



Kiani, Y. S., Ranaghan, K. E., Jabeen, I., & Mulholland, A. J. (2019). Molecular dynamics simulation framework to probe the binding hypothesis of CYP3A4 inhibitors. *International Journal of Molecular Sciences*, 20(18), [4468]. <https://doi.org/10.3390/ijms20184468>

Publisher's PDF, also known as Version of record

License (if available):  
CC BY

Link to published version (if available):  
[10.3390/ijms20184468](https://doi.org/10.3390/ijms20184468)

[Link to publication record in Explore Bristol Research](#)  
PDF-document

This is the final published version of the article (version of record). It first appeared online via MDPI at <https://www.mdpi.com/1422-0067/20/18/4468> . Please refer to any applicable terms of use of the publisher.

## University of Bristol - Explore Bristol Research

### General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:  
<http://www.bristol.ac.uk/pure/about/ebr-terms>



Article

# Molecular Dynamics Simulation Framework to Probe the Binding Hypothesis of CYP3A4 Inhibitors

Yusra Sajid Kiani <sup>1</sup>, Kara E. Ranaghan <sup>2</sup>, Ishrat Jabeen <sup>1,\*</sup> and Adrian J. Mulholland <sup>2,\*</sup>

<sup>1</sup> Research Center for Modeling and Simulation (RCMS), National University of Sciences and Technology (NUST), Islamabad 44000, Pakistan; yusra1233@rcms.nust.edu.pk

<sup>2</sup> Centre for Computational Chemistry, School of Chemistry, University of Bristol, Bristol BS8 1TS, UK; Kara.Ranaghan@bristol.ac.uk

\* Correspondence: ishrat.jabeen@rcms.nust.edu.pk (I.J.); adrian.mulholland@bristol.ac.uk (A.J.M.)

Received: 24 July 2019; Accepted: 1 September 2019; Published: 10 September 2019



**Abstract:** The Cytochrome P450 family of heme-containing proteins plays a major role in catalyzing phase I metabolic reactions, and the CYP3A4 subtype is responsible for the metabolism of many currently marketed drugs. Additionally, CYP3A4 has an inherent affinity for a broad spectrum of structurally diverse chemical entities, often leading to drug–drug interactions mediated by the inhibition or induction of the metabolic enzyme. The current study explores the binding of selected highly efficient CYP3A4 inhibitors by docking and molecular dynamics (MD) simulation protocols and their binding free energy calculated using the WaterSwap method. The results indicate the importance of binding pocket residues including Phe57, Arg105, Arg106, Ser119, Arg212, Phe213, Thr309, Ser312, Ala370, Arg372, Glu374, Gly481 and Leu483 for interaction with CYP3A4 inhibitors. The residue-wise decomposition of the binding free energy from the WaterSwap method revealed the importance of binding site residues Arg106 and Arg372 in the stabilization of all the selected CYP3A4-inhibitor complexes. The WaterSwap binding energies were further complemented with the MM(GB/PB)SA results and it was observed that the binding energies calculated by both methods do not differ significantly. Overall, our results could guide towards the use of multiple computational approaches to achieve a better understanding of CYP3A4 inhibition, subsequently leading to the design of highly specific and efficient new chemical entities with suitable ADMETox properties and reduced side effects.

**Keywords:** docking; molecular dynamics simulation; CYP3A4; CYP3A4 inhibitors; WaterSwap; residue-wise energy decomposition

## 1. Introduction

Drug metabolism has become a focus for research due to its important role in understanding pharmacotoxicology, pharmacotherapy and applications in drug delivery [1–4]. The human Cytochrome P450 family of heme-containing enzymes are one of the most important classes of enzymes involved in drug metabolism due to their major role in phase I metabolic reactions (N- hydroxylation, sulphoxidation, epoxidation, deamination, desulfuration, dehalogenation, peroxidation, O- and S-dealkylation) of various endogenous and exogenous compounds including drugs [5,6]. The CYP3A4 subtype from the CYP3A family is one of the most dominant drug metabolizing enzymes in humans, responsible for the metabolism of ~50% of currently marketed drugs [7]. The CYP isoforms show an inherent ability for the metabolism of structurally diverse molecules which ultimately leads to the poor bioavailability of various drugs [8,9]. Moreover, CYP activity in intestinal and liver microsomes has been associated with a high number of documented drug–drug and drug–food interactions [10–12]. The combination of high abundance in the liver and a large active site cavity capable of the simultaneous

binding of multiple diverse substrates [13,14] means that it is implicated in undesirable drug–drug interactions and toxicological outcomes [15]. Drug–drug interactions arise mainly due to the inhibition or induction of human CYPs where inhibition might reduce the metabolism and elimination of drugs leading to toxicological outcomes, and induction might compromise the bioavailability and therapeutic efficiency [16]. Similarly, the induction of CYP3A4 might lead to reduced bioavailability due to rapid degradation whereas the inhibition of CYP3A4 leads to increased drug levels and toxicity [17]. Therefore, it is of great importance to understand the underlying mechanisms behind the induction and inhibition of major metabolic enzymes including CYP3A4 so that the impact of these processes can be minimized.

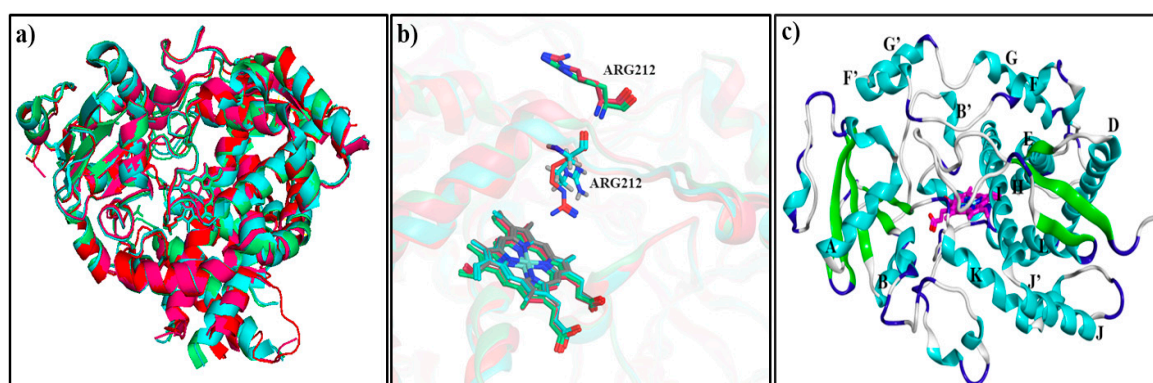
Generally, to probe the molecular basis of CYP450 related metabolism, inhibition and drug–drug interaction potential a number of ligand-, structure-based and integrated modeling studies have been reported in literature [18–25]. Particularly, to gain deeper insights into the binding of ligands to CYP3A4, various *in silico* studies based on quantitative structure–activity relationships (QSAR), pharmacophore models and machine learning methods have been undertaken [26–31]. Structure-based approaches also provide efficient computational methods for the investigation of the interactions responsible for ligand binding to a particular target which is indispensable for understanding metabolic processes and the development of novel leads and therapeutic agents [32]. The availability of X-ray crystallographic structures has also provided a sound basis for computational studies exploring the catalytic mechanism and estimating the binding properties of small molecules with mammalian CYP3A4 [17]. Docking, molecular dynamics simulations and binding energy predictions for the inhibitors and substrates of CYP3A4 have been described extensively [18,33–36]. The binding affinity metric is important and advantageous over expensive experiments for the assessment of ligand interactions with the binding site residues and the desolvation effects [37,38]. For a ligand–protein complex, various methods are available for the prediction of binding energy [39]. In the current study, a combination of docking, molecular dynamics simulations (MD) and binding free energy predictions (using the WaterSwap method [40]), molecular mechanics/Poisson–Boltzmann surface area (MM/PBSA) and molecular mechanics/generalized-Boltzmann surface area (MM/GBSA) [41,42] have been performed for a set of highly efficient CYP3A4 inhibitors to investigate CYP3A4 inhibition. The ligand affinity for CYP3A4 and the overall contribution of the binding site residues towards the stabilization of the inhibitor-bound complex can be identified through WaterSwap calculations, insight which might aid the development of new therapeutic options by reducing drug–drug interactions and the associated undesirable toxic effects.

## 2. Results and Discussion

### 2.1. Protein Structure Selection

To date, about 25 CYP3A4 crystallographic structures in ligand-bound and ligand-free forms are available in the Protein databank and the structures 3NXU (2 Å) [43], 1TQN (2.05 Å) [44], 3UA1 (2.15 Å) [45] and 4K9W (2.40 Å) [46] were selected for further conformational analysis. 1TQN is a structure of the apo-state of the enzyme, whereas the remaining three structures represent ligand-bound states. Moreover, a few N- and C-terminal residues were also missing in all structures. In addition, 3NXU has missing G-H loop residues (Glu264-His267) and H-I loop residues (Ser281-Lys288), 1TQN has missing H-I loop residues (Lys282-Glu285), whereas 3UA1 has missing G-H loop residues (Ser259-Leu261, Lys266-Arg268) and H-I loop residues (Asn280-Ser286) and 4K9W has a missing region of G-H loop (Gln265-His267) and H-I loop residues (Ser281-His287). The structure 1TQN, with fewer missing residues (only at H-I loop in addition to the missing -N/-C terminal residues) was selected as a template for the calculation of root-mean-square deviation of the  $C_{\alpha}$  atom's ( $C_{\alpha}$ RMSD) in comparison to the 3NXU, 3UA1 and 4K9W structures. Overall, a  $C_{\alpha}$ RMSD of 1.28 Å of 1TQN as compared to 3NXU, 3UA1 and 4K9W shows that the backbone conformations of these structures are similar (Figure 1a). It is notable that in 1TQN and 3UA1, Arg212 is positioned in the vicinity of heme within the active site,

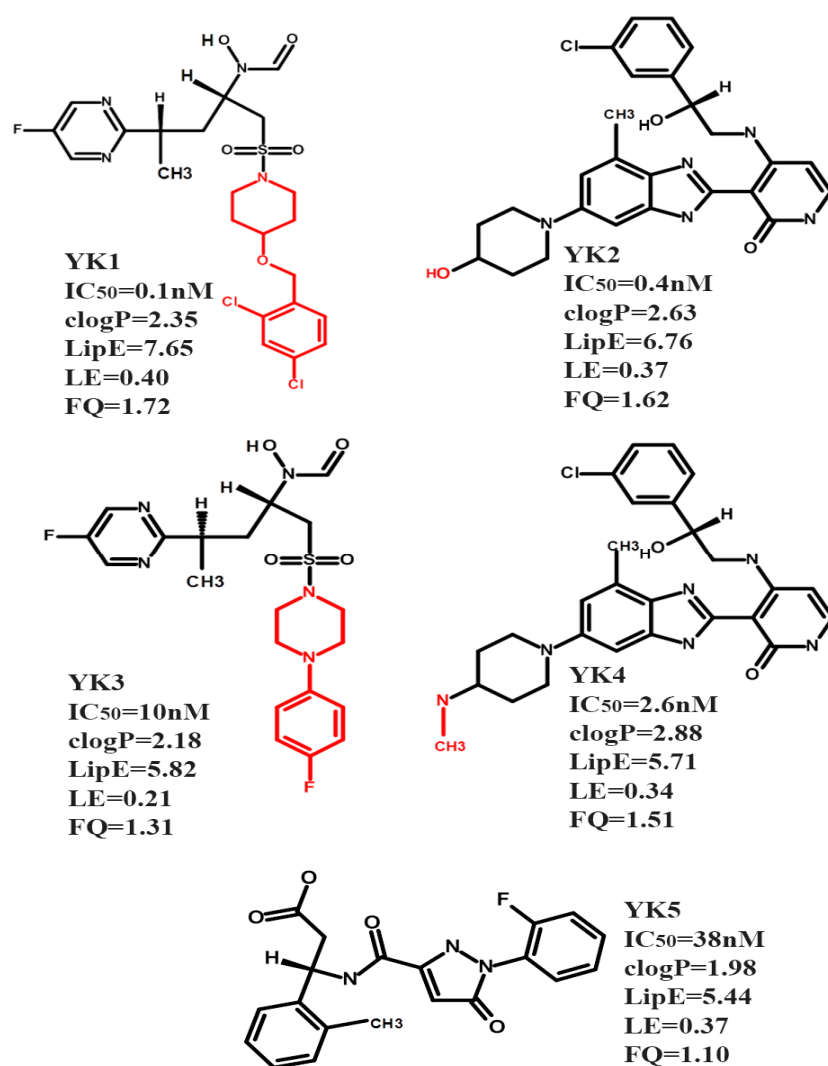
whereas in the other ligand-bound structures it is pointing outwards (Figure 1b). The positioning of Arg212 towards the active site is implicated to be important in terms of substrate/inhibitor binding within the CYP3A4 binding site [45,47]. Therefore, due to a reasonable RMSD of 1TQN to 3UA1 (0.52 Å) and all other structures, fewer missing residues and the position of Arg212 towards binding site, 1TQN was selected as the final structure after building the missing residues using Modeller [48]. The overall structural organization of 1TQN is shown in Figure 1c and it is notable that the helices F-F' and G-G' form the roof of the heme embedded binding site.



