Supporting Information:

In Silico Analysis of the Conservation of Human Toxicity and Endocrine Disruption Targets in Aquatic Species.

Fiona M. McRobb, Virginia Sahagún, Irina Kufareva and Ruben Abagyan*

Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California San Diego, La Jolla, CA

*Corresponding author: rabagyan@ucsd.edu

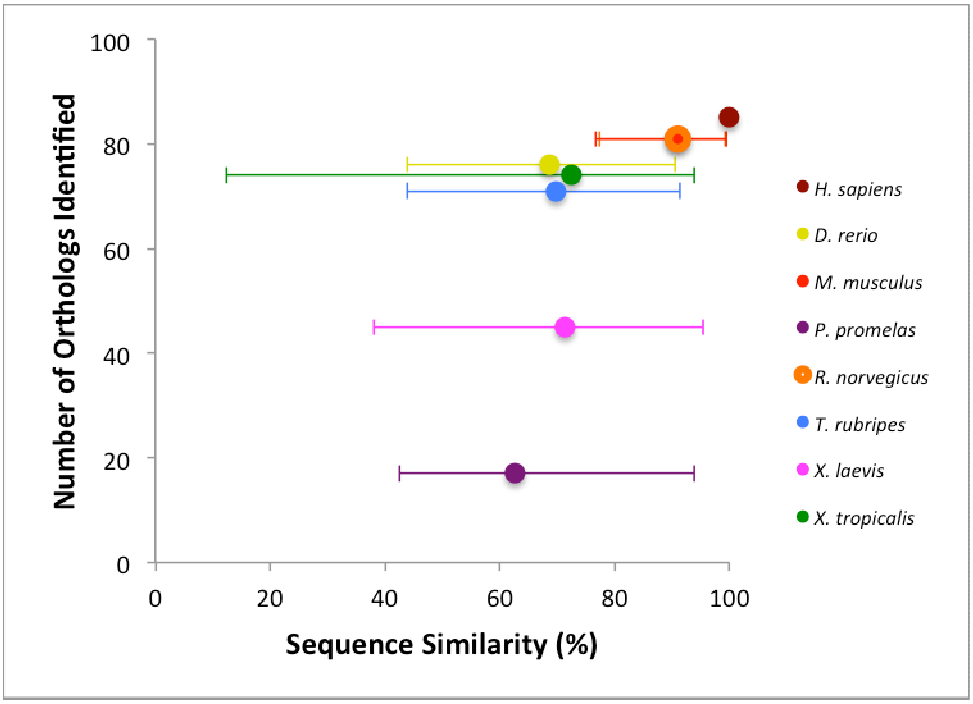

Figure S1. The average sequence similarity of orthologs compared to the corresponding human toxicity target (minimum and maximum similarities are indicated by the error bars).

