.5	2.6 ± 0.8
Replica 9	-	-	-	1.5 ± 0.7	-	-	1.4 ± 1.0	3.1 ± 0.3
Replica 10	-	-	-	1.6 ± 0.8	-	-	1.7 ± 1.0	3.0 ± 0.3
CYP2D6*53	-	-	-	-	2.4 ± 1.0	2.3 ± 0.7	-	-
Replica 11	-	-	-	-	2.0 ± 0.6	2.5 ± 0.8	-	-
Replica 12	-	-	-	-	2.0 ± 0.4	2.2 ± 0.8	-	-

The data was calculated for all mutated amino acids in the wild-type as well as for the mutated amino acids in the variants. The values are given with standard deviation.

Table S11. Active site volume calculated with POVME for production simulations and replicas.

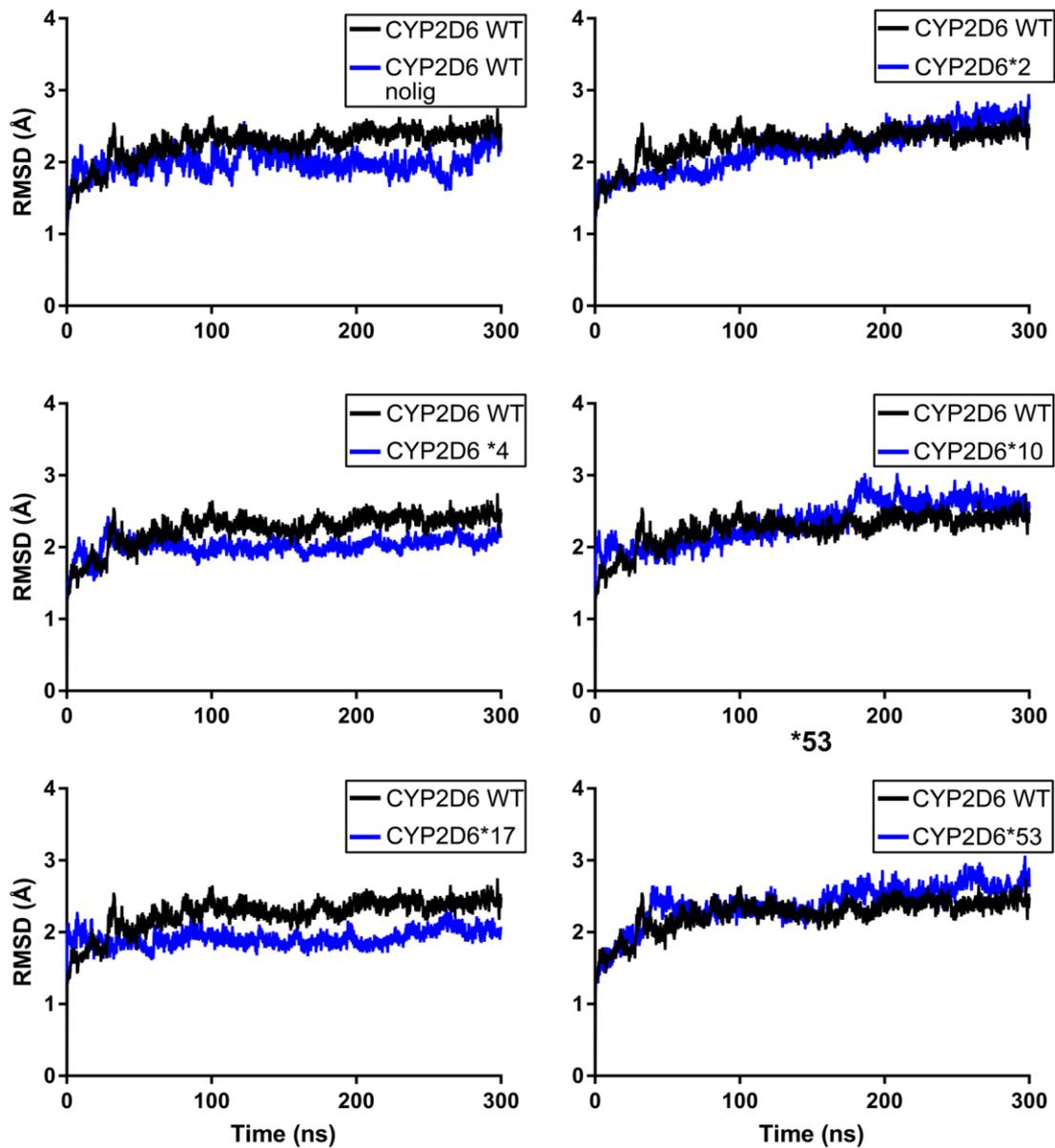
<i>Compared simulations</i>	<i>Volume (\AA^3)</i>	<i>Average Replicas</i>
CYP2D6 WT	674.9 \pm 83.6	
Replica 1	634.0 \pm 64.9	589.7 \pm 62.6
Replica 2	545.4 \pm 52.9	
CYP2D6*2	739.6 \pm 76.4	
Replica 3	776.4 \pm 64.5	898.0 \pm 172.0
Replica 4	1019.6 \pm 95.6	
CYP2D6*4	398.6 \pm 46.1	
Replica 5	514.0 \pm 77.1	455.2 \pm 83.2
Replica 6	396.3 \pm 40.7	
CYP2D6*10	883.0 \pm 80.7	
Replica 7	947.3 \pm 80.5	978.0 \pm 43.3
Replica 8	1008.6 \pm 66.8	
CYP2D6*17	687.9 \pm 66.1	
Replica 9	557.8 \pm 55.1	628.1 \pm 99.4
Replica 10	698.4 \pm 60.3	
CYP2D6*53	1092.3 \pm 108.3	
Replica 11	1074.9 \pm 88.2	1086.3 \pm 16.1
Replica 12	1097.7 \pm 66.7	