**Figure 1.** Superposition of chain A in CYP3A4 crystal structures (a) 1TQN (red), 3UA1 (cyan), 3NXU (green) and 4K9W (pink); (b) positioning of Arg212 within 1TQN (red), 3UA1 (cyan), 3NXU (green) and 4K9W (pink); and (c) the structural organization of 1TQN.

## 2.2. Ligand Selection

Initially, the concept of drug efficiency metrics has been applied for the selection of efficient modulators against true drug targets [49]. Recently, drug efficiency metrics were used to select the highly efficient inhibitors of CYP450 isoforms which represents an anti-target [22]. Only five CYP3A4 inhibitors (Figure 2) out of the large data set of 2747 compounds (Table S1) showing the defined thresholds (Materials and Methods and Section S1) of activity ( $IC_{50}$ : 0.1–38 nM), clogp (1.98–2.88), lipophilic efficiency (LipE) (5.44–7.65), ligand efficiency (LE) (0.21–0.40 kcal mol<sup>-1</sup> HA<sup>-1</sup>) and fit quality  $\geq 1$  (Figure S1a,b) were selected for further docking and molecular dynamics studies to probe the binding hypothesis within the active site of CYP3A4.



**Figure 2.** The chemical structures and experimental biological activity ( $IC_{50}$ ) values of the selected CYP3A4 inhibitors fulfilling the drug efficiency criteria. For structural analogues the common scaffold is shown in black and the R groups are presented in red. The selected CYP3A4 inhibitors YK1–5 have the following ChEMBL IDs CHEMBL1683444 (YK1), CHEMBL520419 (YK2), CHEMBL1683445 (YK3), CHEMBL482102 (YK4) and CHEMBL3145341 (YK5) and are highlighted yellow in Table S1.

### 2.3. Docking and Molecular Dynamics Simulation (MD)

Our aim here is not to investigate drug–drug interactions, but rather to investigate the potential binding modes of the chosen CYP3A4 inhibitors: this is the initial issue and primary question of interest in drug development process, i.e., is it likely that a given compound will inhibit CYP3A4. A secondary question, which is relevant for compounds to be used in the clinic, is whether drug–drug interactions may be expected, and how they occur as reported for warfarin by Lonsdale et al. [35]. This might be an interesting question for similar studies in the future. Therefore, the primary aim of our study is the elucidation of the binding hypothesis of the selected inhibitors using the experimentally determined structure of the soluble portion of the CYP3A4 enzyme (1TQN). Generally, the CYP450 microsomal enzymes are partly immersed in the membrane and anchored through the N-terminal helix. The heme shows a tilt angle due to its orientation with reference to the plane of membrane and the F'/G' loop is deeply buried while the B/C-loop and  $\beta 1$ ,  $\beta 2$ ,  $\beta 4$  and  $\beta 5$  sheets also show substantial contacts with the membrane [50,51]. The CYP proximal site shows an interface with the cytosol while the distal site where the substrate ingress and egress channels are located is positioned towards the lipid bilayer [52].

Docking, molecular dynamics and QM/MM approaches have remained fruitful for the elucidation of oxidation selectivity of substrates whereas to bring about complete perspective of drug metabolism other key factors, such as substrate access through the membrane and binding to the active site, must be taken into consideration [53,54].

Previously, several authors have investigated the effects of including the membrane on binding with major CYP subtypes and it was shown that in comparison to crystallographic structures and single solution simulations the consideration of membrane binding instigates conformational changes to the substrate ingress and egress channels at the distal site [35,51,52,55,56]. We have also previously investigated the effects of including the membrane (and the membrane binding domain) on binding in CYP3A4 [35]. Comparisons of simulations with and without the membrane present showed that the inhibitor binding mode is not significantly affected by the presence of the membrane; as there was a minimal effect on the protein conformational flexibility and residues in the near vicinity of heme proximal site yet, significant differences were observed in the substrate access channels of both systems [35]. However, for other properties (e.g., pathways of association and binding of inhibitors or substrates that pass through the membrane), in contrast, consideration of the membrane is essential [35]. It is also well known, that N-terminal  $\alpha$ -helix connects the human CYPs to the membrane, but due to difficulties in dealing with CYP-membrane-bound structures and to ease the process of crystallization, hydrophilic residues are added as a replacement of this helix [17]. It is also observed that the deletions in the crystallized structures did not reduce the catalytic effect or the ability of the enzyme to attach to membrane surface hence, these engineered structures are biochemically relevant [57,58]. Therefore, owing to the complexity of dealing with CYP-membrane-bound structures, this approach of modeling the soluble domain only has been applied successfully in investigations of inhibition and activity of CYP3A4 and other P450s [36,59–62].

Herein, the docking and molecular dynamics simulations are based on the experimentally determined structure of the soluble part of the enzyme. The AutoDock [63] molecular docking platform was used to generate a maximum of 10 different binding conformations per ligand. The concept of alternative binding modes is clearly very important for binding in particular with the highly plastic, CYP3A4 active site [47]. However, for our dataset, a careful inspection of the observed docking conformations of the selected CYP3A4 inhibitors showed no evidence of binding to multiple sites within the CYP3A4 active site rather, the docked conformations depict clustering at a similar position with reasonable RMSD values (1.46–3.5 Å) as shown in Figure S2a–e. Therefore, as reported in various studies the minimum energy pose ranked through the free energy scoring function was selected for further analysis [36,64]. The docking results of the selected ligand–protein complexes including Autodock energy (kcal/mol), the residues involved in van der Waals interactions and hydrogen bonding in the docked structures are given in Table 1.

AMBER14 [65] was used to explore the conformational flexibility and stability of the inhibitor-bound complexes as described in the Materials and Methods section. The trajectory analysis was performed using CPPTRAJ [65] through the determination of root-mean-square deviation of the  $C_{\alpha}$  atoms ( $C_{\alpha}$ RMSD), root-mean-square fluctuation ( $C_{\alpha}$ RMSF), radius of gyration (Rg) and hydrogen bond analysis for each complex. For the selected five CYP3A4 inhibitor-bound complexes multiple runs of MD (two for each complex) were performed. The conformational analysis of the replicated MDs indicates that the MD results do not differ significantly as shown by difference of  $\sim 0.09$  to  $0.57$  Å and  $\sim 0.01$  to  $0.07$  Å in the  $C_{\alpha}$ RMSD and the Rg values for the ligand-bound complexes (CYP3A4-YK1–5) (for detailed analysis refer to Supplementary Materials Section S2 and Figure S3). Therefore, the initial run of MDs (MD1) with smaller average  $C_{\alpha}$ RMSD and Rg values were considered final for further analysis. The RMSD of  $C_{\alpha}$  atoms in the apo-state of CYP3A4 and the five selected ligand-bound models was measured as a function of time with respect to the initial structure (Figure S4). For the unliganded CYP3A4 (1TQN) protein, high fluctuations were observed in the  $C_{\alpha}$ RMSD up to 10 ns, after which it stabilized at an average  $C_{\alpha}$ RMSD value of  $2.83 (\pm 0.4)$  Å. However, the ligand-bound simulations took longer to stabilize, with high fluctuations in the RMSD over the first  $\sim 25$ – $30$  ns for CYP3A4-YK1 and

CYP3A4-YK4, and ~35 ns for the CYP3A4-YK2, CYP3A4-YK3 and CYP3A4-YK5 complexes (Figure S4). Overall, with the exception of the CYP3A4-YK1 (average  $C_{\alpha}$ RMSD: 3.13 ( $\pm 0.7$ ) Å) the remaining complexes CYP3A4-YK2-YK5 displayed average  $C_{\alpha}$ RMSD values of 2.23 ( $\pm 0.4$ ) Å, 2.11 ( $\pm 0.6$ ) Å, 2.09 ( $\pm 0.3$ ) Å and 2.31 ( $\pm 0.7$ ) Å, respectively, that lie in a comparable range to the average  $C_{\alpha}$ RMSD (2.83 ( $\pm 0.4$ ) Å) of the apo-protein (1TQN) indicating the stability of all the selected systems.

The  $C_{\alpha}$  based RMSF values of the ligand-bound complexes show the highly plastic or stable regions of the proteins during the course of MD simulations. For all five inhibitor-bound complexes our results indicate a fluctuation pattern similar to that obtained for the apo-protein, showing the highest fluctuations in flexible regions including the channel, loop forming regions, the N and C terminal regions as shown in Figure S5. Overall, for all systems a few core regions and the loop regions including the G-H and the built-in missing H-I loop demonstrated the highest RMSF values (within 2.0–5.4 Å). After the assessment of conformational changes in the CYP3A4 structure upon ligand binding through RMSF calculations, the radius of gyration was also measured for the apo-protein and the selected systems to predict the overall compactness of the CYP3A4-inhibitor-bound complexes. The radius of gyration (Rg) parameter shows the equilibrium conformation of a total system and is calculated as the root-mean-square distance of an assemblage of atoms from their common center of mass [66]. The Rg of the apo-protein 1TQN shows a highly compact protein with an average Rg value of 44.3 ( $\pm 0.03$ ) Å for the 50 ns MD, with the ligand-bound complexes giving very similar values at 44.3 ( $\pm 0.05$ ) Å, 44.4 ( $\pm 0.06$ ) Å, 44.5 ( $\pm 0.1$ ) Å, 44.3 ( $\pm 0.03$ ) Å and 44.4 ( $\pm 0.04$ ) Å for complexes of YK1–5, respectively. Overall, the Rg values of the complexes CYP3A4-YK1–5 show that upon ligand binding the CYP3A4 retains a compact conformation as shown in Figure S6.

Detailed interaction profiles of the selected CYP3A4 inhibitors with important binding site residues of CYP3A4 before and after 50 ns of MD, along with the centroid structures for each complex, are provided in Table 1. Here, the centroid structures for all CYP3A4-inhibitor bound complexes were obtained after performing clustering on the basis of RMSD values of important active site residues as described in the Material and Methods section. It is notable that for the selected inhibitors, the docking energies vary from –11.7 to –9.5 kcal/mol, a fairly narrow range for a series of ligands that span one order of magnitude in terms of  $IC_{50}$  values and the predicted ordering of the ligands shows little correlation with the experimental data (Figure S10a). Phe57, Arg105, Ser119, Arg212, Phe215, Thr309, Ser312, Ala370, Arg372, Glu374 and Leu483 were observed to form significant interactions with the ligands (Table 1). The selected docking poses of the inhibitors (YK1 to YK5) were positioned in the close vicinity of the buried helix I, near the heme and below the roof formed by the loop region between the F'-F and G'-G helices. The docking and MD results of the selected compounds (YK1–5) show no evidence of direct coordination to iron (i.e., type II binding), indicating that such close approach does not occur for any of these compounds. However, the known limitations of docking methods, and of empirical (MM) force fields for modeling metalloprotein interactions, are well known and should be considered. Modeling such interactions where the inhibitor may bond directly to iron, would require sophisticated methods capable of representing a bonding interaction, e.g., QM/MM calculations employing good quality density functional theory [35,53,67–69]. Whereas, the approaches used here allowing for movement of the ligand indicate that this mode of binding is not likely.

