Supporting Information

Lipophilic Metabolic Efficiency (LipMetE) and Drug Efficiency Indices to Explore the Metabolic Properties of the Substrates of selected Cytochrome P450 isoforms

Yusra Sajid Kiani[1] and Ishrat Jabeen[1]*

[1]Research Center for Modeling and Simulation (RCMS), National University of Science and Technology (NUST), Sector H-12, 44000, Islamabad, Pakistan

*** Correspondence:** Corresponding Author: Ishrat.jabeen@rcms.nust.edu.pk

Table of Contents

Table S1: Table summarizing the drug classes of the CYP450 substrates included in the CYP1A2, CYP2C9, CYP2C19, CYP2D6 and CYP3A4 datasets that were used for the LipMetE calculations.

CYP450	Drug classes
CYP1A2	antipsychotic, anti-arrhythmic, antineoplastic, analgesic, antipyretic, anti-androgenic, antidepressants and anticoagulants
CYP2C9	anticoagulants, anticonvulsants, antihyperglycemic, proton pump inhibitors, antidepressant of the selective serotonin reuptake inhibitor (SSRI), antidepressant of the serotonin-norepinephrine reuptake inhibitor (SNRI), non-steroidal anti-inflammatory agents (NSAID), anesthetic and antipsychotic agents
CYP2C19	proton pump inhibitors, anticonvulsants, antidepressant of selective serotonin reuptake inhibitor SSRI class, tricyclic antidepressant (TCA), non-steroidal anti-inflammatory agents (NSAIDs), anticoagulants, analgesics, typical and atypical antipsychotic agents.
CYP2D6	tetracyclic antidepressants (TeCA), antidepressants of SNRI and SSRIs, antiarrhythmics, serotoninergic anorectic, antidiarrheals, antipsychotic, non-nucleoside reverse transcriptase inhibitors (NNRTI), antitussive, β1-adrenergic blocking, analgesic, neuroleptic, antihistamine and anticholinergic agents
CYP3A4	antidepressant (TCA, SSRIs), anticoagulants, antidiarrheals, selective estrogen receptor modulators (SERM), proton pump inhibitors, sedatives, nondihydropyridine (non-DHP) calcium channel blockers, NSAIDs, analgesics, antiarrhythmics, HIV protease inhibitors, class I antiarrhythmic, antipsychotic, antiplatelet, antibiotic and vasoactive agents

Section S1: Lipophilic efficiency (LipE) profiling of CYP1A2, CYP2C9. CYP2C19, CYP2D6 and CYP3A4 substrate datasets

Overall, the LipE values of 58 **CYP1A2** substrates (including those used for the LipMetE calculations) range from -1.96 to 7.24 with clogP and K_m values between -2.8 to 6.4 and 0.37 to 3440 µM respectively (Table 2, Table S2). Hitherto, it has been reported in various investigations that CYP1A2 substrates are lipophilic planar hetero-aromatic compounds with clogP values in the range -0.1 to 3.15 [1-5] thus, it represents that our dataset provides a wider range of CYP1A2 substrates. Previously, Leeson and Springthorpe demonstrated a LipE ~5–7 and clogP of ~2.5 for an average oral drug [6]. However, in present dataset only one CYP1A2 substrate, the Aflatoxin B1 exhibits a very high LipE of 7.24 which might be due to its very low clogP (-2.8) and a high affinity against CYP1A2 (K_m=36 µM) as shown in Figure 4a. It might indicate that a very polar small molecule (clogP: -2.8) may not be an ideal candidate for clinic even if it possesses very high LipE because its absorption, binding affinity with CYP1A2, and active transport mediated clearance may hamper it's *in vivo* performance. This is further supported by the study of Lewis et al. who demonstrated that CYP1A2 substrates are hydrophobic in nature and possess high clogP values (0.01- 3.61) that is indicative of the importance of hydrophobic interactions in substrate binding [7].

Similarly, **CYP2C9** substrates are typically medium sized acidic molecules exhibiting clogP in the range of 0.89-5.18 [1-3,5] which is similar to the clogP range (1.61-5.68) of CYP2C9 substrates in our dataset (Table S2). Amongst the 55 CYP2C9 substrates only one compound

