Supplementary Material

Manuscript Title:

Computational modelling of the binding of arachidonic acid to the human monooxygenase CYP2J2

Journal Name

Journal of Molecular Modeling

Authors

G. Proietti, K.K. Abelak, D. Bishop-Bailey, A. Macchiarulo & I. Nobeli

Corresponding author address and email

Dr Irene Nobeli

Institute of Structural and Molecular Biology, Department of Biological Sciences, Birkbeck, Malet Street, London WC1E 7HX, UK.

i.nobeli@bbk.ac.uk

Supplementary Table 1.

A summary of published computational models of CYP2J2.

Authors; Journal - Year	Template(s) (+ligand); PDB (resolution/Å)	Methods	Reason for building the computational model	ASV / Å^3	Comments
Lafite et al; Biochemistry - 2007	CYP2A6 (+methoxsalen); 1z11 (2.05) CYP2B4 (+bifonazole); 2bdm (2.30) CYP2C5 (+diclofenac); 1nr6 (2.10) CYP2C8 (no ligand); 1pq2 (2.7) CYP2D6 (no ligand); 2f9q (3.00)	Model built with Swiss-Model v3.5; Heme added from CYP2C8 Stability validation carried out at 300K; All other MD at 50K or 100K Docking of terfenadone derivatives was carried out with soft-restrained MD in vacuo MD and minimisations were done in vacuo	Explanation of the unusual regioselectivity of the hydroxylation of terfenadone derivatives by CYP2J2	945	Arachidonic acid was not docked in this study
Li et al; Proteins - 2008	CYP2C9 (+warfarin); 1og5 (2.55)	Model built with Modeller 8v2 and relaxed with MD in Amber using ff03 force-field over 1ns at 300K Flexible ligand docking for 4 ligands with Affinity (InsightII) using CVFF force-field Residues within 8Å of ligands were allowed to move freely during docking Monte-Carlo search followed by simulated annealing over 5ps and temperature rescale from 500K to 300K	Examination of the modes of binding of four CYP2J2 inhibitors	NC	Arachidonic acid was not docked in this study RMSD did not plateau and thus model is unlikely to be fully equilibrated
Lee et al; Drug Metab Dispos - 2010	CYP2B4 (no ligand); 1po5 (1.60) CYP2C8 (no ligand); 1pq2 (2.7) CYP2A6 (+coumarin); 1z10 (1.90)	Model built with Modeller 8v1	Comparison of binding sites between CYP2J2 and CYP3A4	1420	No docking or MD simulations were carried out
Cong et al; J Chem Inf Model - 2013	CYP2R1 (+vitamin D3); 3c6g (2.80)	Model built with 'segment matching' method 1ns MD to extract snapshots for docking Flexible docking of ARA with Monte-Carlo simulated annealing in MOE 2008 10ns MD simulations of docked ARA with wild-type and mutant enzymes; All MD at 310K in Amber 10 using Amber force-field	Explanation of polymorphism-induced changes to ARA binding and metabolism T143A and N404Y mutant models were built to allow comparison to wild-type	320	Arachidonic acid modelled in its protonated form The carboxyl group of ARA is H-bonded to Leu378 and Gly486
Xia et al; Curr Drug Metab - 2014	CYP2A6 (no ligand); 2pg6 (2.53) CYP2C8 (+troglitazone); 2vn0 (2.70) CYP2E1 (+indazole); 3e6i (2.20) CYP17A1 (+abiraterone); 3ruk (2.60) CYP2R1 (+vitamin D3); 3c6g (2.80)	Model built with 'segment matching' method Short MD to extract snapshots of the protein for docking Flexible docking of ARA with MOE 10ns MD simulations of docked ARA with wild-type and mutant enzymes; All MD at 310K in Amber 11 using Amber force-field	Explanation of mutation-induced changes to ARA binding T143A, R158C, I192N and N404Y mutant models were built to allow comparison to wild-type	330	Arachidonic acid modelled in its protonated form The carboxyl group of ARA is H-bonded to Leu378 and Gly486

ASV – Active Site Volume; NC – not calculated

Supplementary Table 2.