The volumes of the active site cavities in CYP2D6 WT and its related variants. The volumes were estimated with POVME (v2.0) based on the last 100 frames of the production phase. The values are given with standard deviation and average values for the replicas have been calculated.

Figure S1. Backbone RMSD plots of the simulations MemAs and AnchorSim.

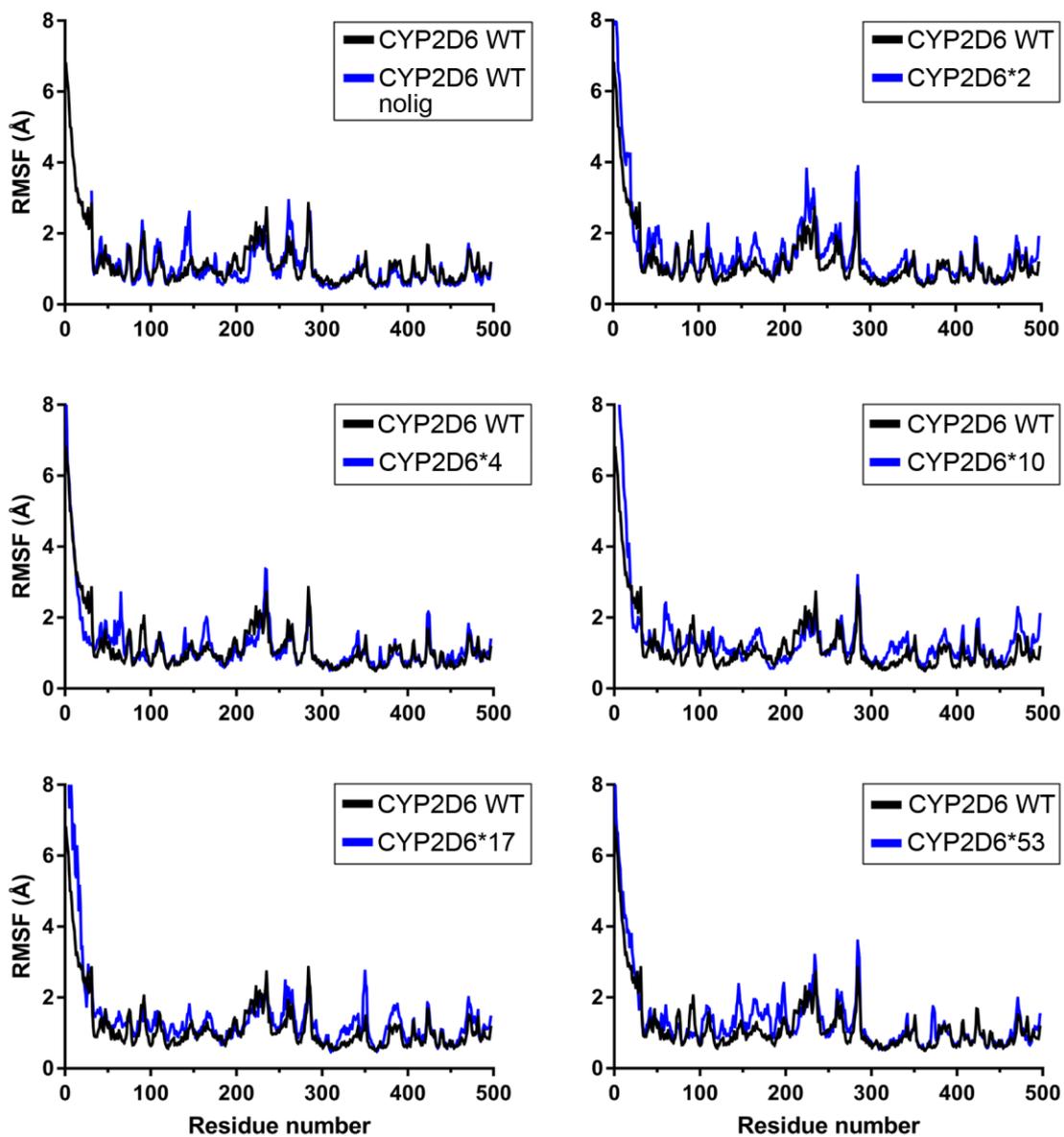
The backbone RMSD values for (A) the MemAs simulation and (B) the AnchorSim simulation. The values are shown for the whole duration of the simulation.