Trimethadione showed a LipE value above 5.0 (K_m: 4.8 µM, clogP: 0.31) and a LipMetE of -3.72 (Figure 4b). Only one dietary flavonoid and two drugs Lornoxicam and Sildenafil displayed LipE ≥3 with K_m values of 0.4/3.6/9.6 µM and clogP values of 2.76/2.34/1.98 respectively. The dietary flavonoid displayed negative LipMetE value of -0.189 whereas, for Sildenafil and Lornoxicam the LipMetE parameter could not be determined due to unavailability of V_{max}, $CL_{int,app}$ and C_{prot} parameters. The LipE ≥5 against CYP2C9 and low metabolic stability might reflect a rapid metabolic turnover mainly due to high lipophilicity and high affinity of substrates towards CYP2C9. This can be further correlated with the substrate binding mechanism of CYP2C9. Previously, various authors reported the presence of hydrophobic residues (Phe69, Phe100, Leu102, Leu208, Leu362, Leu366 and Phe476), acidic residues (Asp293, Glu300) and basic residues forming anionic binding site (Arg 105 and Arg 108) within the active site of CYP2C9[8-12]. However, it is also well established that CYP2C9 ligand binding cannot be elucidated by a single type of interaction rather multiple binding mechanisms including electrostatic, hydrogen bonding and π-π stacking interactions within the CYP2C9 active site along with lipophilicity (logP, logD) and acid-ionization constant (pKa) play important role in defining CYP2C9-substrate binding, which is confirmed by further quantitative structure activity relationships (QSAR), homology modeling and site-directed mutagenesis studies (SDM) [3,13-15].

CYP2C19 is typically known for the metabolism of medium to large molecular weight lipophilic amides and amines with logP values between -1.43 to 4.6 [1,3,13,14,16] which is similar to the range of clogP values (-0.78 - 5.59) of our CYP2C19 substrates dataset (Table S2) used for the calculation of LipE. CYP2C19 substrates exhibit LipE values in the range of -4.2 to 5.13 as shown in Figure 4c.

Additionally, it is well documented that **CYP2D6** substrates are usually medium sized, basic and ionized at physiological pH with clogP values between 0.75-5.40 [1,5,16]. Overall, the current dataset of 73 CYP2D6 substrates for LipE calculations displayed clogP in the range of -0.04 to 6.0, with K_m values of 0.057 to 9541 µM and LipE values of -1.14 to 5.12 respectively as shown in Figure 4d and Table S2. In our dataset, β-carboline Harmaline that has neuroprotective properties is the only candidate that presented a LipE of 5.14 with clogP of 0.31 and K_m value of 3.54 µM respectively.

Similarly, **CYP3A4** plays an important role in the metabolism of structurally diverse compounds having logP values in the range of 0.97 to 7.54 which is in line with the clogP values observed for our dataset (-2.8 to 6.82) [1,3,16]. The CYP3A4 substrates within our dataset span the K_m range of 0.04 to 25100 µM (Table S2). Moreover, it has been demonstrated that CYP3A4 substrates are commonly more hydrophobic in comparison to CYP2C9 substrates [3]. Considering the $logD_{7.4}$ values (0.8 to 8.0) of CYP3A4 substrates it is evident that they are lipophilic with a variable degree of ionization [17,18]. Overall, CYP3A4 substrates including 79 drug compounds presented LipE values between -2.51 to 5.31. However, only in case of Aflatoxin B1, a very high LipE of 6.66 with LipMetE of -2.8, clogP of -2.8 and a K_m value of 139 µM has been observed (Figure 4e).