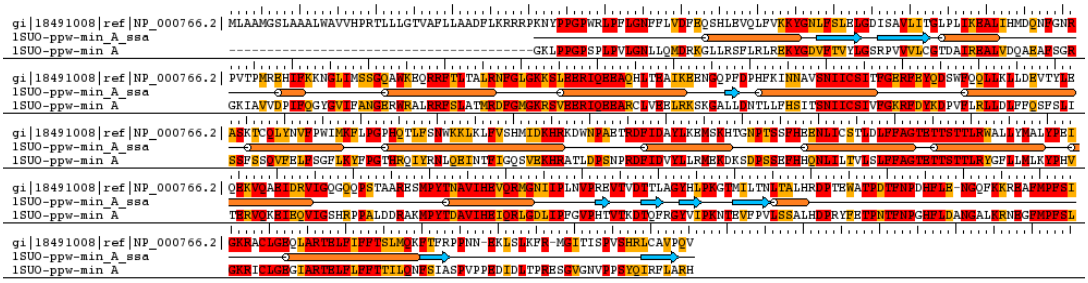
A summary of computational models of CYP2J2 available in online resources and/or automated modelling servers.

Server	Residue range	Template	Seq id/%	Res/ Å	Creation date	QMEAN Z-Score	Verify3d		ProCHECK		ERRAT	Note
							%	Min/Max	Favoured + additionally allowed/%	Disallowed/% (id)	Score	
SWISSMODEL	62-498	3czhB	44	2.30	30/05/2015	-0.03	91.76	0.09/0.65	96.7	0.8 (148V)	78.1	
SWISSMODEL	71-474	3idbA	46	2.00	30/05/2015	-0.11	89.85	-0.05/0.64	99.7	None	85.4	
MODBASE	43-500	3czhA	44	2.30	23/10/2013	-0.07	83.41	-0.22/0.74	96.7	0.7 (52F, 442S, 492V)	69.0	
M4T	3-464	1po5A, 1r9oA			28/06/2015	-1.32	75.16	-0.20/0.69	93.7	0.4 (442S, 501Q)	67.7	
IntFOLD v2	39-501	2pg6A			28/06/2015	-0.45	83.47	-0.49(0.00)/0.73	96.6	0.2 (442S)	65.9	1-38 present with no secondary structure
I-TASSER	1-502	2pg6A, 3czhA, 1z11A, 3swzA, 2nnjA	38, 41, 38, 26, 38		28/06/2015	-1.48	79.48	-0.18/0.73	94.6	1.4 (12L, 23L, 26V, 27A, 48W, 295S)	87.0	Also modelled transmembrane helices (1-40)
PHYRE2	1-501	1nr6A, 2nnjA, 3czhA, 3ebsA, 3e4eA, 1po5A	42, 42, 44, 42, 40, 43		28/06/2015	-1.13	78.88	-0.23/0.74	94.2	1.2 (3A, 12L, 32A, 39R, 116M, 442S)	72.3	Also modelled transmembrane helices (1-40)
RaptorX	36-502	3czhA			28/06/2015	-0.03	88.47	-0.09/0.73	96.2	0.2 (442S)	66.2	
SWISSMODEL	42-501	3tbgA	43	2.10	30/05/2015	-0.34	89.78	-0.22/0.74	98.5	1.0 (224A, 273E, 480S, 487I)	77.7	

Seq – Sequence; id - identity; Res – Resolution;

Supplementary Fig. 1

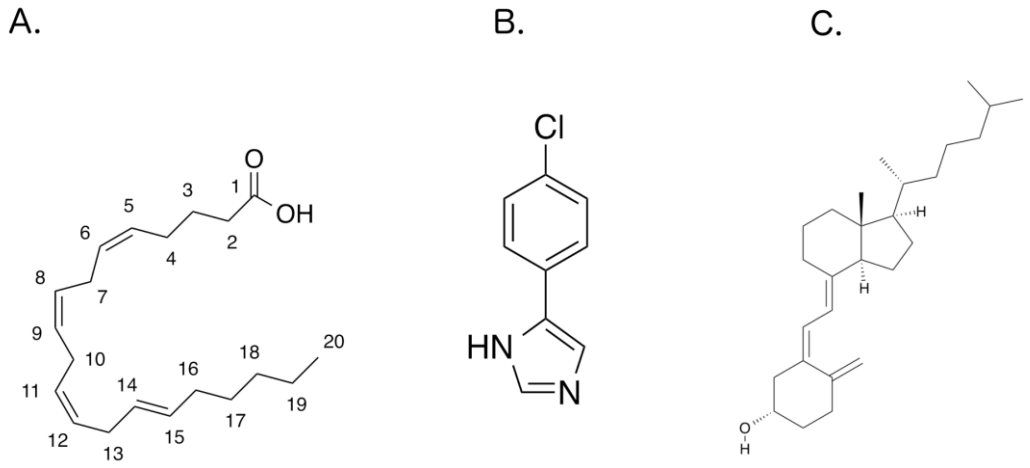
Sequence alignment of human CYP2J2 to the template used for homology modelling (rabbit CYP2B4)