Overall, a close assessment of the binding modes of inhibitors (YK1–5) show that YK1, YK2, YK3 and YK4 protrude from the proximal binding site towards the distal site (Figure S8a,c,e,i) whereas, YK5 binds at the distal binding site (Figure S8g) with the phenylalanine-cluster (Phe108, Phe215, Phe219, and Phe304), Ser119 and arginine residues (Arg105, Arg106, Arg212, Arg372) contributing significantly towards interactions (Table 1). The docked ligands YK1–5 within the proximal and distal sites of CYP3A4 active site are shown in Supplementary Materials Figure S7a–e as discussed by Tanaka et al. [70]. Residues including Ser119, Ile301, Phe304, Ala305, Ile369, Ala370 and Glu374 have been identified as important for binding through site-directed mutagenesis and X-ray crystallographic studies [44,71–74]. Likewise, our docking results are well in line with the previously reported studies that have shown the importance of active site residues Val101, Asn104, Arg105, Met114, Ser119,

Leu211, Arg212, Asp214, Asp217, Pro218, Glu374, Ile301, Ala305, Thr309, Ile369, Ala370, Leu373, Glu374, Ser478, Leu479 and the Phe-cluster for interaction with the inhibitors and substrates of CYP3A4 [33,44,71,74,75]. MD simulations show that the hydrogen bonds identified in the docking poses did not persist throughout the trajectory; however, new hydrogen bonds are formed during the course of the MD simulations. The time dependent hydrogen bond analysis of CYP3A4-inhibitor complexes during the 50 ns simulation demonstrates that the highly potent inhibitors YK1 and YK2 form the highest numbers of hydrogen bonds in comparison to inhibitors YK3-YK5 as shown in Figure S9a–e.

The analysis of hydrogen bond distances between CYP3A4 and YK1 shows that the hydrogen bonds with Arg105 and Arg212 in the docked complexes did not persist during the MD. However, in the CYP3A4-YK1 centroid structure, hydrogen bonds were identified between YK1:Carbonyl=O3-Glu374:NH and YK1:Hydroxyl-O2-Arg372: =O (Figure S8b) which remained stable till the end of the 50 ns MD simulation (Table 1). Other residues including Arg105, Arg106, Arg212 and Ala370 formed hydrogen bonds only during part of the MD simulation as shown in Figure 3a. For YK3, a structural analogue of YK1, three hydrogen bonds were observed between the inhibitor and Arg105, Arg106 and Arg372 of the centroid structure (Figure S8f). Figure 3c represents the time dependent distance analysis of the hydrogen bonds formed between YK3 and Arg105–106 during the 50 ns MD simulation. The hydrogen bond between YK3:O3-Arg105:HH1 remained stable during the last 30 ns of the MD simulation (Figure 3c).

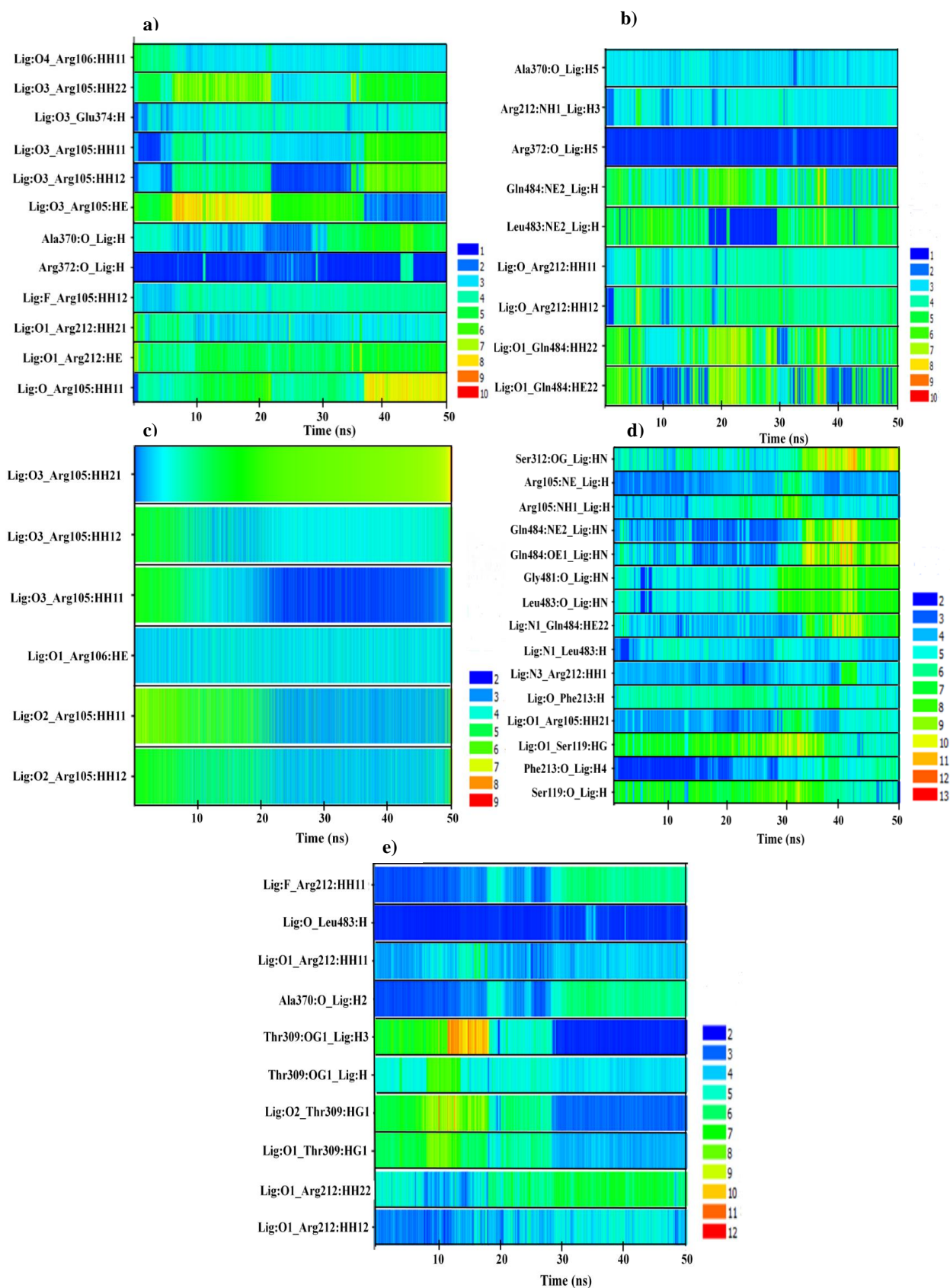
Additionally, the CYP3A4 inhibitors YK2 and YK4, yet another pair of structural analogues, formed hydrogen bonds with Ser312 in the pose from docking (Figure S8c,g), whereas in the centroid structures, YK2 and YK4 formed one and three hydrogen bonds, respectively (Table 1, Figure S8d,h). However, after MD, hydrogen bonds between YK2:Hydroxyl-O2- Arg372:O and YK4: Hydroxyl-O1-Ser119:O were observed as shown in Table 1. The time series distance profiles of CYP3A4-YK2 and CYP3A4-YK4 during the 50 ns MD simulations show that the hydrogen bonds formed between Arg372:O-YK2:H5 remained stable throughout the trajectory and the bond between Phe213:O-YK4:H4 was observed for a short duration as shown in Figure 3b,d. Prior to MD, YK5 displayed hydrogen bonding interactions with Arg212 and Glu374 (Table 1 and Figure S8i), whereas, in the centroid structure Arg212, Ala370 and Leu483 were involved in hydrogen bonding (Table 1 and Figure S8j). A stable hydrogen bond between Leu483:H-YK5:O was observed in the MD trajectory and hydrogen bonds between Arg212:HH11-YK5:F and Thr309:OG1-YK5:H3 were only stable for a fraction of the simulation. The distance of the hydrogen bonds between amino acid residues and YK5 functional groups are shown in Figure 3e.



**Table 1.** Ligand–protein interaction profiles of the selected CYP3A4 inhibitors within the CYP3A4 binding site.

CYP3A4 Inhibitors	Autodock Energy (kcal/mol)	van der Waals Interactions Before MD	Hydrogen Bond Interactions in the Docked Poses			van der Waals Interactions in Centroid Structure from Clustering	Hydrogen Bond Interactions in the Centroid Structures			van der Waals Interactions after MD	Hydrogen Bond Interactions at 50 ns		
			Acceptor Atom	Donor Atom	Distance Å		Acceptor Atom	Donor Atom	Distance Å		Acceptor Atom	Donor Atom	Distance Å
CYP3A4-YK1	−9.9	Leu482, Thr309, Phe304, Asp76, Thr224, Arg372, Leu373, Arg375	Lig: Sulphonyl=O3 Lig: Carbonyl=O10	Arg212-NH1 Arg105-NE	3.01 2.93	Phe220, Gly109, Pro107, Leu373, Tyr53, Thr224, Glu374, Gly481, Ala370, Leu482, Arg212	Arg372=O Lig: Carbonyl=O3 Lig: Carbonyl=O3	Lig: Hydroxyl-O2 Arg105-NE Glu374-NH	2.62 3.68 3.70	Tyr53, Arg105, Arg106, Pro107, Gly109, Phe213, Leu216, Phe220, Ala370, Leu373, Gly481, Leu482	Arg372=O Lig: Carbonyl=O3	Lig: Hydroxyl-O2 Glu374-NH	2.58 3.05
CYP3A4-YK2	−11.6	Phe316, Glu308, Gln484, Thr309, Phe304, Phe108, Arg105, Phe213, Glu374, Gly481, Leu482, Leu483	Ser312-OG Ala370=O Arg372=O	Lig: Hydroxyl-O Lig: Hydroxyl-O2 Lig: Hydroxyl-O2	3.01 2.81 3.13	Ser312, Glu308, Arg212, Thr309, Phe304, Phe108, Arg105, Phe57, Glu374, Leu373, Leu483, Gln484	Arg372=O	Lig: Hydroxyl-O2	2.71	Phe57, Arg105, Arg212, Phe304, Thr309, Ser312, Leu373, Glu374, Leu462, Gln484	Arg372=O	Lig: Hydroxyl-O2	2.96
CYP3A4-YK3	−9.5	Phe57, Arg105, Arg106, Phe108, Ala305, Phe304, Thr309, Met371	Lig: Sulphonyl=O2 Lig: Carbonyl=O1 Lig: Carbonyl=O1 Lig: Carbonyl=O1	Arg212-NH1 - Ser119-OG Arg212-NH1 Arg212-NH2	3.05 3.53 3.18 3.02	Ile50, Tyr53, Asp76, Leu216, Leu221, Thr224, Val225, Gly481, Leu482, Arg212, Ala370, Leu373	Lig: Sulphonyl=O1 Lig: Hydroxyl-O3 Lig: N2	Arg106:NE Arg105:NH1 Arg372:NH	3.11 3.20 3.50	Asp76, Ile47, Arg105, Phe215, Leu216, Phe220, Leu221, Thr224, Ala370, Leu373, Glu374, Leu482	No hydrogen bonds formed	No hydrogen bonds formed	–
CYP3A4-YK4	−11.7	Phe108, Thr309, Phe304, Glu308, Ser312, Phe316, Leu373, Met371, Leu482,	Ser312-OG Leu483=O	Lig: Amine-N5 Lig: Amine-N5	3.16 3.12	Ser312, Gln484, Leu482, Thr309, Asp214, Phe316, Leu483, Pro485, Pro368, Met371, Ala370, Phe108, Ser119, Arg105	Phe213=O Arg212-NE Arg212-NH1	Lig: Amine-N2 Lig: Amine-N3 Lig: Amine-N3	3.22 3.64 3.66	Phe57, Arg106, Phe215, Phe241, Ile301, Pro368, Ala370, Glu374, Gly480, Gly481, Leu482	Ser119-O	Lig: Hydroxyl-O1	2.96
CYP3A4-YK5	−10.4	Arg105, Arg106, Ser119, Phe241, Ile301, Phe215,	Glu374-OE2 Lig: Carbonyl=O3	Lig: Hydroxyl-O Arg212-NH1	3.32 3.13	Phe316, Ile369, Leu483, Met371, Arg372, Phe215, Glu374, Ser312, Gln484, Glu308,	Lig: Carbonyl=O Ala349=O Lig: Carbonyl=O1	Leu483-NH Lig: Amine-N2 Arg212:NH1	2.82 3.38 2.91	Arg105, Ser119, Phe304, Gly306, Glu308, Ser312, Phe316, Ile369, Ala370, Met371, Gln484, Pro485	Thr309-OG1 Lig: Carbonyl=O	Lig: Hydroxyl-O3 Leu483-N	2.77 3.00

The terms MD and Lig are used as an abbreviation for Molecular Dynamics and Ligand respectively in Table 1.