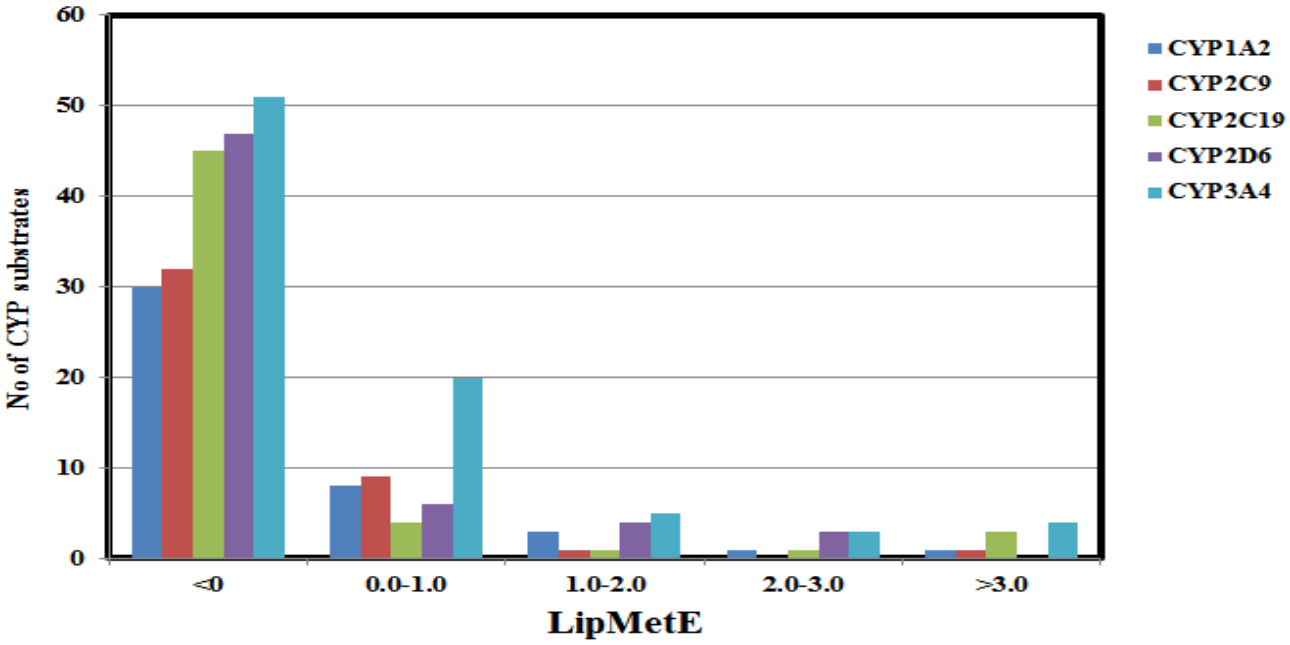

Figure S1. Distribution of LipMetE values across each dataset of the CYP450 substrates.

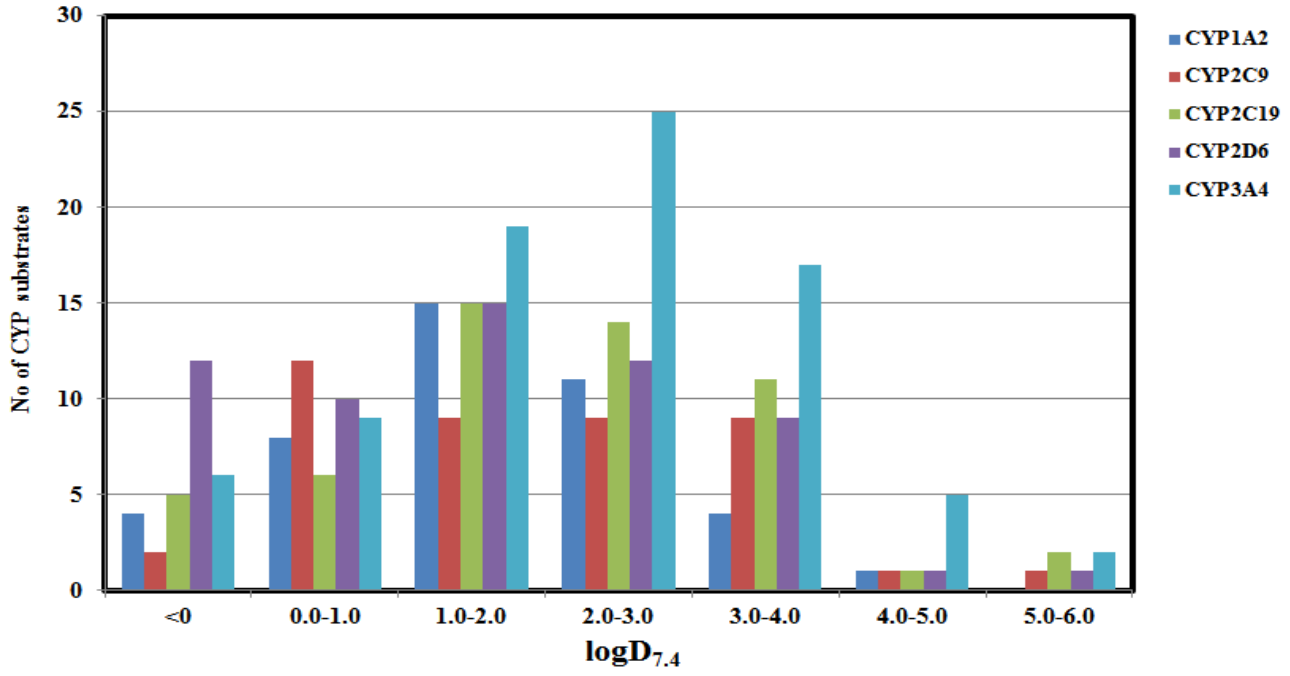

Figure S2. Distribution of $logD_{7.4}$ values across each dataset of the CYP450 substrates.