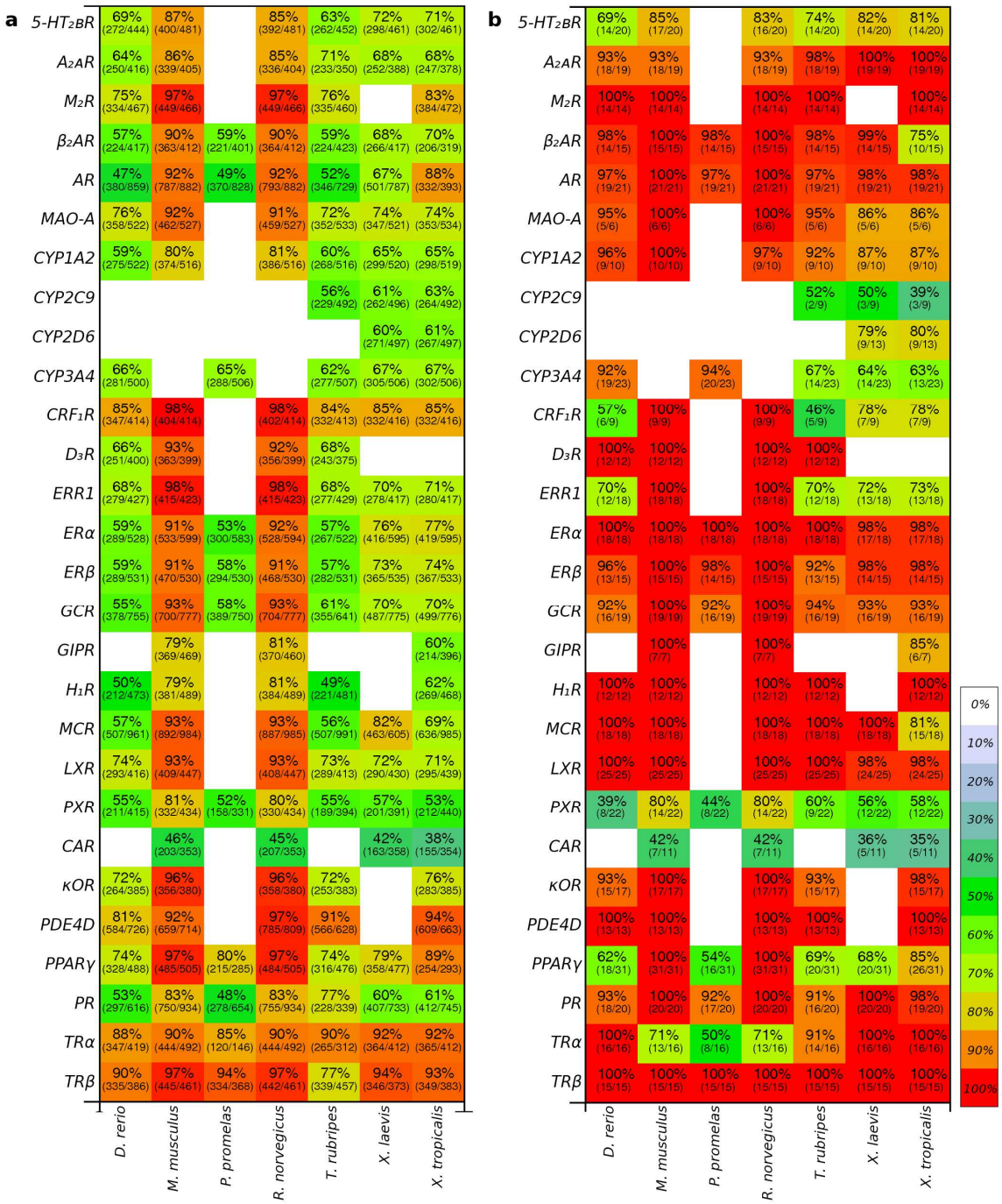

Figure S2. Sequence similarity (percentage and color) and sequence identity (number of identical residues/number of aligned residues is shown in parenthesis) for the 28 toxicity target proteins of **a.** the full sequence and **b.** the ligand-contact residues conserved for 80% of the co-crystalized ligands. White spaces indicate that no ortholog was identified (often due to an incomplete proteome).