**Figure 3.** The time dependent distance analysis of (a) CYP3A4-YK1; (b) CYP3A4-YK2; (c) CYP3A4-YK3; (d) CYP3A4-YK4; and (e) CYP3A4-YK5 residues and inhibitor groups involved in hydrogen bonding. The distances are measured in Angstrom ( $\text{\AA}$ ) and the time is shown along the x-axis.

Generally, the MM(GB/PB)SA method is used for the calculation of binding free energy by selecting snapshots at regular intervals over an entire MD simulation and it does not take into account the protein–water and ligand–water interaction details as it uses an implicit water model. The WaterSwap method uses explicit water molecules [76], so does not suffer from the same limitations as the MM(GB/PB)SA method. The WaterSwap method has been applied successfully to other systems [77–84]. Herein, we apply the WaterSwap method to estimate the binding free energies for the five highly potent inhibitors of CYP3A4. Overall, the WaterSwap calculations for the five selected complexes gave binding energies in the range  $-46.7 (\pm 0.5)$  to  $-30.9 (\pm 0.6)$  kcal/mol as shown in Table 2, indicating that the inhibitors all bind strongly to CYP3A4. The free energy values calculated through the BAR, FEP and TI methods differ by  $\leq 1$  kcal/mol indicating good convergence of the calculations. Two values are reported for YK4 and YK5 as the cluster analysis showed that structures belonging to the dominant cluster did not exist across the entire trajectory, so the centroid of the second dominant cluster was also included in the analysis. The binding energies predicted for YK4 are  $-39.7 (\pm 0.5)$  kcal/mol and  $-30.9 (\pm 0.6)$  kcal/mol for the centroid structures of the first two dominant clusters from the MD simulations and for YK5 the corresponding values are  $-36.0 (\pm 0.4)$  kcal/mol and  $-40 (\pm 0.3)$  kcal/mol, respectively. For YK4, a decrease of  $\sim 8.8$  kcal/mol was observed in the calculated binding energy for the second centroid structure because the residues (Arg106, Arg212, Arg372 and Gln434) previously favoring the stabilization of the inhibitor–protein complex of the first centroid structure now contribute significantly towards the stabilization of the water cluster as shown in Figure 5. However, for YK5 a higher binding energy for the second centroid structure might be due the stabilization of the inhibitor–protein complex by residues including Arg106, Arg212, Glu308, Thr309 and Gly481 that were contributing negatively towards inhibitor–protein stabilization in the previous calculation (Figure 5). There is a significant difference in energy for the two starting geometries, showing that even though there will be some configurational sampling through the Monte Carlo simulations, care should be taken in the selection of a starting geometry for the WaterSwap calculations.

**Table 2.** Binding free energies of the selected inhibitor-bound complexes (CYP3A4-YK1 to CYP3A4-YK5) from Monte Carlo calculations in WaterSwap, MM/PBSA and MM/GBSA. Note that two different WaterSwap calculations (using centroid clusters c0 and c1) were performed for YK4 and YK5 as the dominant cluster did not include structures from the entire simulation.

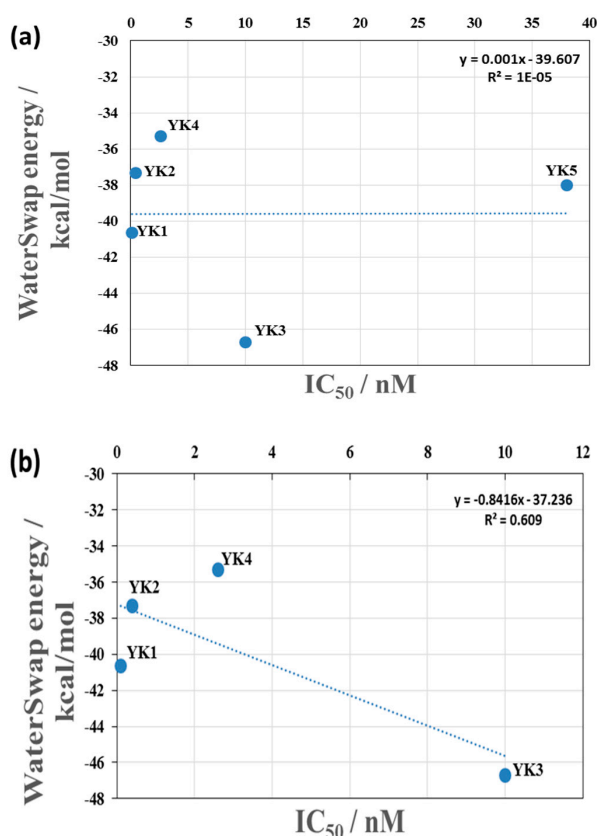
Inhibitor-Bound Complex	IC <sub>50</sub> nM	Autodock Score kcal/mol	MM/PBSA kcal/mol	MM/GBSA kcal/mol	WaterSwap			
					BAR kcal/mol	FEP kcal/mol	TI kcal/mol	Average kcal/mol
CYP3A4-YK1	0.1	-9.9	$-40.78 \pm 0.43$	$-61.22 \pm 0.43$	-41.1	-40.2	-40.6	$-40.6 \pm 0.5$
CYP3A4-YK2	0.4	-11.6	$-25.50 \pm 0.33$	$-37.44 \pm 0.24$	-37.5	-36.8	-37.7	$-37.3 \pm 0.5$
CYP3A4-YK3	10	-9.5	$-32.37 \pm 0.31$	$-49.48 \pm 0.31$	-47.3	-46.3	-46.5	$-46.7 \pm 0.5$
CYP3A4-YK4	2.6	-11.7	$-30.87 \pm 0.31$	$-34.96 \pm 0.23$	-40.3 -31.6	-39.5 -30.6	-39.3 -30.5	$-39.7 \pm 0.5$ $-30.9 \pm 0.6$
CYP3A4-YK5	38	-10.4	$-22.52 \pm 0.34$	$-38.46 \pm 0.25$	-36.1 -40.4	-36.3 -39.8	-35.6 -39.9	$-36.0 \pm 0.4$ $-40.0 \pm 0.3$

Furthermore, binding energy calculations were performed using the Molecular Mechanics/Generalized Born Surface Area (MM/GBSA) and Molecular Mechanics/Poisson–Boltzmann (MM/PBSA) methods, to analyze the interactions of the selected CYP3A4 inhibitors with the binding site residues (by taking into account both entropic and enthalpic contributions) and for the comparative analysis of the binding energies calculated through the WaterSwap method [41,42]. The binding energy calculations through these methods are reproducible and computationally inexpensive with a relatively good accuracy [85]. The calculated binding free energies extracted from 100 snapshots of the MD trajectories were averaged. The binding energy values calculated by the MM/PBSA and MM/GBSA methods were predicted in the range  $-40.78$  to  $-22.52$  kcal/mol and  $-61.22$  to  $-34.96$  kcal/mol, respectively, as shown in Table 2. The important energy components of these calculations are shown in Supplementary Materials Table S2. Although the difference in solvation and gas phase Gibbs free energy is similar between the

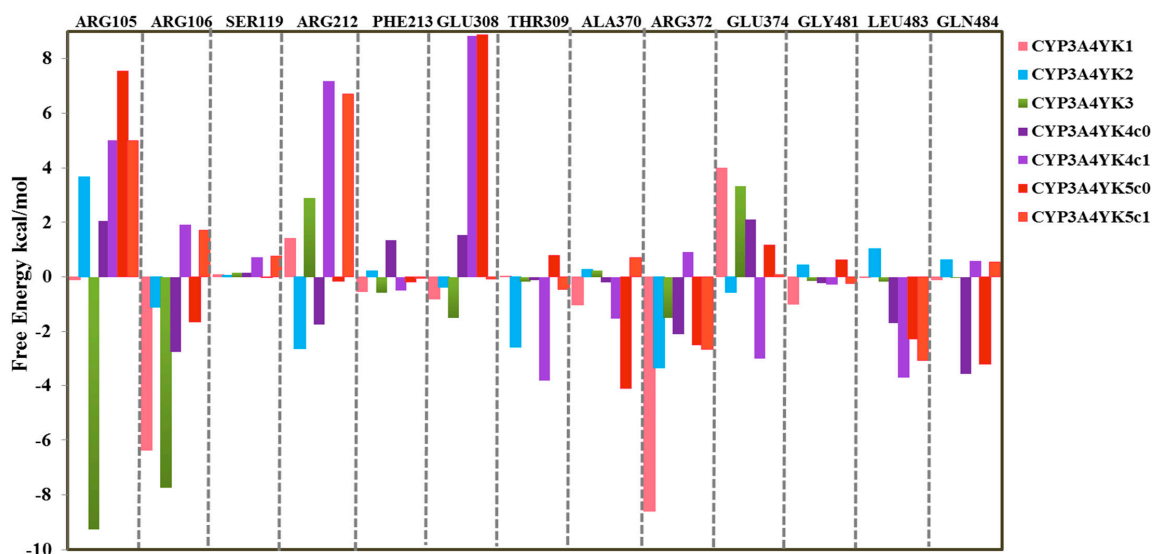
two methods but the comparison of binding energy predictions for the selected CYP3A4-inhibitor complexes indicates that higher binding energies were obtained by MM/GBSA method which might be due to the fact that MM/PBSA has an additional dispersion free energy component for calculation of binding energy and due to the difference in  $\Delta E_{\text{EGB/EPB}}$  and  $\Delta G_{\text{solv}}$  values in both methods. Overall, for all selected complexes the MM/PBSA and MM/GBSA calculations show that the van der Waals and electrostatic interactions had the highest contribution towards the binding energy as shown in Table S2.

For the selected CYP3A4-inhibitor complexes docking scores show that, this is a relatively small range of energies compared to the order of magnitude difference in the  $\text{IC}_{50}$  values of the inhibitors. Figure 4a shows a plot of the binding free energy (calculated through WaterSwap) vs  $\text{IC}_{50}$  value for the complexes, where it can be seen that there is no correlation with the experimental data. Figure 4b shows the same plot but with the data for YK5 removed as, with an  $\text{IC}_{50}$  value of 38 nM, it lies far outside the values of the other inhibitors. The correlation between the binding free energy and experimental  $\text{IC}_{50}$  value is significantly improved with a correlation coefficient of  $R^2 = 0.6$  (when an average of the two values are used for YK4). Correlation of the docking scores with the experimental data is also improved by the omission of the YK5 data, but the correlation remains weak ( $R^2 = 0.31$ ) as shown in Figure S10b. Overall, comparison of the binding energies obtained using MM/GBSA, MM/PBSA and WaterSwap methods shows that the calculated values do not differ significantly. However, MM/GBSA and MM/PBSA based binding energy values show little or no correlation with the  $\text{IC}_{50}$  values of the selected compounds (as shown in Figure S11a,b). The removal of compound YK5, a significant outlier in terms of activity, did not improve the  $R^2$  values for both methods. Therefore, WaterSwap provides a good estimate of the general trends for these compounds and is better than using the docking score alone; however, it is not very good at accounting for the differences in binding for structurally quite similar compounds. YK1 is a better inhibitor than YK3 by an order of magnitude in terms of  $\text{IC}_{50}$  value, but YK3 has a higher calculated binding free energy than YK1. A larger test set of ligands is needed to fully benchmark the ranking of ligands in terms of binding free energy.