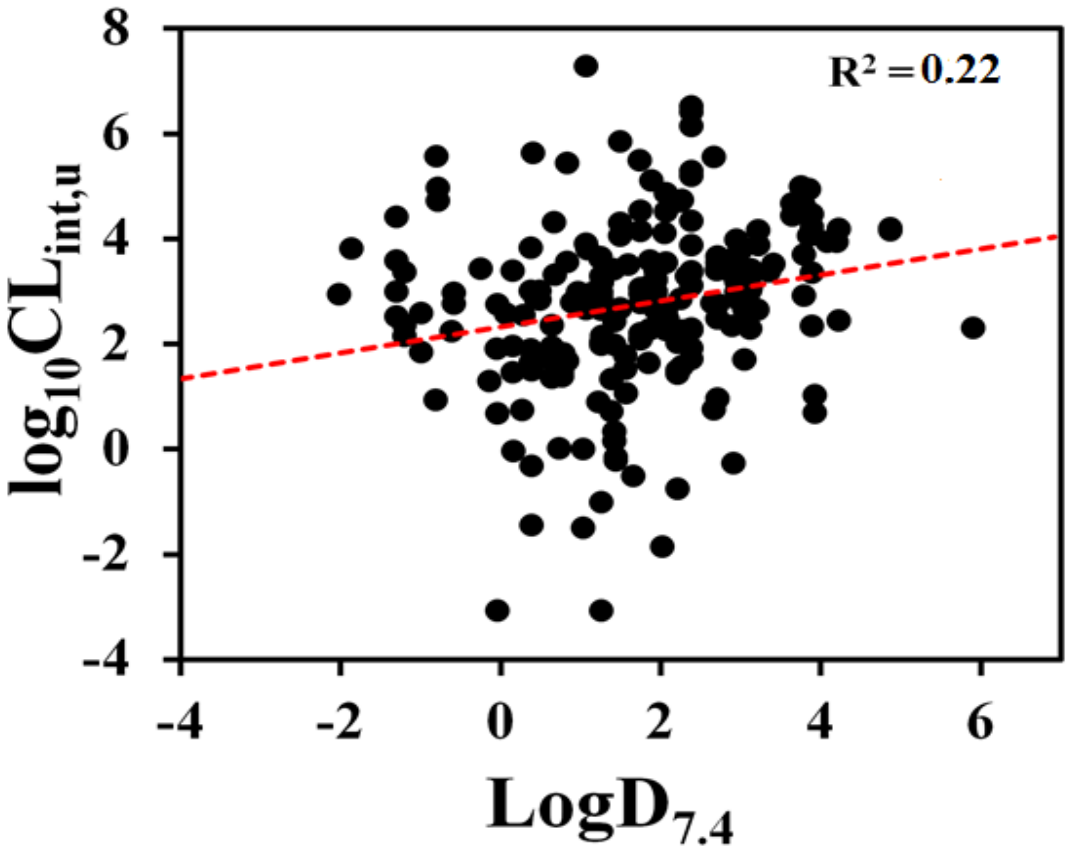

Figure S3: Correlation between intrinsic clearance ($log_{10}CL_{int,u}$) and $logD_{7.4}$ of the entire dataset of CYP450 substrates. A poor or no correlation was observed between intrinsic clearance ($log_{10}CL_{int,u}$) and $logD_{7.4}$ for our dataset.