Alignment of human CYP2J2 (Uniprot P51589; top row) to rabbit CYP2B4 (PDB ID: 1suo; bottom row) produced by the CLUSTALW implementation within Schrödinger's homology modelling suite[1]. The middle row represents the secondary structure assignment for all template residues, following protein preparation of the structure using Schrödinger's Protein Preparation Wizard. Colouring of the residues is based on the "Coloring by Homology" scheme, where red indicates strict conservation and orange indicates a conservative substitution. There are three one-residue gaps in the alignment and none of them are part of a secondary structure element.

Supplementary Fig. 2

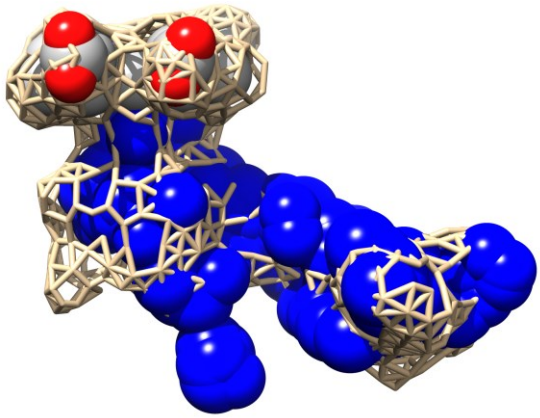
2D diagrams of ligands relevant to this study



The structures of A) arachidonic acid (AA); B) 4-(4-chlorophenyl) imidazole, the ligand bound to rabbit CYP2B4 in PDB ID 1suo, and C) vitamin D3, the ligand bound to human CYP2R1 in PDB ID 3c6g. All structures were drawn using ChemBioDraw[2].

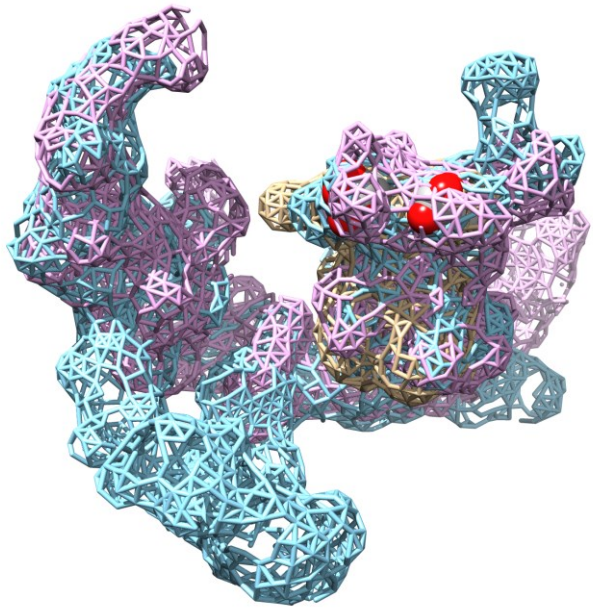
Supplementary Fig. 3

Cavity detection in models of CYP2J2.



A.
The active site cavity of the human CYP2J2 homology model produced in this study (prior to docking or MD), as perceived by PDBsum (cavity depicted in points connected by golden rods) and *fpocket* (blue spheres represent the program's "alpha spheres" output). The corresponding volumes calculated by the two programs are: 1977Å^3 and 1339Å^3, both a lot larger than the 234Å^3 calculated by SiteMap (as described in the main manuscript). The PDBsum calculation is the only one to include the cavity volume that is filled by heme.