Raza et al. have shown that in comparison to the MM/PBSA analysis, the WaterSwap method identifies more residues with significant contribution to the calculated binding energies [86]. Additional insight can be obtained from the WaterSwap calculations as the binding free energy can be decomposed in terms of individual residue contributions. It is important to emphasize that the interaction energies calculated through this method are not directly comparable to experimental results (e.g.,  $\Delta\Delta G$  for mutated enzymes) nor will they sum to the total binding energy; however, this type of decomposition analysis might be useful for the identification of residues that have the largest effects on the binding energy. The residue decomposition can be visualized and analyzed using the CHEWD plugin for Chimera or PyMol [86]. The residue decomposition indicates favorable and unfavorable interactions, information that is useful for lead development. Here, the WaterSwap residue-wise binding energy decompositions were analyzed for the residues with greatest contributions to the inhibitor binding throughout the course of MD simulations. These include 13 residues (Arg105, Arg106, Ser119, Arg212, Phe213, Glu308, Thr309, Ala370, Arg372, Glu374, Gly481, Leu483, and Gln484) for which the total  $G_{\text{residue}}$  components are shown in Figure 5. For YK1–5 the binding energy calculations based on the first centroid clusters show negative free energies of Arg106 and Arg372 for all five selected CYP3A4 inhibitors, favoring the stabilization of the inhibitor–protein complexes through van der Waals and hydrogen bond interactions. Whereas, the binding energy calculations for YK4 based on the second cluster show positive free energies for both Arg106 and Arg372 and only Arg106 for YK5 (Figure 5).



**Figure 4.** (a) The correlation between the binding free energy of the complex predicted by WaterSwap and the  $IC_{50}$  values for YK1–5. (b) The correlation between the binding free energy of the complex predicted by WaterSwap and the  $IC_{50}$  values with the data for YK5 removed as it is a significant outlier compared to the other compounds. Note that the values for YK4 and YK5 are the average of WaterSwap calculations using two different starting structures to ensure that the structures were representative of the entire MD trajectory.



**Figure 5.** The decomposition of the WaterSwap binding free energy into residue-wise components. The values for the selected CYP3A4 inhibitors are shown, where negative values indicate that the residue stabilizes the inhibitor–protein complex and positive values indicates stabilization of the water cluster rather than the inhibitor.

YK1 has strong interactions with both arginines 106 and 372, whereas YK3 has a strong interaction with Arg106 and a weaker interaction with Arg372 but has an additional very favorable interaction with Arg105 not seen in the other complexes. This additional stabilization from Arg105 may be the reason why YK3 has a greater binding free energy than YK1. The residue contribution for Ser119 is very small in all complexes, weakly positive for complexes of CYP3A4 with YK1–4, showing a slight preference for the water rather than the ligand, and weakly negative for YK5 (first cluster c0). The interaction between the ligand and Glu374 is unfavorable in all complexes except CYP3A4-YK2 and CYP3A4-YK4 (second cluster c1), where it is slightly favorable due its involvement in van der Waals interactions. For both structurally similar inhibitors YK2 and YK4, the binding energy calculations based on the first cluster show favorable interactions with Arg212 which are not present in complexes of YK1 and YK3. For YK4 the calculated values using both clusters also show stabilizing interactions with Leu483 whereas, Gln484 favors the stabilization of the inhibitor–protein complex for the first cluster only. YK5 has the smallest binding free energy of all the ligands studied when using the first cluster for calculation, and the residue-wise decomposition shows that it does not have as strong favorable interactions with the arginine residues in the binding site and that the most significant residues for binding in this complex are Ala370 and Gln484. Overall, Arg106 and Arg372 provide the most significant contribution to the binding of the inhibitors studied here.

The availability of CYP crystallographic structures and the use of various in silico approaches including docking, molecular dynamic simulations and the MM(GB/PB)SA method for the investigation of ligand binding have been described extensively in literature to further supplement the understanding of the underlying mechanism of ligand recognition by the CYP isoforms [36,87–92]. The MM(GB/PB)SA method is commonly used for the calculation of binding free energies based on the MD simulation trajectories [76]. The MM(GB/PB)SA analysis and per residue energy decompositions for a set of CYP3A4 substrates used clinically in combinatorial cancer therapy (cytarabine, daunorubicin, doxorubicin and vincristine) has shown the importance of Asp61, Asp76, Ala174, Glu122, Asp214, Arg105, Arg212, Ala217, Glu234, Ala305, Glu308, Ala370, Ile369 and Glu374 residues for favorable interactions with ligands [36]. Herein, the WaterSwap results were analyzed further to find the contribution of 13 important binding site residues to the binding of the inhibitors. Arg106, Arg212, Phe213, Glu308, Ala370, Arg372, Gly481, Leu483 and Gln484 play a major role in the stabilization of the selected CYP3A4-inhibitor complexes. Interactions with Arg105, Ser119, Thr309 and Glu374 are unfavorable as they show a preference for the water cluster rather than the inhibitor. This information could be useful to help guide the development of these lead molecules to enhance binding. Principally, the WaterSwap method is fully automated and robust, providing new tools for the analysis and visualization of the important driving forces in binding of protein–ligand complexes and is broadly applicable to study binding or assess the impact of mutations on binding [76,77,82,83].

### 3. Materials and Methods

#### 3.1. Ligand Selection

For the identification of highly potent inhibitors of CYP3A4, a large dataset of CYP3A4 inhibitors was extracted from the ChEMBL database (<https://www.ebi.ac.uk/chembl/>) [93] using a filtering criteria of  $IC_{50} \leq 100 \mu\text{M}$ . After removal of duplicates and fragments the remaining data set of 2747 compounds was subjected to LipE- and LE calculations [22] as described by Hopkins et al. [49]. A brief detail of LipE and LE calculations of the CYP3A4 inhibitors has been provided in the Supplementary Materials (SM) (Section S1, Figure S1a,b and Table S1). Finally, five CYP3A4 inhibitors showing optimal LipE and LE values were selected for further MD studies to estimate the binding free energy and stability of the inhibitor-bound complexes.

### 3.2. Molecular Docking Studies

CYP3A4 inhibitors were built in GaussView 5 [94] and their geometry optimized using density functional theory with the B3LYP functional and the 6–31G(d) and 6–31 + G(2d,p) basis sets. The 2.05 Å resolution crystal structure of human CYP3A4 1TQN [44] was retrieved from the protein data bank to provide the starting point for docking and MD simulations. The missing 1TQN H-I loop residues were built as described by Webb and Sali [95] using Modeller version 9.15 [48]. The modeled structure of CYP3A4 (1TQN) [44] was prepared using AutoDockTools [63] where polar hydrogens were added after computing the Kollman partial charges and solvent parameters for the macromolecule. After protein modification, non-polar and polar hydrogens were merged, Gasteiger charges were computed for optimized ligands and the flexible torsions were defined.

The CYP3A4 binding site is large and highly promiscuous allowing multiple ligands to bind simultaneously [96]. Additionally, for CYP3A4 two different binding cavity volumes (1,386 Å<sup>3</sup> and 520 Å<sup>3</sup>) have been reported previously [44,71]. Therefore, in order to provide the appropriate docking area and dimensions of the CYP3A4 binding site, the Grid Box pointer was centered above the heme at grid points x: –20, y: –21 and, z: –11 with a grid size of 60 × 60 × 60 Å and a grid spacing of 0.375 Å. The conformational space of ligands was explored using a Lamarckian genetic algorithm (LGA) that is based on a local search method for the energy minimization [97]. AutoDock evaluates conformations during docking simulations using a semi-empirical free energy force field that accounts for the intramolecular and desolvation terms along with directionality in hydrogen bonds [98]. The free energy scoring function involves two steps; firstly, the evaluation of the intramolecular energetics for the transition from the unbound state to the bound conformation for each molecule separately (Equation (1)), followed by the intermolecular energetics evaluation of bringing the two molecules together into the bound complex (Equation (2)) [99].

$$\Delta G = \left( V_{Bound}^{L-L} - V_{Unbound}^{L-L} \right) + \left( V_{Bound}^{P-P} - V_{Unbound}^{P-P} \right) + \left( V_{Bound}^{P-L} \right) - V_{Unbound}^{P-L} + \Delta S_{conf} \quad (1)$$

Here,  $V$  indicates the six pairwise evaluations,  $\Delta S_{conf}$  gives an estimate of the conformational entropy lost upon binding,  $L$  and  $P$  refer to the ligand and protein in a protein–ligand complex. Equation (2) includes a pairwise atomic term in the evaluation of dispersion/repulsion, hydrogen bonding, electrostatics, and desolvation:

$$V = W_{vdw} \sum_{i,j} \left( \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \right) + W_{hbond} \sum_{i,j} E(t) \left( \frac{C_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r_{ij}^{10}} \right) + W_{elec} \sum_{i,j} \frac{q_i q_j}{\epsilon(r_{ij}) r_{ij}} + W_{sol} \sum_{i,j} \left( S_i V_j + S_i V_j \right) e^{-\frac{r_{ij}^2}{2\sigma^2}} \quad (2)$$

A maximum of 10 different conformations were generated for each ligand during the docking run and the minimum energy pose was selected for further investigation. The CYP3A4-inhibitor interaction profiles were analyzed to assess whether the binding conformations agree with available data reported for other CYP3A4 inhibitors. The stability of the selected CYP3A4-inhibitor docked complexes was evaluated using MD simulations.

### 3.3. Molecular Dynamic Simulations using Graphical Processing Units (GPUs)

The unliganded 1TQN structure and the finally selected inhibitor-bound complexes were prepared using the LEaP module in AMBER14 [65]. All protein complexes were solvated using TIP3PBOX [99] with a margin of 12.0 Å and an appropriate number of Cl<sup>–</sup> ions were added to neutralize the overall charge on the system. The GAFF force field [100] was used to parameterize the ligands using the ANTECHAMBER program with AM1-BCC charge assignment method [65], while the proteins were modeled using AMBER force field FF14SB [101]. The topology and coordinate files for the entire system

were generated by the LEaP module of AMBER 14 [65]. Heme parameters were adopted from an earlier study [102] and the bond between CYS442 SG atom and Fe atom of heme was explicitly defined.

MD simulations of the selected conformers were performed after optimizing the positions of hydrogen atoms, water molecules and all atoms within each complex by energy minimizations through the SANDER module in AMBER14 [65]. MD simulations were performed using a 2 fs time step with a cutoff radius of 8.0 Å for nonbonded interactions. The particle-mesh Ewald method [103] was used for the calculation of long-range electrostatic interactions and all bonds containing hydrogen were constrained using the SHAKE algorithm [104]. Heating was performed by gradually increasing the temperature from 0 to 300 K over a period of 20 ps. Langevin temperature control with a random number assignment to each trajectory (random initial velocities) was used. A 100 ps equilibration was carried out at a temperature of 300 K with 1 atm pressure which was followed by 140 ps under the same temperature and pressure conditions. A 50 ns MD simulation was performed for all the complexes using PMEMD.CUDA from AMBER14 [65]. Trajectory analysis was performed using the CPPTRAJ module in AMBER14 [65]. The stability of each CYP3A4-inhibitor complex was assessed by evaluating the  $C_{\alpha}$ RMSD,  $C_{\alpha}$ RMSF, Rg and hydrogen bond analysis with respect to the starting structures.

### 3.4. Binding Energy Predictions Using WaterSwap

The WaterSwap method from the Sire package [40] was used for the prediction of absolute binding free energies for the CYP3A4-inhibitor docked complexes. The conformational landscapes of the CYP3A4-inhibitor complexes obtained after the molecular dynamics simulations were clustered using CPPTRAJ [65] based on the hierarchical agglomerative approach using average-linkage [105]. Clustering was performed based on the RMSD of important active site residues including Tyr53, Phe57, Asp76, Arg105, Arg106, Phe108, Ser119, Arg212, Phe215, Ile301, Phe304, Ala305, Thr309, Ile369, Ala370, Met371, Arg372, Glu374, Leu482 and Leu483 with no fit. For each complex the centroid of these clusters was used as the starting structure for WaterSwap calculations [40].