Figure S2. Backbone RMSD profiles of the membrane-bound simulations and the WT nolog simulations of CYP2D6.



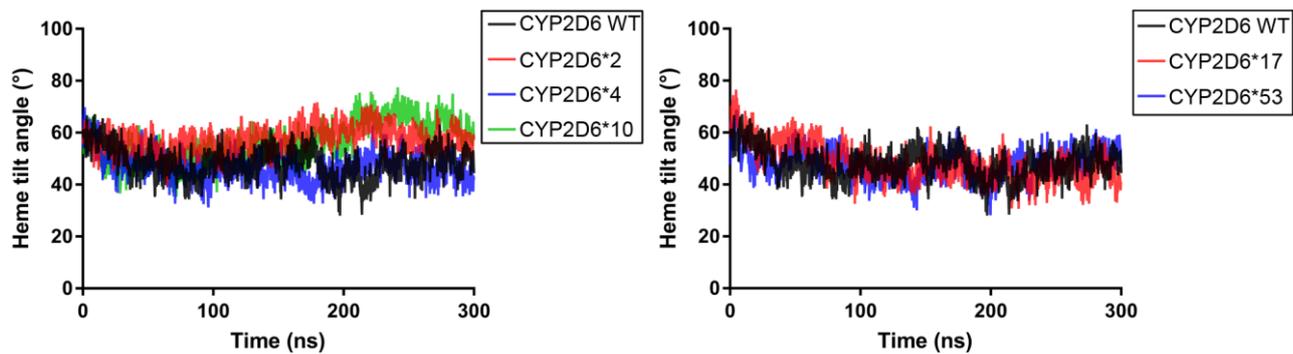
The backbone RMSD values of membrane-bound simulations as well as the WT nolog simulation. The values are shown for the whole 300 ns simulation time.

Figure S3. Backbone RMSF profiles of membrane-bound simulations and the WT nolog simulation.



The backbone RMSF values of membrane-bound simulations, as well as the WT nolog simulation. The values are shown for the whole 300 ns simulation time.

Figure S4. Heme tilt angle of membrane-bound simulations of CYP2D6.



The heme tilt angle of the membrane-bound simulations. It is defined as the angle between the heme plane and the membrane normale corresponding to the z-axis. The values are shown for the whole simulation time of 300 ns.

Figure S5. Burying depth of membrane-bound systems of CYP2D6.