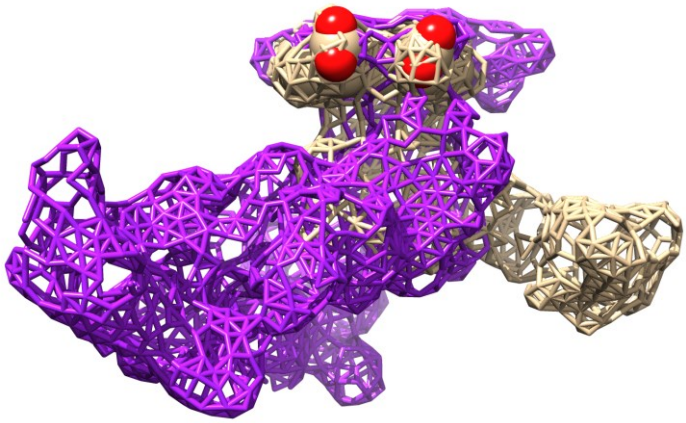
B.



The active site cavities as perceived by PDBsum for three snapshots from the MD simulation: at 0 ns (golden rods), 25 ns (cyan rods) and 50 ns (pink rods). The corresponding volumes are: 1948Å^3, 8119Å^3 and 7602Å^3. Only one cavity is shown per snapshot, together with the heme molecule (shown in a spheres representation). It

is clear that during the simulation a narrow channel connecting neighbouring cavities opens up, allowing the creation of much more extended cavities at 25 and 50ns. In such cases, relying on just volume estimates calculated by a program is misleading, as opening of channels leads to the incorporation of multiple cavities that are far from the active site.

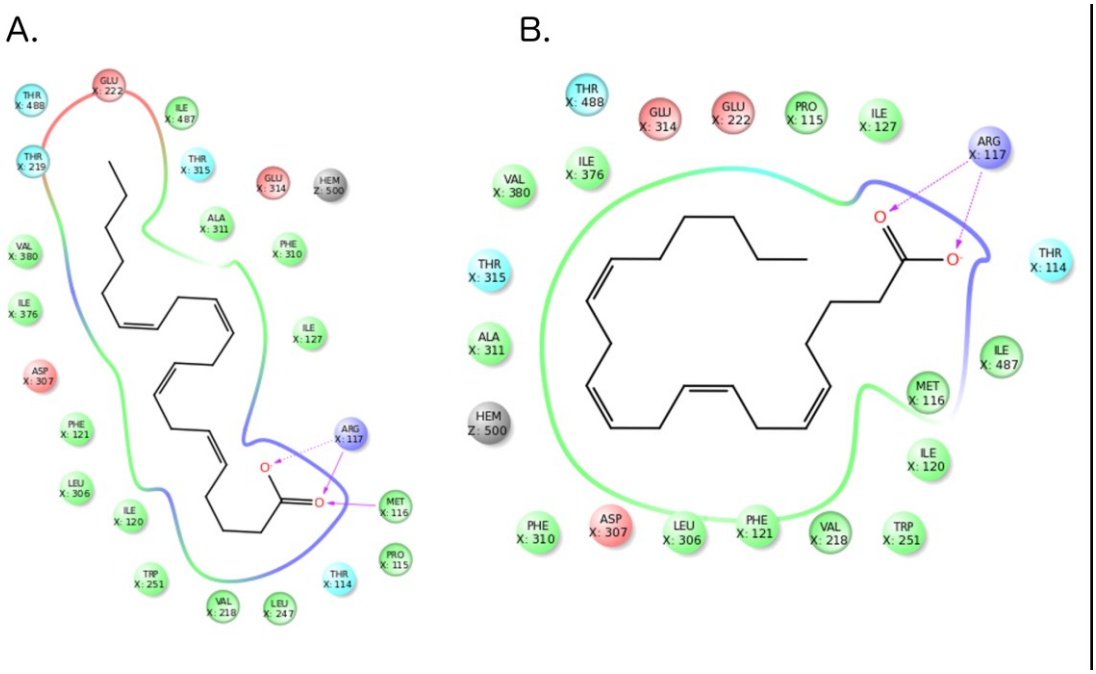
C.