The WaterSwap method calculates the absolute protein–ligand binding free energy using an explicit water model by essentially swapping the ligand dimensions with an equal volume of water within the binding site [40]. The method takes into account the protein–water, ligand–water and protein–water–ligand interactions thus, removing the double decoupling cavitation problems existing in the implicit solvent methods for the estimation of binding free energies [76]. The WaterSwap method uses the Monte Carlo simulations [106] for attaining the WaterSwap reaction coordinates (WSRC) representative of the ligand swapping with water and also for the calculation of binding free energies [40]. The WSRC is augmented with the identity constraint [107] that identifies a cluster of water molecules equivalent to the volume and shape of the ligand within the protein binding site which is achieved by coupling two periodic boxes including the protein box (protein–ligand complex solvated in water box) and the bulk-water-box (box of water centered on the ligand coordinates) to the same thermostat and barostat [40], followed by the subsequent swapping of the two systems using a dual topological approach [108,109]. Equation (5) is used for the calculation of binding energy of all systems:

$$E(\lambda) = E_{proteinbox} + E_{waterbox} + E_{ligand} + E_{cluster} + (1 - \lambda)(E_{ligand:proteinbox} + E_{cluster:waterbox}) + (\lambda)(E_{cluster:proteinbox} + E_{ligand:waterbox}) \quad (3)$$

where,  $E_{proteinbox}$  is the energy of all molecules within the protein excluding the ligand,  $E_{waterbox}$  is the energy of all molecules in the water box excluding the water cluster,  $E_{ligand}$  denotes the intramolecular energy of the ligand,  $E_{cluster}$  represents the interaction energy between all the water molecules present within the water cluster,  $E_{ligand:proteinbox}$  shows the interaction energy between all atoms of the protein box and the ligand, and  $E_{cluster:waterbox}$  is the interaction energy between all water molecules in bulk water box and the identified water cluster.

The reaction involves simultaneous decoupling of the ligand from the protein box and the water cluster from the water box followed by parallel coupling of the ligand to the water box and the water



cluster to the protein box.  $\lambda$  is the WaterSwap reaction coordinate that transforms from  $\lambda = 0$  to  $\lambda = 1$  to assist the transfer of the ligand from the protein bound state (at  $\lambda = 0$ ) to the water box (at  $\lambda = 1$ ). Additionally, the transformation also corresponds to the transfer of water cluster from water box to fill the resultant cavity in the protein box. For a WaterSwap calculation of a system, the absolute binding free energy is estimated using replica exchange thermodynamic integration method (RETI) [77,110,111] that takes into account replica exchange moves along the WSRC on the basis of calculated energy gradients using finite-difference TI (FDTI) with respect to  $\lambda$  (Equation (5)). However, to calculate the free energy gradients along  $\lambda$ , the collective averages are computed using Equation (5).

$$dE/d\lambda = (E_{cluster:proteinbox} + E_{ligand:waterbox}) - (E_{ligand:waterbox} + E_{cluster:waterbox}) \quad (4)$$

$$dG/d\lambda_{\lambda} = dE/d\lambda_{\lambda} \quad (5)$$

Monte Carlo simulations [110] are used to accept or reject neighboring replicas periodically during the replica exchange, thus allowing full sampling of ligand-bound and ligand-free states during the simulation of the connected protein box/bulk-water-box. Moreover, the quality of the WaterSwap estimations is also dependent on the number of Monte Carlo simulations used to calculate the energy gradients across  $\lambda$ . The potential of mean force (PMF) along the WSRC was obtained by the integration of the energy gradients along  $\lambda$  using Equation (7) [40].

$$G_{Bind} = -\int_0^1 DG/d\lambda_{\lambda} d\lambda \quad (6)$$

The ligand unbinding from the protein is represented as the negative integral in WSRC. The absolute binding free energy in WaterSwap is based on a single reaction coordinate which is approximated by averaging the total energy gradient along  $\lambda$  to achieve free energy gradient followed by the integration over the entire WSRC. Since the WaterSwap method for calculating binding free energy is based on averages of the gradients of total WaterSwap energies across the sampled configurations, it is therefore also reasonable to average the gradients of the total energy components ( $G_{proteinbox}$ ,  $G_{waterbox}$ ) with respect to  $\lambda$ . These decompositions are approximations and as a result,  $G_{proteinbox}$  and  $G_{waterbox}$  will not be exactly equal to  $G_{bind}$ . Therefore, the decomposition of components would assist in revealing whether a favorable binding free energy is due to ligand showing a strong affinity for protein or a low affinity for water [76].

Finally, in this study the most representative cluster for each complex was selected for further analysis, where the ligands were swapped with the representative water clusters after the construction of WaterSwap coordinates for the calculation of the PMF using the RETI method. For all CYP3A4-inhibitor-bound complexes WaterSwap calculations were performed for 1000 iterations using 25 million moves of Monte Carlo (MC) sampling across each of the 16  $\lambda$ s (0.005, 0.071, 0.203, 0.269, 0.335, 0.401, 0.467, 0.533, 0.599, 0.665, 0.731, 0.797, 0.863, 0.929, 0.995) at a temperature of 298.15 K and pressure of 1 atm on each replica. For the WaterSwap calculations the MC across the 16  $\lambda$  windows have been used since it is already reported that sixteen MC simulations are adequate, generating sixteen free energy gradients spaced across  $\lambda$  [76]. The free energy of binding was evaluated using three different statistical techniques: free energy perturbation (FEP), thermodynamic integration (TI) and the Bennett's acceptance ratio (BAR) algorithm. Furthermore, the degree of convergence of the calculated results can be assessed by considering the level of agreement between these three energy values, where ideally the deviation should be within 1 kcal/mol [76].

Furthermore, the application of MM/PBSA and MM/GBSA methods to assure the accurate ranking of inhibitors on the basis of their binding energy can add value to the drug design research. The binding energies of the five CYP3A4 inhibitor docked complexes were estimated through (MM/PBSA) and (MM/GBSA) method using Equations (7) and (8) [41,42].

$$\Delta G_{binding} = G_{complex} - G_{protein} - G_{ligand} = \Delta E_{MM} + \Delta G_{PB} + \Delta G_{nonpolar} - T\Delta S \quad (7)$$

$$\Delta\Delta G_{binding} = \Delta E_{MM} - \Delta G_{solv} + \Delta G_{SA} \quad (8)$$

Herein, the  $\Delta E_{MM}$  accounts for the difference between the minimized energies of the CYP3A4 inhibitor-bound complexes and the total energies of the enzyme and inhibitor (including electrostatic and the van der Waals energies),  $T\Delta S$  shows the change in entropy of the ligand binding conformations,  $G_{solv}$  is the summation of contributions from the polar states and  $\Delta G_{SA}$  parameter accounts for the difference in the surface area energies and is estimated from the solvent accessible surface area (SASA) (using a water probe radius of 1.4 Å).

#### 4. Conclusions

The CYP3A4 binding site is highly promiscuous in nature and therefore, it is more prone to drug–drug interactions due to the inhibition or induction of the metabolic enzyme. The interaction of CYP3A4 with a broad spectrum of chemical entities leads to greater chances of inhibition. Thus, highly potent inhibitors were selected to elucidate the mechanism of CYP3A4 inhibition and extend the understanding of CYP3A4 mediated drug–drug interactions. The current study provides an overview of the most probable binding poses and energy values of CYP3A4 inhibitors which may further help to address the primary question of CYP3A4 inhibition. Ultimately, the probable binding modes and interaction profiles of the selected CYP3A4 inhibitors were explored by docking and MD simulations. Binding site residues including Phe57, Arg105, Arg106, Ser119, Arg212, Phe213, Thr309, Ser312, Ala370, Arg372, Glu374, Gly481 and Leu483 were observed to be significant in terms of interactions with CYP3A4. The binding energies were calculated within considerable range using MM/GBSA, MM/PBSA and the WaterSwap methods. The WaterSwap calculations demonstrate that the highly potent inhibitors show an overall greater affinity towards CYP3A4 and the binding site residues Arg106 and Arg372 favor the stabilization of the all CYP3A4-inhibitor complexes. The ultimate goal of this study and a follow up study is to design a pipeline for CYP3A4 inhibition to be used in the lead optimization programs. However, herein, we present ligand binding and the estimation of free energy of binding with CYP3A4. The estimation of free energy could pave the way towards understanding the strength of CYP inhibition. In a follow up study the ligand–protein complexes with minimum binding free energies could be used as a reference/template for building predictive models for CYP3A4 inhibition that may have advantage over the already existing predictive models due to the use of more appropriate templates based on binding free energy. This study provides an understanding of CYP3A4 inhibition process and our results could guide the use of multiple approaches, including simulations, relevant to CYP3A4 inhibition for the virtual screening of new chemical entities during lead optimization programs, to provide basis for early screening of CYP inhibitors ultimately leading to the design/selection of new chemical entities with suitable ADME-Tox properties and reduced side effects.

**Supplementary Materials:** Supplementary materials can be found at <http://www.mdpi.com/1422-0067/20/18/4468/s1>.

**Author Contributions:** Y.S.K.; K.E.R., I.J. and A.J.M. conceived and designed the project. Y.S.K.; K.E.R., I.J. and A.J.M. carried out all the work, analyzed the results and wrote the paper.

**Funding:** This research received no external funding.

**Acknowledgments:** We would acknowledge the Higher Education Commission (HEC) Pakistan, for their support and assistance. We would also thank the Advanced Computing Research Center, University of Bristol (<http://www.bris.ac.uk/acrc>) for providing the computational facilities and easy access to BlueCrystal Phase 3 cluster. A.J.M. thanks EPSRC for funding (grant number EP/M022609/1, CCP-BioSim, <http://www.ccpbiosim.ac.uk>) and K.E.R thanks BBSRC for funding (grant number BB/M000354/1).