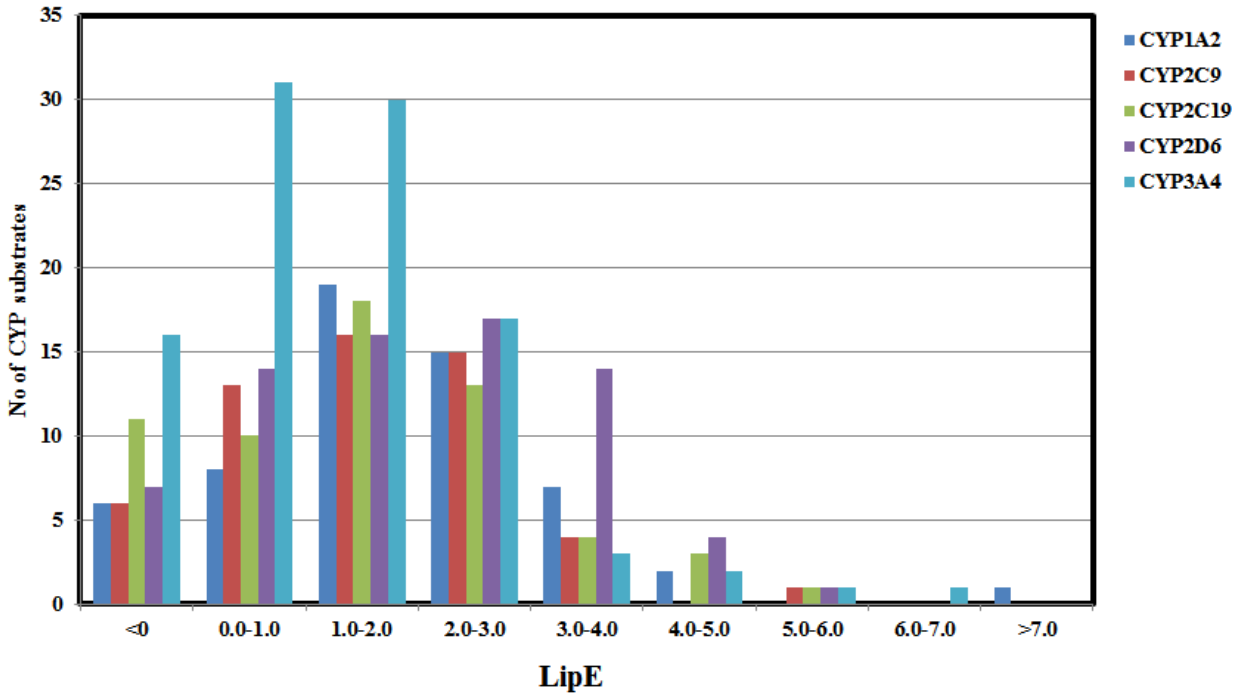

Figure S4. Distribution of LipE values across each dataset of the CYP450 substrates.

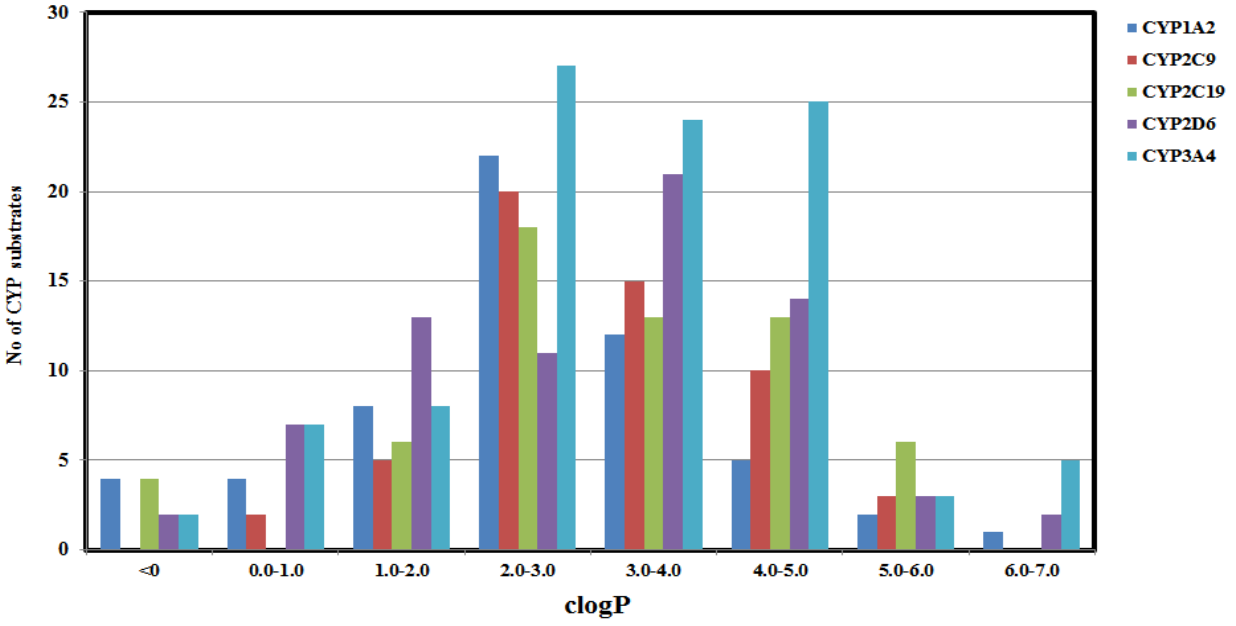

Figure S5. Distribution of clogP values across each dataset of the CYP450 substrates.