Cavities perceived by PDBsum in the original hCYP2J2 model constructed in this study (golden rods) and a model constructed by SwissModel using as template the PDB structure 3czh (purple rods). The corresponding volumes are: 1977Å^3 and 4944Å^3. The cavities of the two models appear to be very different despite the fact that the two model structures differ by only 2.3Å RMSD (over all 437 Cα pairs; calculated with Chimera's MatchMaker program). Comparison of this figure to (B) shows that it is possible to observe comparably large differences in the perceived cavities of snapshots from the same model undergoing molecular dynamics simulation.

Supplementary Fig. 4

Protein-ligand interaction diagrams for induced fit docking (IFD) poses



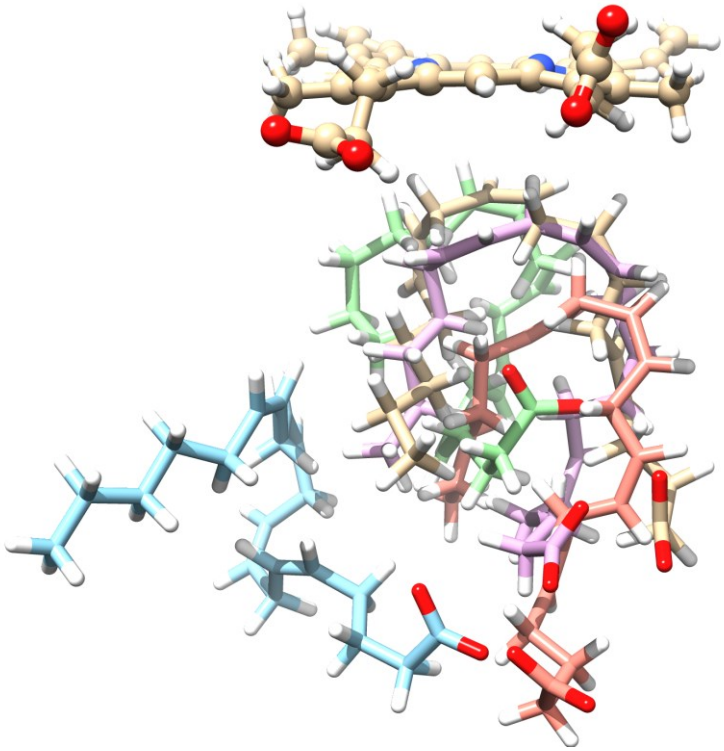
Protein-ligand interaction diagrams (created with the Schrödinger software[1]) corresponding to two poses of AA docked with IFD to our model of CYP2J2 are shown. In these diagrams protein residues are shown as circles coloured by the type of the residue (green=hydrophobic, cyan=polar, red=negatively charged, blue=positively charged). Hydrogen bonds are shown as arrows linking the protein residue to a ligand atom. Dashed and solid arrows correspond to hydrogen bonds involving side-chain and main-chain (backbone) hydrogen bonds respectively.

A) The top pose from IFD shows both Arg117 and Met116 hydrogen bonding to AA. In this pose both the backbone and side-chain of Arg117 are involved in hydrogen bonds to the AA carboxylate group.

B) Only Arg117 is hydrogen-bonded to AA in pose number 8 and both hydrogen bonds are to the side-chain of the arginine.