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Cross, D.; Bayliss, M.M. A commentary on the use of hepatocytes in drug metabolism studies during drug discovery and development. *Drug Metab. Rev.* **2000**, *32*, 219–240. [[CrossRef](#)] [[PubMed](#)]
2. Davila, J.C.; Rodriguez, R.J.; Melchert, R.B.; Acosta, D.D., Jr. Predictive value of in vitro model systems in toxicology. *Annu. Rev. Pharmacol. Toxicol.* **1998**, *38*, 63–96. [[CrossRef](#)] [[PubMed](#)]
3. Forbes, B.; Ehrhardt, C.C. Human respiratory epithelial cell culture for drug delivery applications. *Eur. J. Pharm. Biopharm.* **2005**, *60*, 193–205. [[CrossRef](#)] [[PubMed](#)]
4. Ekins, S.; Ring, B.J.; Grace, J.; McRobie-Belle, D.J.; Wrighton, S.A.A. Present and future in vitro approaches for drug metabolism. *J. Pharmacol. Toxicol. Methods* **2000**, *44*, 313–324. [[CrossRef](#)]
5. Nebert, D.W.; Russell, D.W.W. Clinical importance of the cytochromes P450. *Lancet* **2002**, *360*, 1155–1162. [[CrossRef](#)]
6. Almira Correia, M.; Sinclair, P.R.; De Matteis, F.F. Cytochrome P450 regulation: The interplay between its heme and apoprotein moieties in synthesis, assembly, repair, and disposal. *Drug Metab. Rev.* **2011**, *43*, 1–26. [[CrossRef](#)]
7. Wrighton, S.A.; Schuetz, E.G.; Thummel, K.E.; Shen, D.D.; Korzekwa, K.R.; Watkins, P.B.B. The human CYP3A subfamily: Practical considerations. *Drug Metab. Rev.* **2000**, *32*, 339–361. [[CrossRef](#)]
8. Quintieri, L.; Palatini, P.; Nassi, A.; Ruzza, P.; Floreani, M.M. Flavonoids diosmetin and luteolin inhibit midazolam metabolism by human liver microsomes and recombinant CYP 3A4 and CYP3A5 enzymes. *Biochem. Pharmacol.* **2008**, *75*, 1426–1437. [[CrossRef](#)]
9. Dresser, G.K.; Spence, J.D.; Bailey, D.G.G. Pharmacokinetic-pharmacodynamic consequences and clinical relevance of cytochrome P450 3A4 inhibition. *Clin. Pharmacokinet.* **2000**, *38*, 41–57. [[CrossRef](#)]
10. Obach, R.S.; Walsky, R.L.; Venkatakrishnan, K.; Gaman, E.A.; Houston, J.B.; Tremaine, L.M.M. The utility of in vitro cytochrome P450 inhibition data in the prediction of drug–drug interactions. *J. Pharmacol. Exp. Ther.* **2006**, *316*, 336–348. [[CrossRef](#)]
11. Ohshima, K.; Nakajima, M.; Suzuki, M.; Shimada, N.; Yamazaki, H.; Yokoi, T.T. Inhibitory effects of amiodarone and its N-deethylated metabolite on human cytochrome P450 activities: Prediction of in vivo drug interactions. *Br. J. Clin. Pharmacol.* **2000**, *49*, 244–253. [[CrossRef](#)] [[PubMed](#)]
12. Wang, R.W.; Newton, D.J.; Liu, N.; Atkins, W.M.; Lu, A.Y.Y. Human cytochrome P-450 3A4: In vitro drug–drug interaction patterns are substrate-dependent. *Drug Metab. Dispos.* **2000**, *28*, 360–366. [[PubMed](#)]
13. De Groot, M.J.J. Designing better drugs: Predicting cytochrome P450 metabolism. *Drug Discov. Today* **2006**, *11*, 601–606. [[CrossRef](#)] [[PubMed](#)]
14. Zhou, S.-F.F. Drugs behave as substrates, inhibitors and inducers of human cytochrome P450 3A4. *Curr. Drug Metab.* **2008**, *9*, 310–322. [[CrossRef](#)] [[PubMed](#)]
15. Scripture, C.D.; Figg, W.D.D. Drug interactions in cancer therapy. *Nat. Rev. Cancer* **2006**, *6*, 546. [[CrossRef](#)] [[PubMed](#)]
16. Nettleton, D.O.; Einolf, H.J. Assessment of cytochrome p450 enzyme inhibition and inactivation in drug discovery and development. *Curr. Top. Med. Chem.* **2011**, *11*, 382–403. [[CrossRef](#)] [[PubMed](#)]
17. Sevrioukova, I.F.; Poulos, T.L.L. Understanding the mechanism of cytochrome P450 3A4: Recent advances and remaining problems. *Dalton Trans.* **2013**, *42*, 3116–3126. [[CrossRef](#)] [[PubMed](#)]
18. Martiny, V.Y.; Carbonell, P.; Chevillard, F.; Moroy, G.; Nicot, A.B.; Vayer, P.; Villoutreix, B.O.; Miteva, M.A.A. Integrated structure-and ligand-based in silico approach to predict inhibition of cytochrome P450 2D6. *Bioinformatics* **2015**, *31*, 3930–3937. [[PubMed](#)]
19. Bello, M.; Mendieta-Wejebe, J.E.; Correa-Basurto, J.J. Structural and energetic analysis to provide insight residues of CYP2C9, 2C11 and 2E1 involved in valproic acid dehydrogenation selectivity. *Biochem. Pharmacol.* **2014**, *90*, 145–158. [[CrossRef](#)] [[PubMed](#)]
20. Shao, C.Y.; Su, B.-H.; Tu, Y.S.; Lin, C.; Lin, O.A.; Tseng, Y.J.J. CypRules: A rule-based P450 inhibition prediction server. *Bioinformatics* **2015**, *31*, 1869–1871. [[CrossRef](#)] [[PubMed](#)]
21. Lee, J.; Basith, S.; Cui, M.; Kim, B.; Choi, S.S. In silico prediction of multiple-category classification model for cytochrome P450 inhibitors and non-inhibitors using machine-learning method. *SAR QSAR Environ. Res.* **2017**, *28*, 863–874. [[CrossRef](#)] [[PubMed](#)]

22. Kiani, Y.S.; Jabeen, I.I. Exploring the chemical space of cytochrome p450 inhibitors using integrated physicochemical parameters, drug efficiency metrics and decision tree models. *Computation* **2019**, *7*, 26. [[CrossRef](#)]
23. Nembri, S.; Grisoni, F.; Consonni, V.; Todeschini, R.R. In silico prediction of cytochrome P450-drug interaction: QSARs for CYP3A4 and CYP2C9. *Int. J. Mol. Sci.* **2016**, *17*, 914. [[CrossRef](#)] [[PubMed](#)]
24. Lewis, D.F.; Modi, S.; Dickins, M.M. Structure—Activity relationship for human cytochrome P450 substrates and inhibitors. *Drug Metab. Rev.* **2002**, *34*, 69–82. [[CrossRef](#)] [[PubMed](#)]
25. Gottlieb, A.; Stein, G.Y.; Oron, Y.; Ruppin, E.; Sharan, R. INDI: A computational framework for inferring drug interactions and their associated recommendations. *Mol. Syst. Biol.* **2012**, *8*, 1–12. [[CrossRef](#)] [[PubMed](#)]
26. Ekins, S.; Stresser, D.M.; Williams, J.A. In vitro and pharmacophore insights into CYP3A enzymes. *Trends Pharmacol. Sci.* **2003**, *24*, 161–166. [[CrossRef](#)]
27. Ekins, S.; Bravi, G.; Binkley, S.; Gillespie, J.S.; Ring, B.J.; Wikel, J.H.; Wrighton, S.A. Three- and four-dimensional quantitative structure activity relationship analyses of cytochrome P-450 3A4 inhibitors. *J. Pharmacol. Exp. Ther.* **1999**, *290*, 429–438. [[PubMed](#)]
28. Ekins, S.; Bravi, G.; Wikel, J.H.; Wrighton, S.A. Three-dimensional-quantitative structure activity relationship analysis of cytochrome P-450 3A4 substrates. *J. Pharmacol. Exp. Ther.* **1999**, *291*, 424–433. [[PubMed](#)]
29. Silva, F.C.; Varlamova, E.V.; Braga, R.C.; Andrade, C.H. Development of QSAR models for identification of CYP3A4 substrates and inhibitors. *Mol2Net* **2015**, *1*, 1–6.
30. Mao, B.; Gozalbes, R.; Barbosa, F.; Migeon, J.; Merrick, S.; Kamm, K.; Wong, E.; Costales, C.; Shi, W.; Wu, C. QSAR modeling of in vitro inhibition of cytochrome P450 3A4. *J. Chem. Inf. Model.* **2006**, *46*, 2125–2134. [[CrossRef](#)]
31. Kriegl, J.M.; Arnhold, T.; Beck, B.; Fox, T. A support vector machine approach to classify human cytochrome P450 3A4 inhibitors. *J. Comput. Aided Mol. Des.* **2005**, *19*, 189–201. [[CrossRef](#)] [[PubMed](#)]
32. Lionta, E.; Spyrou, G.; Vassilatis, D.K.; Cournia, Z. Structure-based virtual screening for drug discovery: Principles, applications and recent advances. *Curr. Top. Med. Chem.* **2014**, *14*, 1923–1938. [[CrossRef](#)] [[PubMed](#)]
33. Marechal, J.-D.; Yu, J.; Brown, S.; Kapelioukh, I.; Rankin, E.M.; Wolf, C.R.; Roberts, G.C.; Paine, M.J.; Sutcliffe, M.J. In silico and in vitro screening for inhibition of cytochrome P450 CYP3A4 by comedication commonly used by patients with cancer. *Drug Metab. Dispos.* **2006**, *34*, 534–538. [[CrossRef](#)] [[PubMed](#)]
34. Mukhtar, S.; Kiani, Y.S.; Jabeen, I. Molecular docking simulations and GRID-independent molecular descriptor (GRIND) analysis to probe stereoselective interactions of CYP3A4 inhibitors. *Med. Chem. Res.* **2017**, *26*, 2322–2335. [[CrossRef](#)]
35. Lonsdale, R.; Rouse, S.L.; Sansom, M.S.; Mulholland, A.J. A multiscale approach to modelling drug metabolism by membrane-bound cytochrome P450 enzymes. *PLoS Comput. Biol.* **2014**, *10*, e1003714. [[CrossRef](#)] [[PubMed](#)]
36. Panneerselvam, S.; Yesudhas, D.; Durai, P.; Anwar, M.; Gosu, V.; Choi, S. A combined molecular docking/dynamics approach to probe the binding mode of cancer drugs with cytochrome P450 3A4. *Molecules* **2015**, *20*, 14915–14935. [[CrossRef](#)] [[PubMed](#)]
37. Calabrò, G.; Woods, C.J.; Powlesland, F.; Mey, A.S.; Mulholland, A.J.; Michel, J. Elucidation of nonadditive effects in protein–ligand binding energies: Thrombin as a case study. *J. Phys. Chem. B* **2016**, *120*, 5340–5350. [[CrossRef](#)] [[PubMed](#)]
38. Zhao, H.; Caflisch, A. Molecular dynamics in drug design. *Eur. J. Med. Chem.* **2015**, *91*, 4–14. [[CrossRef](#)] [[PubMed](#)]
39. Ge, Y.; van der Kamp, M.; Malaisree, M.; Liu, D.; Liu, Y.; Mulholland, A.J. Identification of the quinolinedione inhibitor binding site in Cdc25 phosphatase B through docking and molecular dynamics simulations. *J. Comput. Aided Mol. Des.* **2017**, *31*, 995–1007. [[CrossRef](#)] [[PubMed](#)]
40. Woods, C.J.; Malaisree, M.; Hannongbua, S.; Mulholland, A.J. A water-swap reaction coordinate for the calculation of absolute protein–ligand binding free energies. *J. Chem. Phys.* **2011**, *134*, 02B611. [[CrossRef](#)] [[PubMed](#)]
41. Kollman, P.A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W. Calculating structures and free energies of complex molecules: Combining molecular mechanics and continuum models. *Acc. Chem. Res.* **2000**, *33*, 889–897. [[CrossRef](#)] [[PubMed](#)]