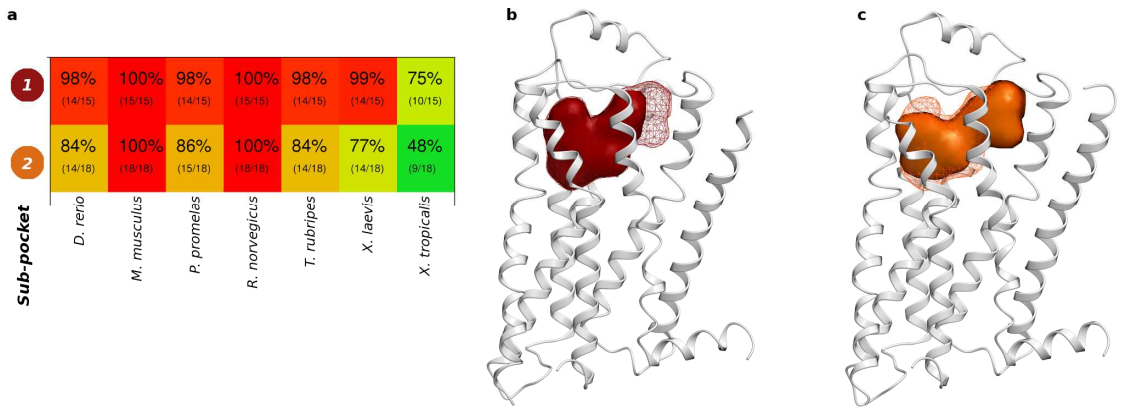

Figure S3. a. Heat-map of the sequence similarity of the two β_2AR sub-pockets (number of identical sub-pocket residues/number of sub-pocket residues). **b.-c.** β_2AR crystal structures (grey ribbons), all co-crystalized ligands (mesh) and sub-pockets 1-2 (solid surface).

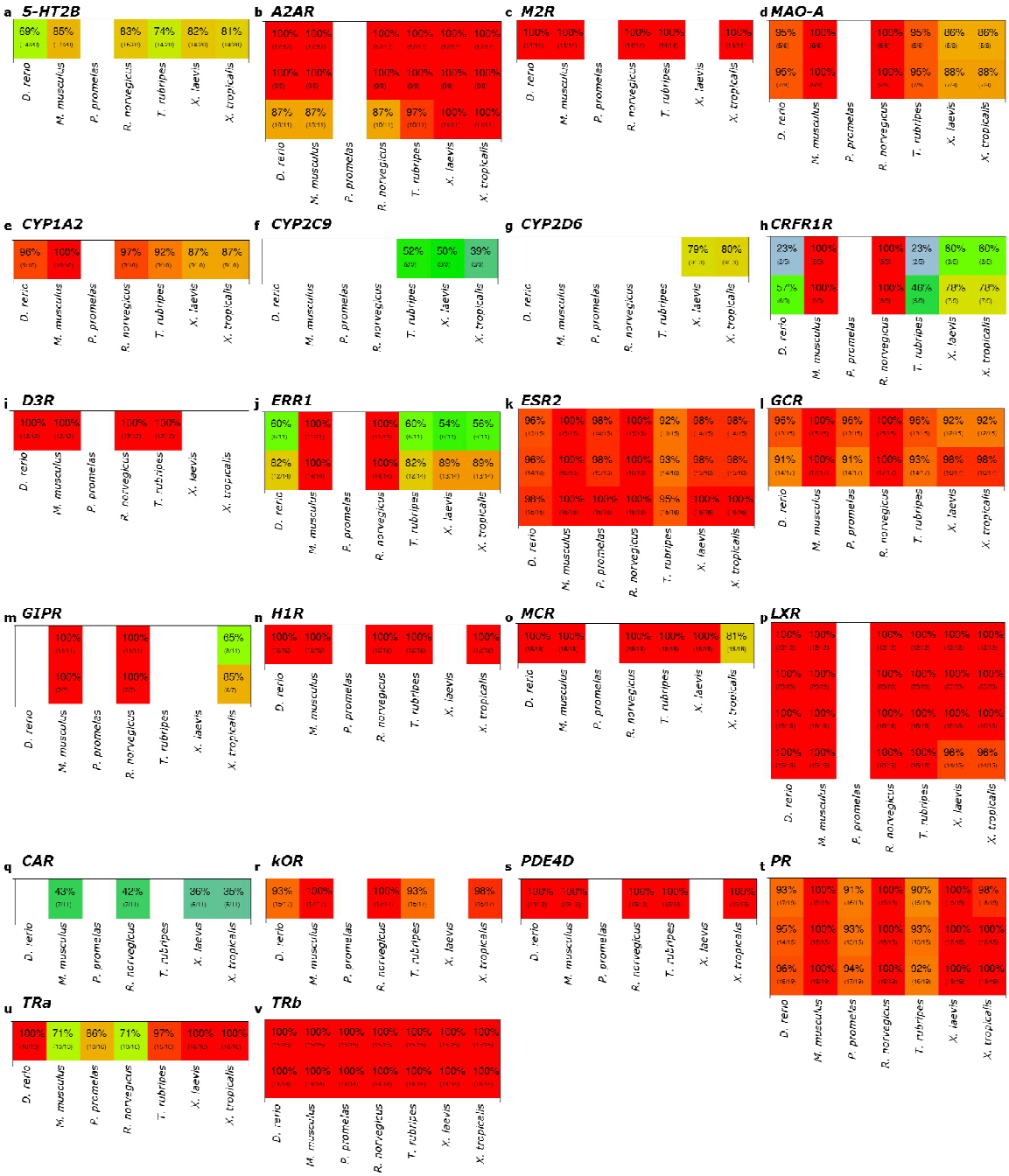

Figure S4. Heat-maps of the sequence similarity and conservation. White spaces indicate that no ortholog was identified (often due to incomplete genome sequencing).