Supplementary Fig. 5

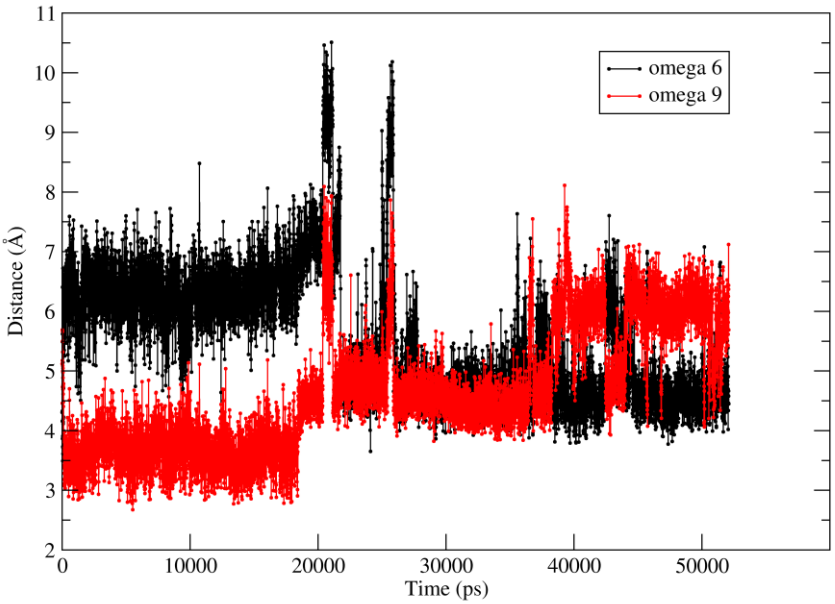
Data from additional MD simulations



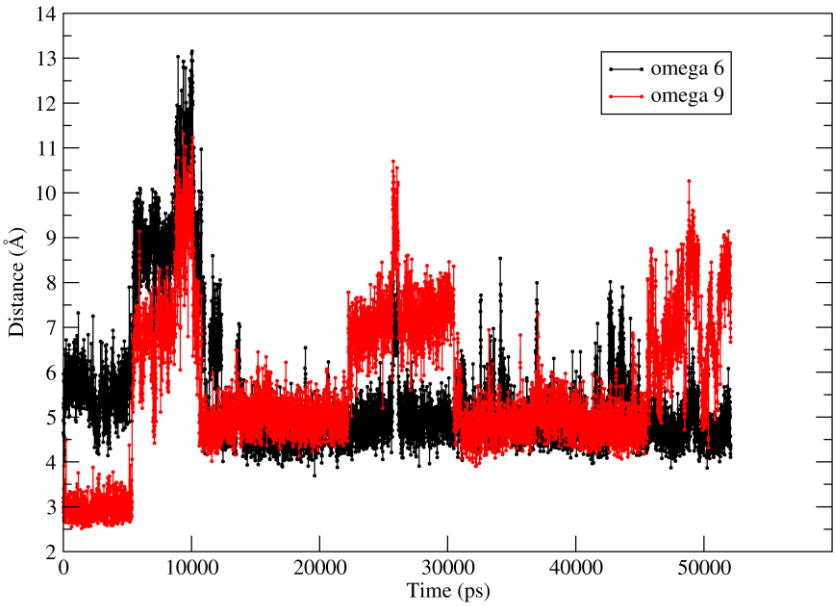
A. Superposition of the last frames from 5 independent MD simulations

The last frames from 5 independent MD simulations of AA binding to human CYP2J2 have been superimposed on one of the five runs (run 2, atoms in gold), using the heme molecule atoms to guide the superposition. RMSD values for the AA atoms in each run relative to run 2 (gold) have been calculated with Chimera's "match" program and are listed here:
Run 1 (cyan): 9.3Å
Run 3 (pink): 3.4Å
Run 7 (green): 5.7Å
Run 8 (salmon): 5.1Å
Clearly runs 1 and 8 represent binding modes that are too far from the heme to be ready for catalysis.

Pose 3



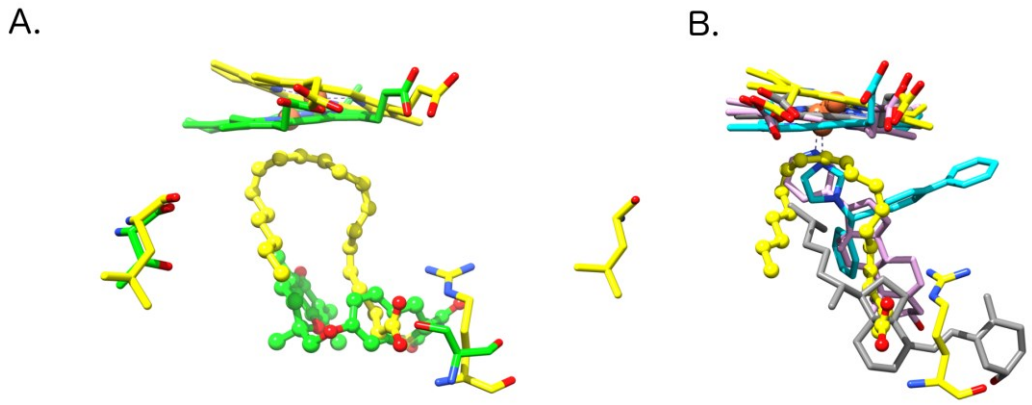
Pose 7



B. Monitoring of the distance between the heme iron and ω9/ω6 bonds of AA
During simulations 3 and 7 there are time slots when the ω9 bond is positioned within 4Å of the heme iron, but in both MD runs the complex eventually moves away from this arrangement.