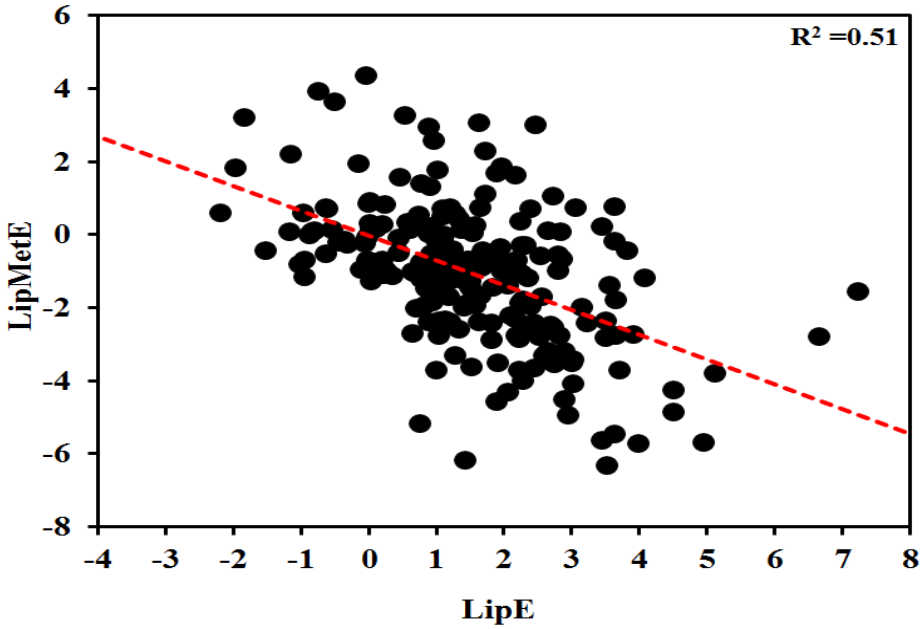

Figure S6: Relationship between metabolic stability (LipMetE) and lipophilic efficiency (LipE) of the CYP1A2, CYP2C9, CYP2C19, CYP2D6 and CYP3A4 substrate datasets.

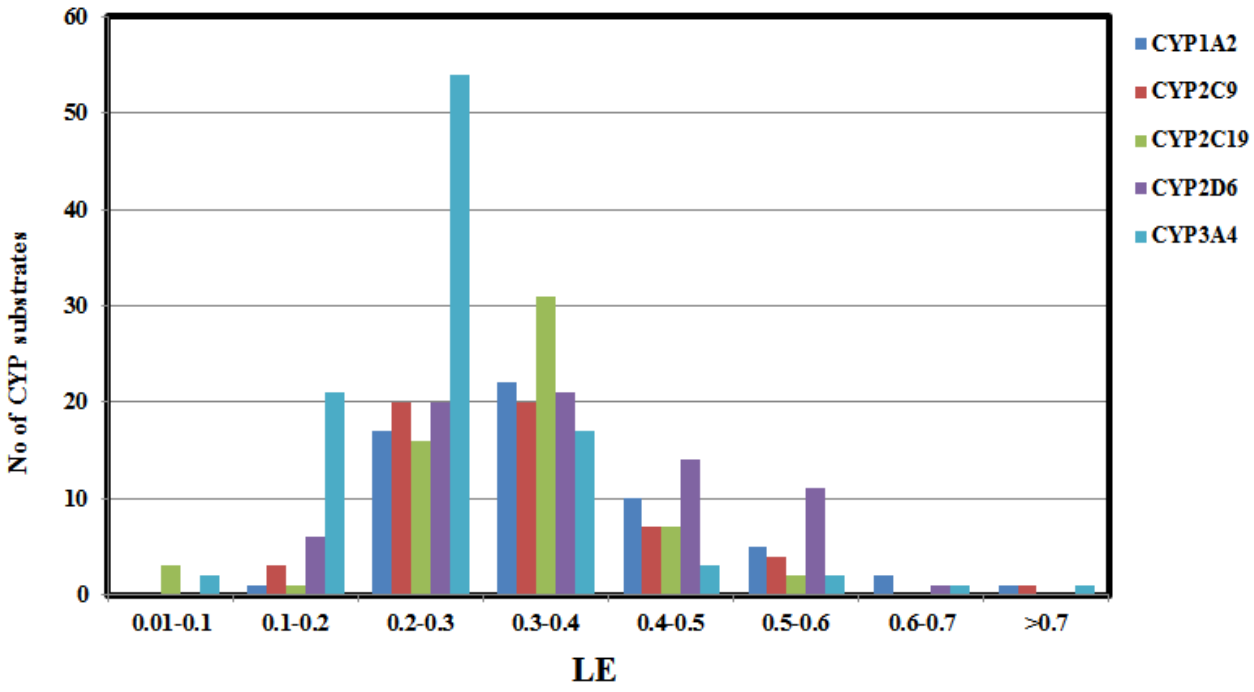

Figure S7. Distribution of ligand efficiency (LE) values across each dataset of the CYP450 substrates.