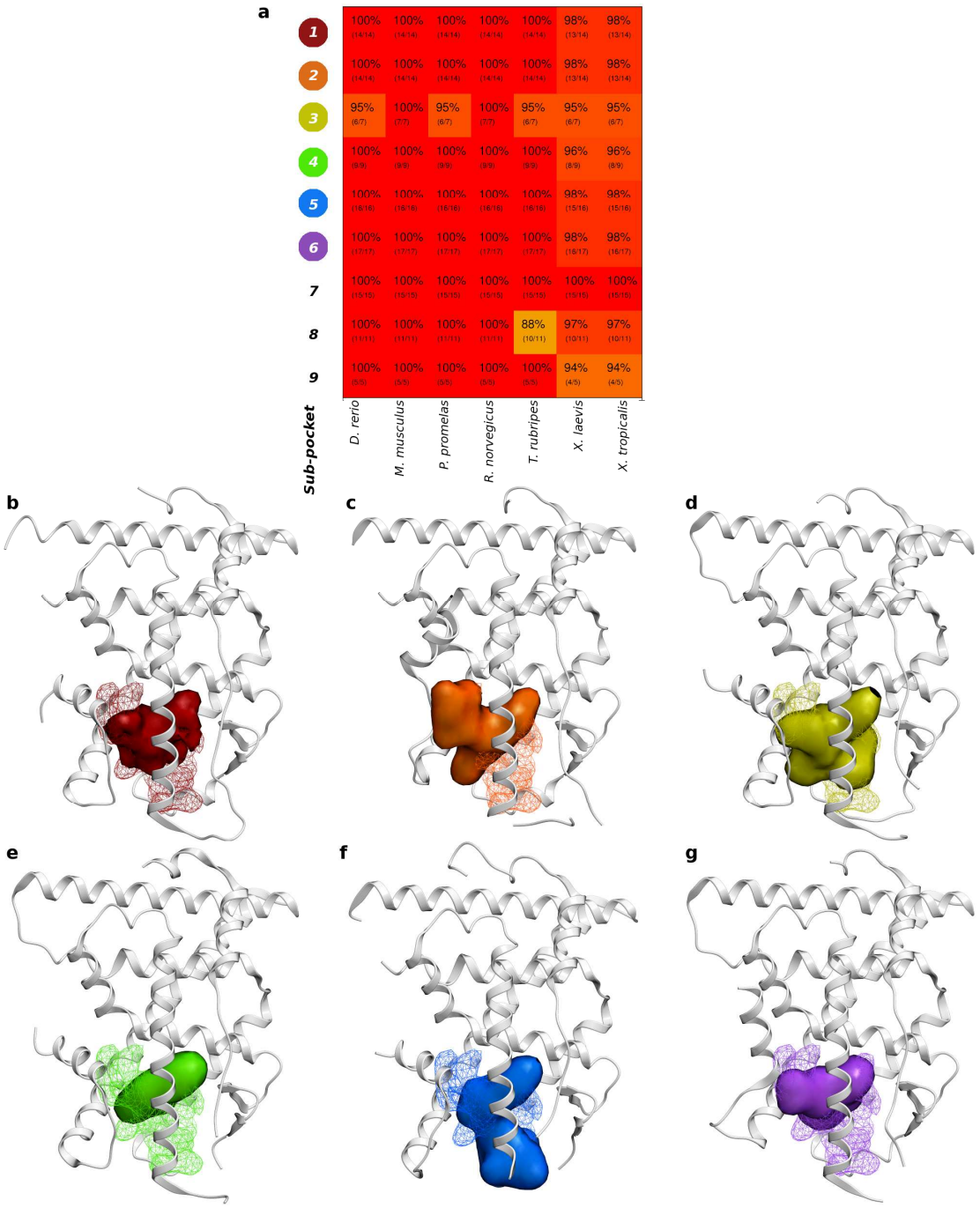

Figure S5. a. Heat-map of the sequence similarity of the six out of nine of the ERα sub-pockets (number of identical sub-pocket residues/number of sub-pocket residues). **b.-g.** ERα crystal structures (grey ribbons), all co-crystalized ligands (mesh) and sub-pockets 1-6 (solid surface).