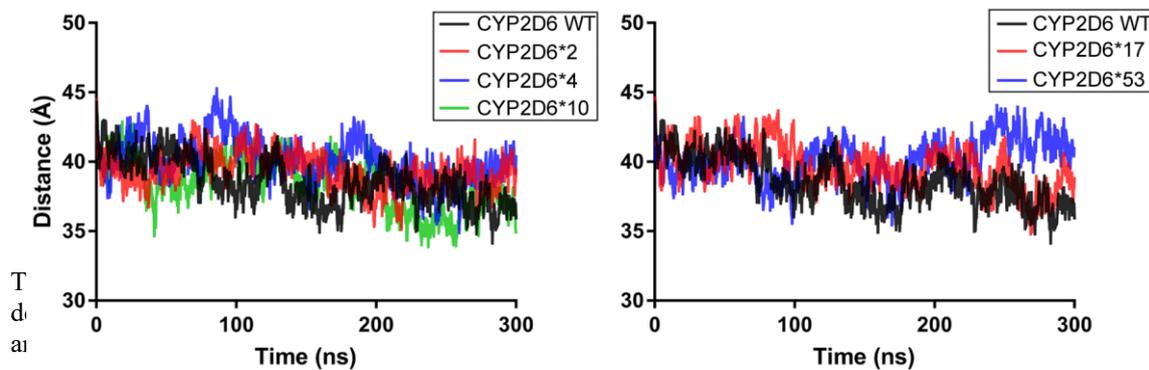
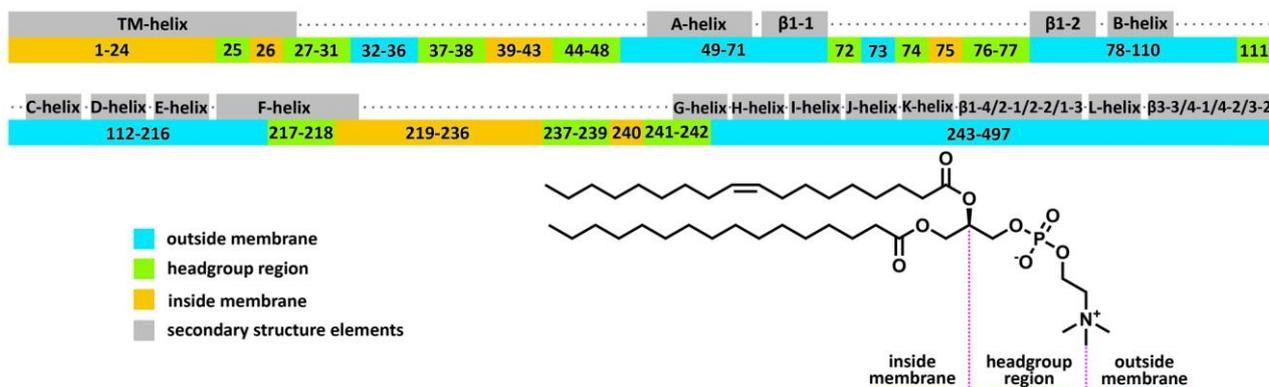
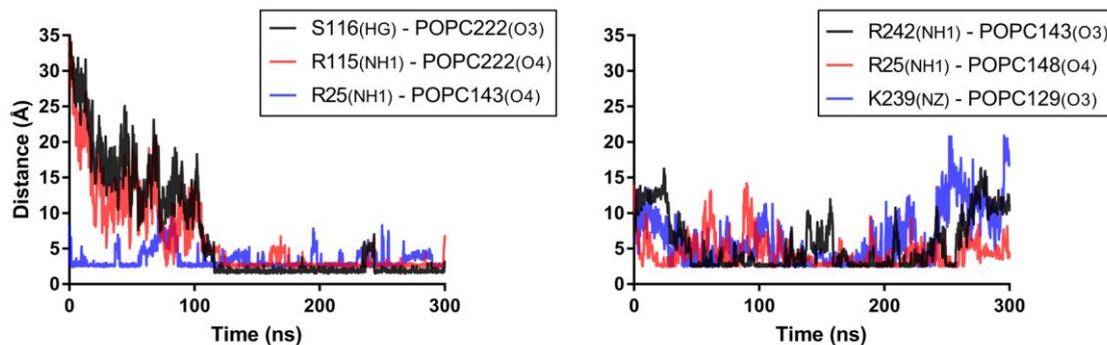


Figure S6. Residues in contact with the membrane, the head group region, or the cytosol in the wild-type model of CYP2D6.