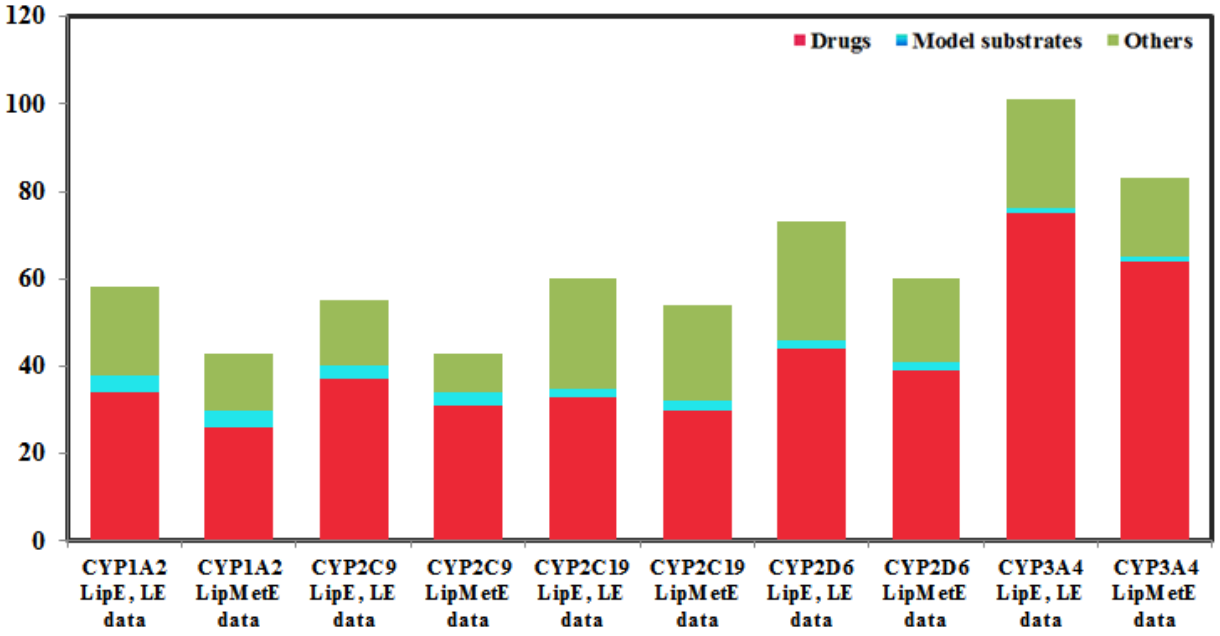

Figure S8: Bar chart representation of marketed drugs, model substrates and other substrates included in the CYP substrate datasets used for LipMetE, LipE and LE profiling against each CYP isoform.