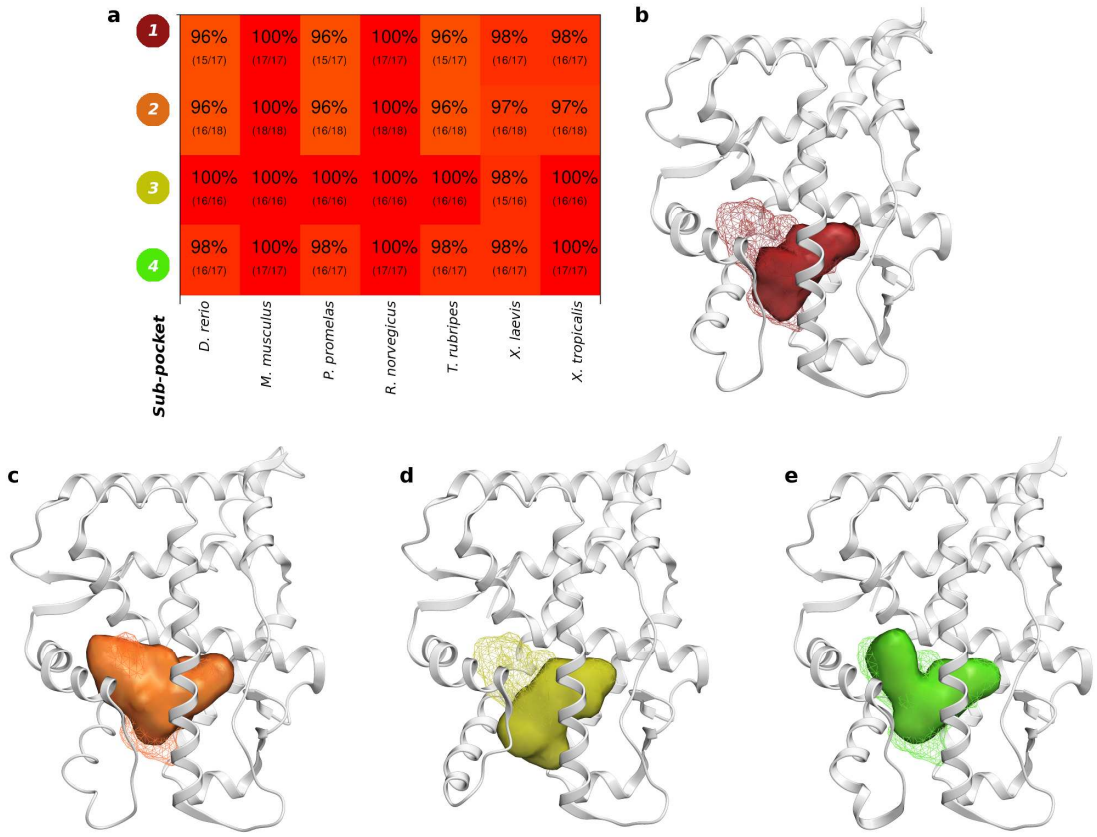

Figure S6. a. Heat-map of the sequence similarity of the four AR sub-pockets (number of identical sub-pocket residues/number of sub-pocket residues). **b.-e.** AR crystal structures (grey ribbons), all co-crystalized ligands (mesh) and sub-pockets 1-4 (solid surface).