Supplementary Fig. 6

CYP450 inhibitors from a number of crystal structures are bound in the same region as the one suggested for AA in this study

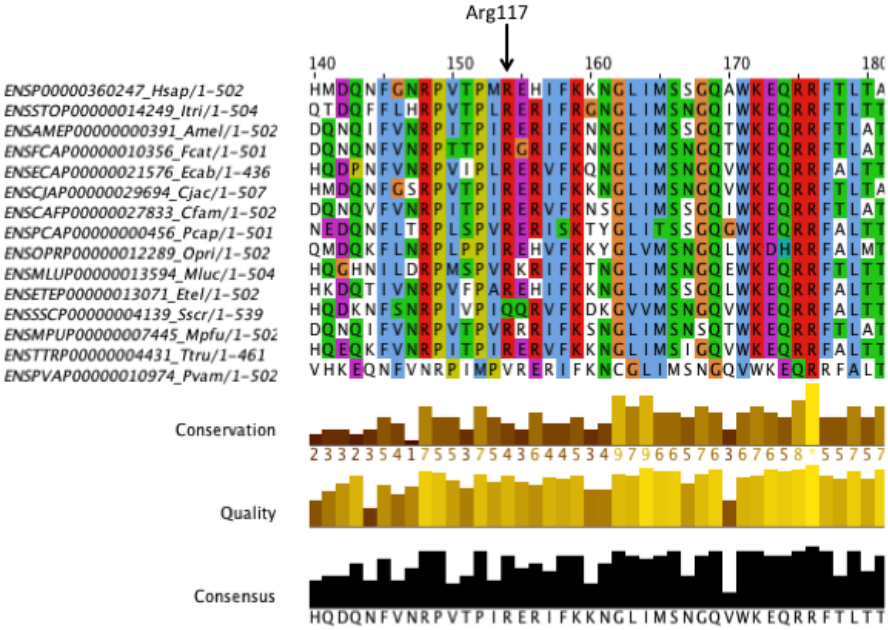


A) Superposition of the CYP2J2-AA complex from the last frame of our MD simulation (yellow) onto human CYP2C8 from the PDB structure 2vn0[3] (green) shows that the inhibitor troglitazone (green balls and sticks) is interacting with Ser103, the residue that aligns with Arg117 in the sequence alignment of the two related enzymes. Thr364 that aligns with Leu378 exhibits no interactions with the inhibitor.

B) Three cytochrome P450 enzymes and their small molecule inhibitors are shown superimposed onto the CYP2J2-AA complex from the last frame of our MD simulation (yellow). For simplicity only the heme molecules are shown at the top of the figure, together with all ligands (all in stick format, except AA, shown as ball and stick) as well as Arg117 and Leu378 from CYP2J2. The complexes shown are: CYP17A1 with bound abiraterone (PDB: 3ruk[4], pink), CYP2B4 with bound bifonazole (PDB: 2bdm[5], cyan) and CYP2R1 with bound vitamin D3 (PDB: 3c6g[6], grey). In all these structures an inhibitor molecule occupies a region much closer to the one occupied by AA in our model, than to the region surrounding Leu378.