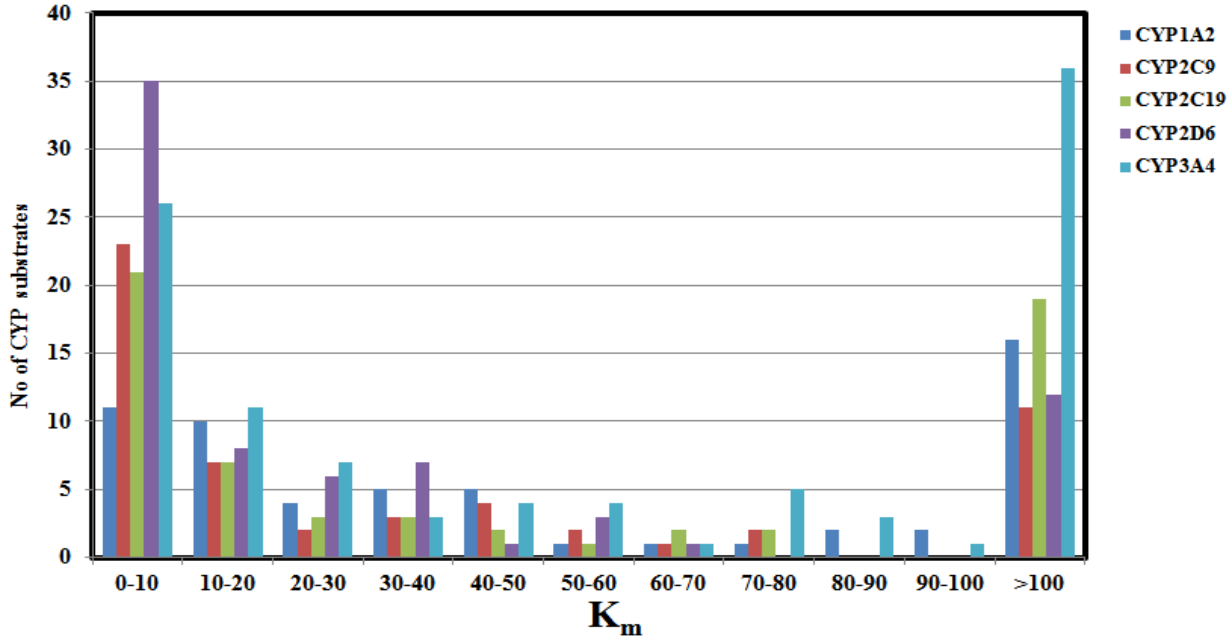

Figure S9. Distribution of K_m values across each dataset of the CYP450 substrates.

References:

1 Lewis, D. F., Jacobs, M. N. & Dickins, M. Compound lipophilicity for substrate binding to human P450s in drug metabolism. *Drug discovery today* **9**, 530-537 (2004).

2 Smith, D. A., Ackland, M. J. & Jones, B. C. Properties of cytochrome P450 isoenzymes and their substrates part 2: properties of cytochrome P450 substrates. *Drug Discovery Today* **2**, 479-486 (1997).

3 Lewis, D. F., Modi, S. & Dickins, M. Quantitative structure-activity relationships (QSARs) within substrates of human cytochromes P450 involved in drug metabolism. *Drug metabolism and drug interactions* **18**, 221-242 (2001).

4 Lewis, D. F. & Dickins, M. Substrate SARs in human P450s. *Drug discovery today* **7**, 918-925 (2002).

5 Lewis, D. F. Structural characteristics of human P450s involved in drug metabolism: QSARs and lipophilicity profiles. *Toxicology* **144**, 197-203 (2000).

6 Leeson, P. D. & Springthorpe, B. The influence of drug-like concepts on decision-making in medicinal chemistry. *Nature Reviews Drug Discovery* **6**, 881-890 (2007).

7 Lewis, D., Lake, B., Dickins, M., Ueng, Y.-F. & Goldfarb, P. Homology modelling of human CYP1A2 based on the CYP2C5 crystallographic template structure. *Xenobiotica* **33**, 239-254 (2003).

8 Ridderström, M. *et al.* Arginines 97 and 108 in CYP2C9 are important determinants of the catalytic function. *Biochemical and Biophysical Research Communications* **270**, 983-987 (2000).

9 de Groot, M. J., Alex, A. A. & Jones, B. C. Development of a combined protein and pharmacophore model for cytochrome P450 2C9. *Journal of medicinal chemistry* **45**, 1983-1993 (2002).

10 Mancy, A., Broto, P., Dijols, S., Dansette, P. M. & Mansuy, D. The substrate binding site of human liver cytochrome P450 2C9: an approach using designed tienilic acid derivatives and molecular modeling. *Biochemistry* **34**, 10365-10375 (1995).

11 Jones, B. *et al.* Putative active site template model for cytochrome P4502C9 (tolbutamide hydroxylase). *Drug metabolism and disposition* **24**, 260-266 (1996).

12 Williams, P. A. *et al.* Crystal structure of human cytochrome P450 2C9 with bound warfarin. *Nature* **424**, 464 (2003).

13 Lewis, D. F., Modi, S. & Dickins, M. Structure–activity relationship for human cytochrome P450 substrates and inhibitors. *Drug metabolism reviews* **34**, 69-82 (2002).

14 Lewis, D. *et al.* Molecular modelling of human CYP2C subfamily enzymes CYP2C9 and CYP2C19: rationalization of substrate specificity and site-directed mutagenesis experiments in the CYP2C subfamily. *Xenobiotica* **28**, 235-268 (1998).

15 Haining, R. L. *et al.* Enzymatic determinants of the substrate specificity of CYP2C9: role of B ‘− C loop residues in providing the π-stacking anchor site for warfarin binding. *Biochemistry* **38**, 3285-3292 (1999).

16 Lewis, D. F. & Dickins, M. Baseline lipophilicity relationships in human cytochromes P450 associated with drug metabolism. *Drug metabolism reviews* **35**, 1-18 (2003).

17 Lewis, D. F., Dickins, M., Eddershaw, P. J., Tarbit, M. & Goldfarb, P. S. Cytochrome P450 substrate specificities, substrate structural templates and enzyme active site geometries. *Drug metabolism and drug interactions* **15**, 1-50 (1999).

18 Smith, D. A. Species differences in metabolism and pharmacokinetics: are we close to an understanding? *Drug metabolism reviews* **23**, 355-373 (1991).