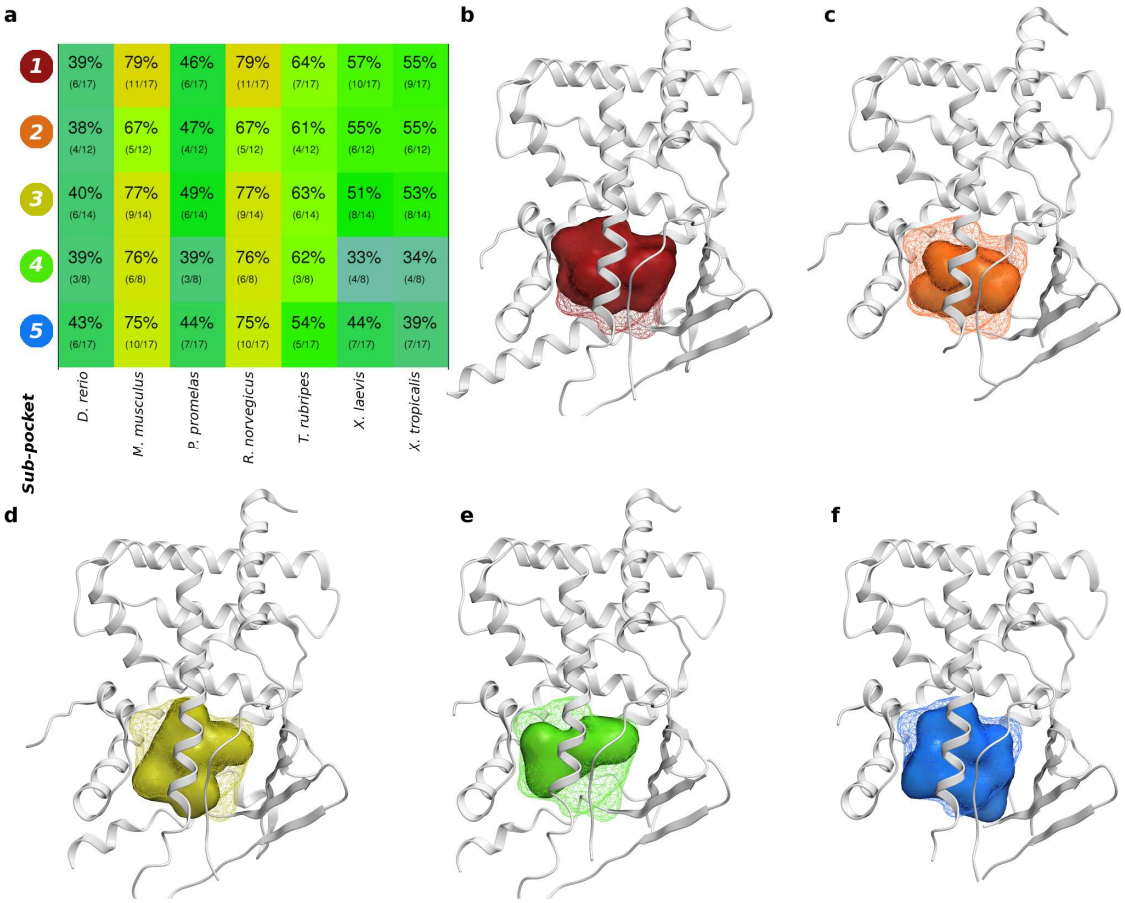

Figure S7. a. Heat-map of the sequence similarity of the five PXR sub-pockets (number of identical sub-pocket residues/number of sub-pocket residues). **b.-f.** PXR crystal structures (grey ribbons), all co-crystalized ligands (mesh) and sub-pockets 1-5 (solid surface).