Supplementary Fig. 7

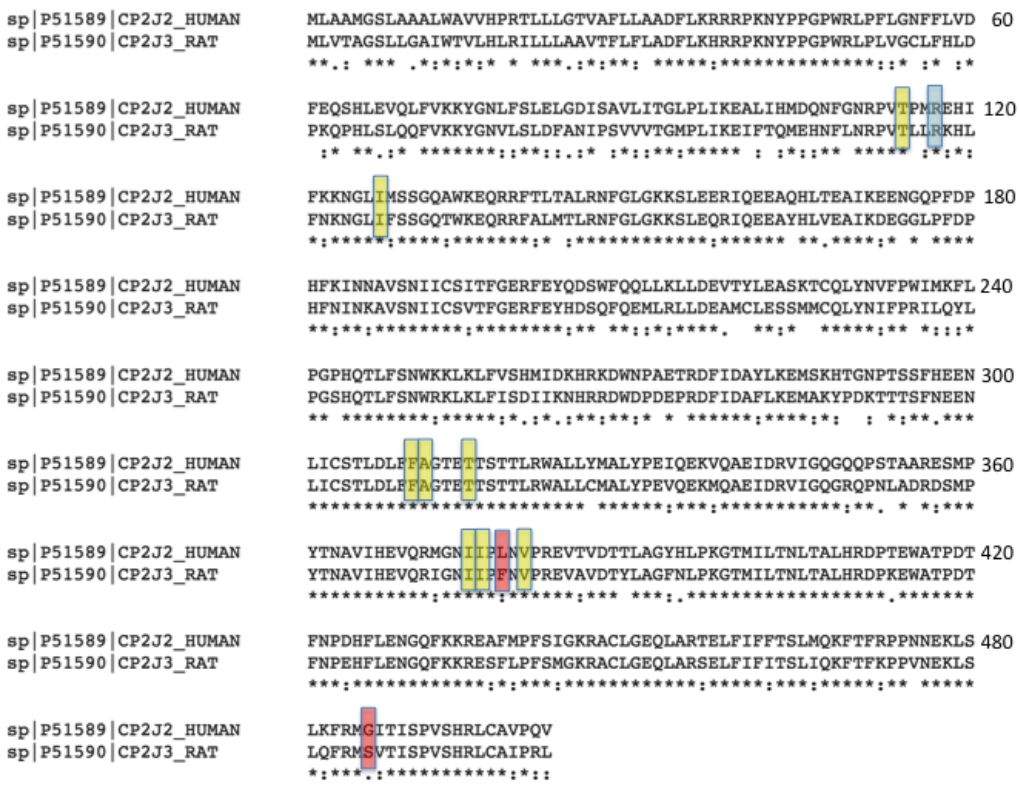
Multiple sequence alignment of human CYP2J2 and a range of closely related proteins from other organisms



From a list of all orthologues of human CYP2J2 reported by Ensembl, we have selected only those proteins with higher sequence identity (at least 70% reported for both target and query genes) and those that were labelled as CYP2J2 within the Ensembl annotation. We then carried out a multiple sequence alignment of these 15 proteins (Ensembl IDs are shown on the left of the figure) using ClustalOmega as implemented within Jalview[7] and displayed the output using JalView. Only the relevant part of the alignment (surrounding Arg117) is shown above for simplicity.

Supplementary Fig. 8

Sequence alignment between human CYP2J2 and its homologue rat CYP2J3 reveals conservation of the residues important for binding of AA in our model



Human CYP2J2 (Uniprot ID: P51589) aligned to rat CYP2J3 (Uniprot ID: P51590) using the ClustalOmega server at the European Bioinformatics Institute. Rat CYP2J3 is known to metabolise arachidonic acid into EETs, just like human CYP2J2. Residues found to interact persistently with arachidonic acid across our MD simulation are conserved in these two sequences and are highlighted with yellow-coloured rectangles, except Arg117, which is highlighted with a cyan-coloured rectangle. Two residues highlighted as important in the study of Cong et al. (Leu378 and Gly486) are not conserved in the rat enzyme.

References for Supporting Information

1. Schrödinger Release 2014-1: Schrödinger, LLC, New York, NY, 2014.
2. ChemBioDraw: version 14, PerkinElmer Informatics, Inc. 2014.
3. Schoch, G. A. *et al.* Determinants of cytochrome P450 2C8 substrate binding: structures of complexes with montelukast, troglitazone, felodipine, and 9-cis-retinoic acid. *J. Biol. Chem.* **283,** 17227–17237 (2008).
4. DeVore, N. M. & Scott, E. E. Structures of cytochrome P450 17A1 with prostate cancer drugs abiraterone and TOK-001. *Nature* **482,** 116–119 (2012).
5. Zhao, Y. *et al.* Structure of microsomal cytochrome P450 2B4 complexed with the antifungal drug bifonazole: insight into P450 conformational plasticity and membrane interaction. *J. Biol. Chem.* **281,** 5973–5981 (2006).
6. Strushkevich, N., Usanov, S. A., Plotnikov, A. N., Jones, G. & Park, H.-W. Structural Analysis of CYP2R1 in Complex with Vitamin D3. *J. Mol. Biol.* **380,** 95–106 (2008).
7. Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M. & Barton, G. J. Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25,** 1189–1191 (2009).