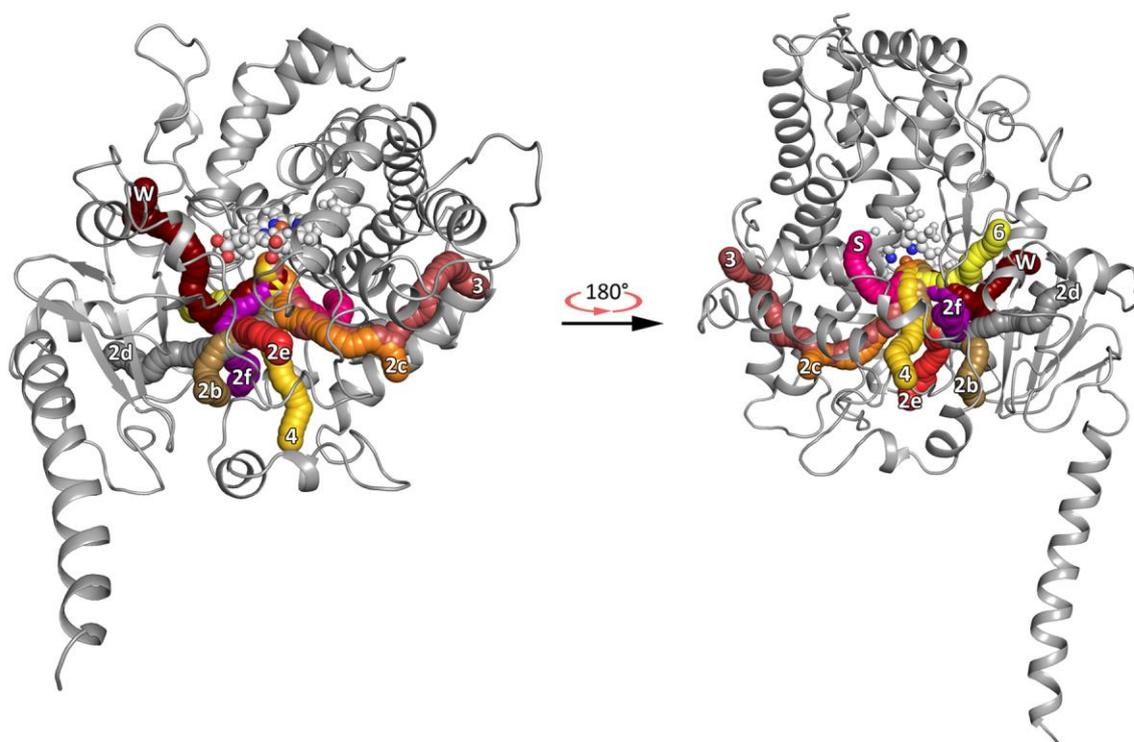
Residues of the CYP2D6 WT model that are in contact with the membrane, the head group region, or the cytosol. We defined the head group region to be between mass centers of the C2-atoms and the N-atoms of the POPC molecules. The localization of the residue was determined considering the position of its α -carbon atom.

Figure S7. Distance between residues and POPC molecules and selected protein residues.



The distance between selected atoms of protein residues and membrane molecules plotted against the simulation time. Distances in the range of 2 Å indicate a possible interaction between the residues.

Figure S10. Spatial distribution of enzyme tunnels in CYP2D6 WT.



Identified enzyme tunnels in all simulations. We considered a tunnel to be present if it occurred in at least five frames during the production phase. We observed 13 of 14 known tunnels in the WT inhibitor simulation.

Figure S11. Identified tunnels in all production simulations.

Variants	Tunnels													
	1	2ac	2b	2c	2d	2e	2f	3	4	5	6	S	W	
CYP2D6 WT inhibitor	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	
CYP2D6 WT nolig	x	x	✓	✓	✓	✓	x	✓	✓	x	✓	✓	✓	
CYP2D6 WT	x	x	✓	✓	✓	✓	✓	✓	✓	x	✓	✓	✓	
CYP2D6*2	x	x	✓	✓	x	✓	x	✓	✓	✓	✓	✓	✓	
CYP2D6*4	x	x	✓	✓	✓	✓	x	✓	✓	✓	✓	✓	✓	
CYP2D6*10	x	x	✓	x	✓	✓	x	x	✓	✓	✓	✓	✓	
CYP2D6*17	x	x	✓	✓	✓	✓	x	x	✓	x	✓	✓	✓	
CYP2D6*53	x	x	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	

Identified enzyme tunnels in all production simulations. We considered a tunnel to be present if it occurred in at least five frames during the production phase. We observed 13 of 14 known tunnels in the WT inhibitor simulation.