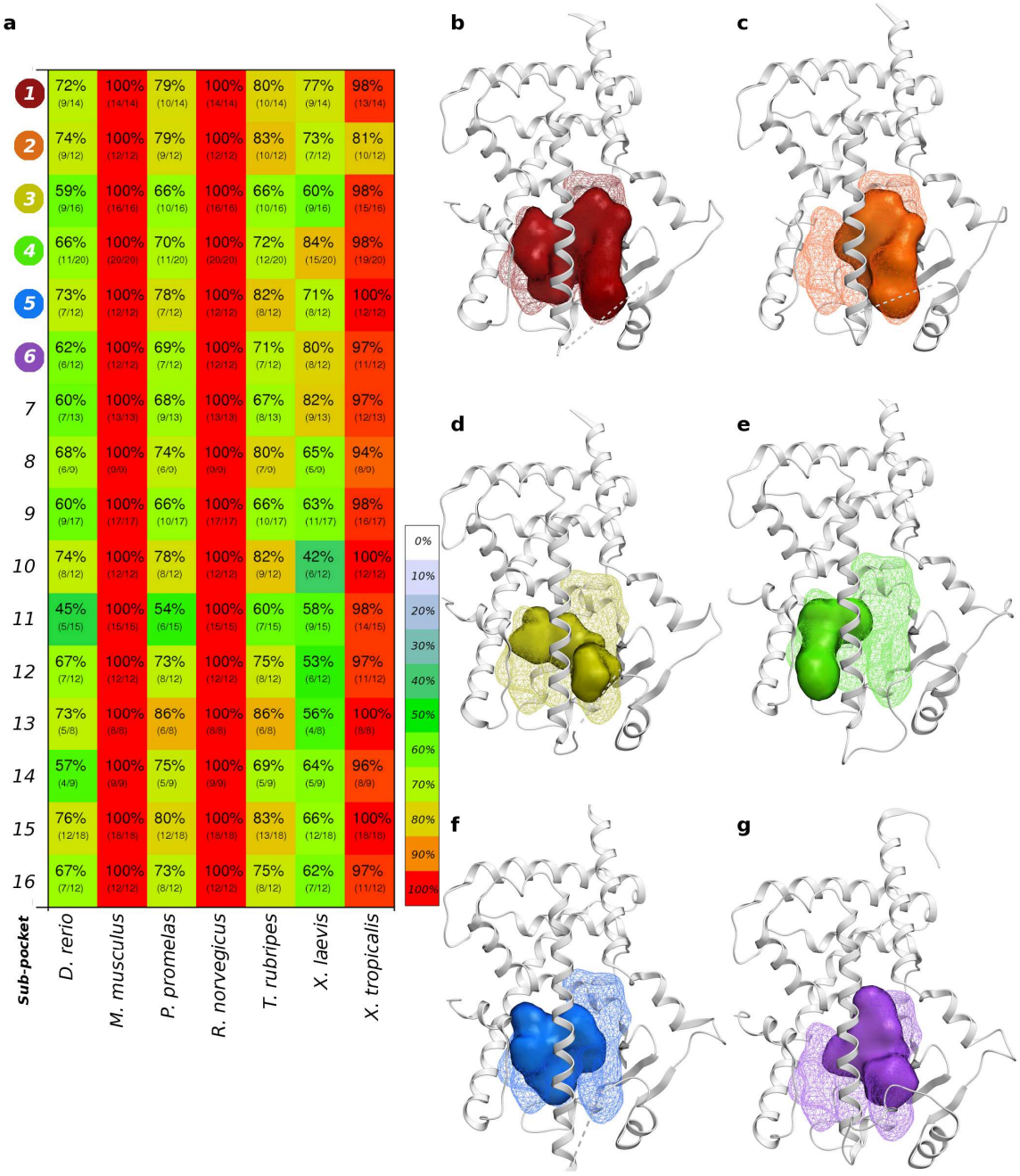

Figure S8. a. Sequence similarity (percentage and color) and sequence identity (number of identical residues/number of aligned residues is shown in parenthesis) for the 16 PPARγ sub-pockets. b-g. Representative PPARγ crystal structures (grey ribbons), all co-crystalized ligands (mesh) and sub-pocket (solid surface); **b.** agonist sub-pocket 1, **c.** partial agonist sub-pocket 2, **d.** partial agonist sub-pocket 3, **e.** partial agonist sub-pocket 4, **f.** partial agonist sub-pocket 5 and **g.** partial agonist sub-pocket 6.

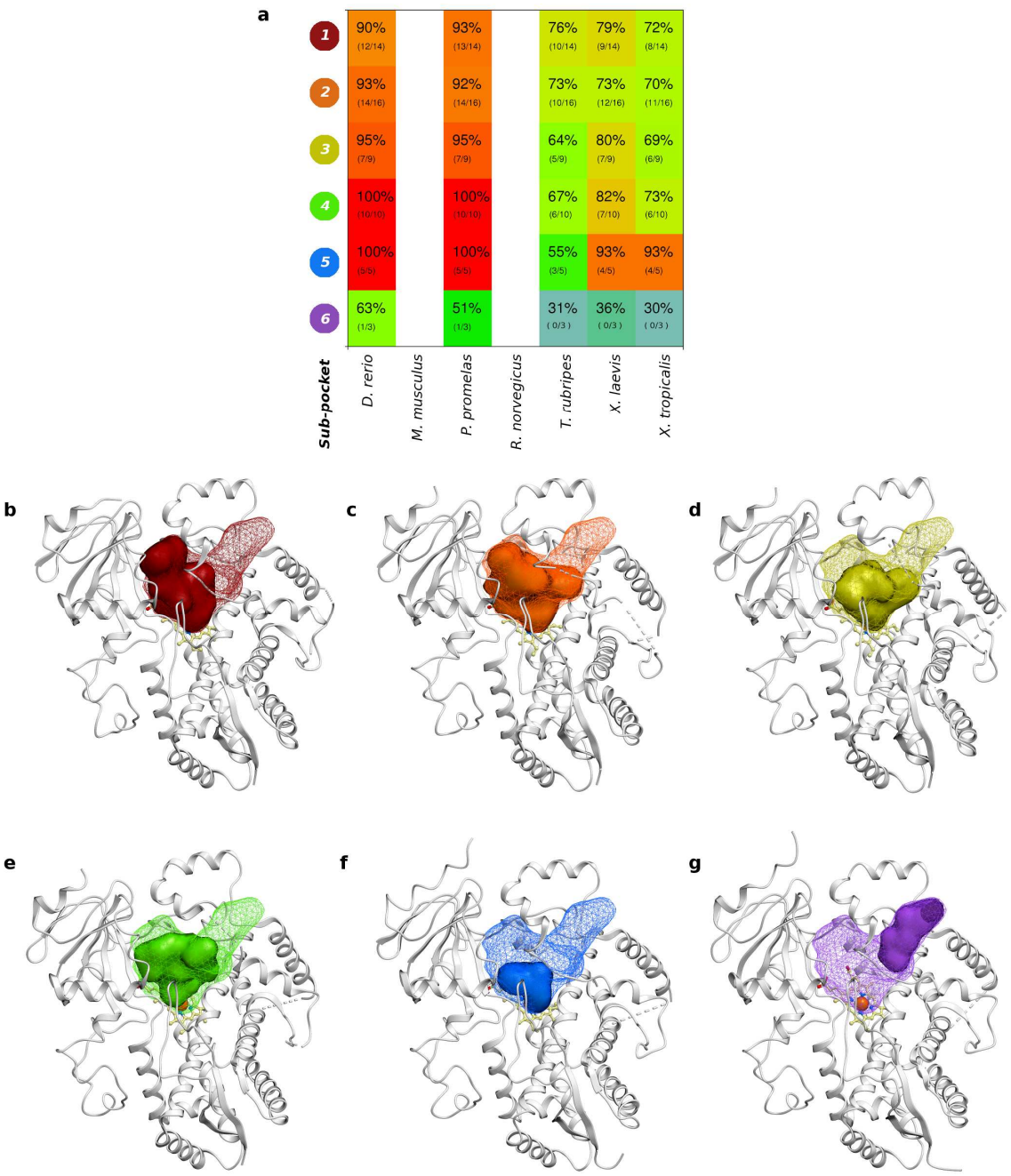

Figure S9. a. Heat-map of the sequence similarity of the six CYP3A4 sub-pockets (number of identical sub-pocket residues/number of sub-pocket residues). **b.-g.** CYP3A4 crystal structures (grey ribbons), all co-crystalized ligands (mesh), sub-pockets 1-6 (solid surface) and the heme co-factor shown as sticks.