42. Genheden, S.; Ryde, U. The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert Opin. Drug Discov.* **2015**, *10*, 449–461. [[CrossRef](#)] [[PubMed](#)]
43. Sevrioukova, I.F.; Poulos, T.L. Structure and mechanism of the complex between cytochrome P4503A4 and ritonavir. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 18422–18427. [[CrossRef](#)] [[PubMed](#)]
44. Yano, J.K.; Wester, M.R.; Schoch, G.A.; Griffin, K.J.; Stout, C.D.; Johnson, E.F. The structure of human microsomal cytochrome P450 3A4 determined by X-ray crystallography to 2.05-Å resolution. *J. Biol. Chem.* **2004**, *279*, 38091–38094. [[CrossRef](#)]
45. Sevrioukova, I.F.; Poulos, T.L. Structural and mechanistic insights into the interaction of cytochrome P4503A4 with bromoergocryptine, a type I ligand. *J. Biol. Chem.* **2012**, *287*, 3510–3517. [[CrossRef](#)] [[PubMed](#)]
46. Sevrioukova, I.F.; Poulos, T.L. Dissecting cytochrome P450 3A4–ligand interactions using ritonavir analogues. *Biochemistry* **2013**, *52*, 4474–4481. [[CrossRef](#)]
47. Shahrokh, K.; Cheatham, T.E., III; Yost, G.S. Conformational dynamics of CYP3A4 demonstrate the important role of Arg212 coupled with the opening of ingress, egress and solvent channels to dehydrogenation of 4-hydroxy-tamoxifen. *Biochim. Biophys. Acta BBA Gen. Subj.* **2012**, *1820*, 1605–1617. [[CrossRef](#)] [[PubMed](#)]
48. Šali, A.; Blundell, T.L. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **1993**, *234*, 779–815. [[CrossRef](#)]
49. Hopkins, A.L.; Keserü, G.M.; Leeson, P.D.; Rees, D.C.; Reynolds, C.H. The role of ligand efficiency metrics in drug discovery. *Nat. Rev. Drug Discov.* **2014**, *13*, 105–121. [[CrossRef](#)]
50. Navrátilová, V.; Paloncýová, M.T.; Berka, K.; Otyepka, M. Effect of lipid charge on membrane immersion of cytochrome P450 3A4. *J. Phys. Chem. B* **2016**, *120*, 11205–11213. [[CrossRef](#)]
51. Šrejber, M.; Navrátilová, V.; Paloncýová, M.; Bazgier, V.; Berka, K.; Anzenbacher, P.; Otyepka, M. Membrane-attached mammalian cytochromes P450: An overview of the membrane's effects on structure, drug binding, and interactions with redox partners. *J. Inorg. Biochem.* **2018**, *183*, 117–136. [[CrossRef](#)] [[PubMed](#)]
52. Berka, K.; Paloncýová, M.T.; Anzenbacher, P.; Otyepka, M. Behavior of human cytochromes P450 on lipid membranes. *J. Phys. Chem. B* **2013**, *117*, 11556–11564. [[CrossRef](#)] [[PubMed](#)]
53. Lonsdale, R.; Houghton, K.T.; Žurek, J.; Bathelt, C.M.; Foloppe, N.; de Groot, M.J.; Harvey, J.N.; Mulholland, A.J. Quantum mechanics/molecular mechanics modeling of regioselectivity of drug metabolism in cytochrome P450 2C9. *J. Am. Chem. Soc.* **2013**, *135*, 8001–8015. [[CrossRef](#)] [[PubMed](#)]
54. Yuki, H.; Honma, T.; Hata, M.; Hoshino, T. Prediction of sites of metabolism in a substrate molecule, instanced by carbamazepine oxidation by CYP3A4. *Bioorg. Med. Chem.* **2012**, *20*, 775–783. [[CrossRef](#)] [[PubMed](#)]
55. Magistrato, A.; Sgrignani, J.; Krause, R.; Cavalli, A. Single or multiple access channels to the CYP450s active site? An answer from free energy simulations of the human aromatase enzyme. *J. Phys. Chem. Lett.* **2017**, *8*, 2036–2042. [[CrossRef](#)] [[PubMed](#)]
56. Spinello, A.; Ritacco, I.; Magistrato, A. The catalytic mechanism of steroidogenic cytochromes P450 from all-atom simulations: Entwinement with membrane environment, redox partners, and post-transcriptional regulation. *Catalysts* **2019**, *9*, 81. [[CrossRef](#)]
57. Williams, P.A.; Cosme, J.; Sridhar, V.; Johnson, E.F.; McRee, D.E. Mammalian microsomal cytochrome P450 monooxygenase: Structural adaptations for membrane binding and functional diversity. *Mol. Cell* **2000**, *5*, 121–131. [[CrossRef](#)]
58. Kumar, S. Engineering cytochrome P450 biocatalysts for biotechnology, medicine and bioremediation. *Expert Opin. Drug Metab. Toxicol.* **2010**, *6*, 115–131. [[CrossRef](#)]
59. Gharaghani, S.; Khayamian, T.; Keshavarz, F. Docking, molecular dynamics simulation studies, and structure-based QSAR model on cytochrome P450 2A6 inhibitors. *Struct. Chem.* **2012**, *23*, 341–350. [[CrossRef](#)]
60. Sato, A.; Yuki, H.; Watanabe, C.; Saito, J.-I.; Konagaya, A.; Honma, T. Prediction of the site of CYP3A4 metabolism of tolterodine by molecular dynamics simulation from multiple initial structures of the CYP3A4-tolterodine complex. *Chem Bio Inform. J.* **2017**, *17*, 38–52. [[CrossRef](#)]
61. Seifert, A.; Tatzel, S.; Schmid, R.D.; Pleiss, J. Multiple molecular dynamics simulations of human p450 monooxygenase CYP2C9: The molecular basis of substrate binding and regioselectivity toward warfarin. *Proteins Struct. Funct. Bioinform.* **2006**, *64*, 147–155. [[CrossRef](#)] [[PubMed](#)]
62. Shao, Y.-X.; Zhao, P.; Li, Z.; Liu, M.; Liu, P.; Huang, M.; Luo, H.-B. The molecular basis for the inhibition of human cytochrome P450 1A2 by oroxylin and wogonin. *Eur. Biophys. J.* **2012**, *41*, 297–306. [[CrossRef](#)] [[PubMed](#)]

63. Morris, G.M.; Huey, R.; Lindstrom, W.; Sanner, M.F.; Belew, R.K.; Goodsell, D.S.; Olson, A.J. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J. Comput. Chem.* **2009**, *30*, 2785–2791. [[CrossRef](#)] [[PubMed](#)]
64. Moore, C.D.; Shahrokh, K.; Sontum, S.F.; Cheatham, T.E., III; Yost, G.S. Improved cytochrome P450 3A4 molecular models accurately predict the Phe215 requirement for raloxifene dehydrogenation selectivity. *Biochemistry* **2010**, *49*, 9011–9019. [[CrossRef](#)] [[PubMed](#)]
65. Case, D.; Babin, V.; Berryman, J.; Betz, R.; Cai, Q.; Cerutti, D.; Cheatham, T., III; Darden, T.; Duke, R.; Gohlke, H. *AMBER 14*; University of California: San Francisco, CA, USA, 2014.
66. Lobanov, M.Y.; Bogatyreva, N.; Galzitskaya, O. Radius of gyration as an indicator of protein structure compactness. *Mol. Biol.* **2008**, *42*, 623–628. [[CrossRef](#)]
67. Lonsdale, R.; Fort, R.M.; Rydberg, P.; Harvey, J.N.; Mulholland, A.J. Quantum mechanics/molecular mechanics modeling of drug metabolism: Mexiletine N-hydroxylation by cytochrome P450 1A2. *Chem. Res. Toxicol.* **2016**, *29*, 963–971. [[CrossRef](#)] [[PubMed](#)]
68. Harvey, J.N.; Bathelt, C.M.; Mulholland, A.J. QM/MM modeling of compound I active species in cytochrome P450, cytochrome C peroxidase, and ascorbate peroxidase. *J. Comput. Chem.* **2006**, *27*, 1352–1362. [[CrossRef](#)] [[PubMed](#)]
69. Lonsdale, R.; Mulholland, J.A. QM/MM modelling of drug-metabolizing enzymes. *Curr. Top. Med. Chem.* **2014**, *14*, 1339–1347. [[CrossRef](#)] [[PubMed](#)]
70. Tanaka, T.; Okuda, T.; Yamamoto, Y. Characterization of the CYP3A4 active site by homology modeling. *Chem. Pharm. Bull.* **2004**, *52*, 830–835. [[CrossRef](#)]
71. Williams, P.A.; Cosme, J.; Vinković, D.M.; Ward, A.; Angove, H.C.; Day, P.J.; Vornrhein, C.; Tickle, I.J.; Jhoti, H. Crystal structures of human cytochrome P450 3A4 bound to metyrapone and progesterone. *Science* **2004**, *305*, 683–686. [[CrossRef](#)]
72. Fowler, S.M.; Taylor, J.M.; Friedberg, T.; Wolf, C.R.; Riley, R.J. CYP3A4 active site volume modification by mutagenesis of leucine 211. *Drug Metab. Dispos.* **2002**, *30*, 452–456. [[CrossRef](#)] [[PubMed](#)]
73. Fowler, S.M.; Riley, R.J.; Pritchard, M.P.; Sutcliffe, M.J.; Friedberg, T.; Wolf, C.R. Amino acid 305 determines catalytic center accessibility in CYP3A4. *Biochemistry* **2000**, *39*, 4406–4414. [[CrossRef](#)] [[PubMed](#)]
74. Roussel, F.; Khan, K.K.; Halpert, J.R. The importance of SRS-1 residues in catalytic specificity of human cytochrome P450 3A4. *Arch. Biochem. Biophys.* **2000**, *374*, 269–278. [[CrossRef](#)] [[PubMed](#)]
75. Szklarz, G.D.; Halpert, J.R. Molecular modeling of cytochrome P450 3A4. *J. Comput. Aided Mol. Des.* **1997**, *11*, 265–272. [[CrossRef](#)] [[PubMed](#)]
76. Woods, C.J.; Malaisree, M.; Michel, J.; Long, B.; McIntosh-Smith, S.; Mulholland, A.J. Rapid decomposition and visualisation of protein–ligand binding free energies by residue and by water. *Faraday Discuss.* **2014**, *169*, 477–499. [[CrossRef](#)] [[PubMed](#)]
77. Woods, C.J.; Malaisree, M.; Long, B.; McIntosh-Smith, S.; Mulholland, A.J. Computational assay of H7N9 influenza neuraminidase reveals R292K mutation reduces drug binding affinity. *Sci. Rep.* **2013**, *3*, 3561. [[CrossRef](#)] [[PubMed](#)]
78. Callegari, D.; Ranaghan, K.; Woods, C.; Minari, R.; Tiseo, M.; Mor, M.; Mulholland, A.; Lodola, A. L718Q mutant EGFR escapes covalent inhibition by stabilizing a non-reactive conformation of the lung cancer drug osimertinib. *Chem. Sci.* **2018**, *9*, 2740–2749. [[CrossRef](#)] [[PubMed](#)]
79. Rinaldi, S.; Van der Kamp, M.W.; Ranaghan, K.E.; Mulholland, A.J.; Colombo, G. Understanding Complex Mechanisms of Enzyme Reactivity: The Case of Limonene-1, 2-Epoxyde Hydrolases. *ACS Catal.* **2018**, *8*, 5698–5707. [[CrossRef](#)]
80. Thai, K.-M.; Le, D.-P.; Tran, T.-D.; Le, M.-T. Computational assay of Zanamivir binding affinity with original and mutant influenza neuraminidase 9 using molecular docking. *J. Theor. Biol.* **2015**, *385*, 31–39. [[CrossRef](#)] [[PubMed](#)]
81. Ahmad, S.; Raza, S.; Abbasi, S.W.; Azam, S.S. Identification of natural inhibitors against *Acinetobacter baumannii*-alanine-d-alanine ligase enzyme: A multi-spectrum in silico approach. *J. Mol. Liq.* **2018**, *262*, 460–475. [[CrossRef](#)]
82. Ahmad, S.; Raza, S.; Uddin, R.; Rungrotmongkol, T.; Azam, S.S. From phylogeny to protein dynamics: A computational hierarchical quest for potent drug identification against an emerging enteropathogen “*Yersinia enterocolitica*”. *J. Mol. Liq.* **2018**, *265*, 372–389. [[CrossRef](#)]

83. Ahmad, S.; Raza, S.; Abro, A.; Liedl, K.R.; Azam, S.S. Toward novel inhibitors against KdsB: A highly specific and selective broad-spectrum bacterial enzyme. *J. Biomol. Struct. Dyn.* **2019**, *37*, 1326–1345. [[CrossRef](#)] [[PubMed](#)]
84. Ahmad, S.; Murtaza, U.A.; Raza, S.; Azam, S.S. Blocking the catalytic mechanism of MurC ligase enzyme from *Acinetobacter baumannii*: An in Silico guided study towards the discovery of natural antibiotics. *J. Mol. Liq.* **2019**, *281*, 117–133. [[CrossRef](#)]
85. Agoni, C.; Ramharack, P.; Soliman, M.E. Co-inhibition as a strategic therapeutic approach to overcome rifampin resistance in tuberculosis therapy: Atomistic insights. *Future Med. Chem.* **2018**, *10*, 1665–1675. [[CrossRef](#)]
86. Raza, S.; Ranaghan, K.E.; van der Kamp, M.W.; Woods, C.J.; Mulholland, A.J.; Azam, S.S. Visualizing protein–ligand binding with chemical energy-wise decomposition (CHEWD): Application to ligand binding in the kallikrein-8 S1 Site. *J. Comput. Aided Mol. Des.* **2019**, *33*, 461–475. [[CrossRef](#)] [[PubMed](#)]
87. Kang, Y.; Park, S.; Yim, C.; Kwak, H.; Gajendrarao, P.; Krishnamoorthy, N.; Yun, S.C.; Lee, K.; Han, K. The CYP3A4\* 18 genotype in the cytochrome P450 3A4 gene, a rapid metabolizer of sex steroids, is associated with low bone mineral density. *Clin. Pharmacol. Ther.* **2009**, *85*, 312–318. [[CrossRef](#)]
88. Fan, J.-R.; Zheng, Q.-C.; Cui, Y.-L.; Li, W.-K.; Zhang, H.-X. Investigation of ligand selectivity in CYP3A7 by molecular dynamics simulations. *J. Biomol. Struct. Dyn.* **2015**, *33*, 2360–2367. [[CrossRef](#)]
89. Cui, Y.-L.; Xu, F.; Wu, R. Molecular dynamics investigations of regioselectivity of anionic/aromatic substrates by a family of enzymes: A case study of diclofenac binding in CYP2C isoforms. *Phys. Chem. Chem. Phys.* **2016**, *18*, 17428–17439. [[CrossRef](#)]
90. Hou, T.; Wang, J.; Li, Y.; Wang, W. Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations. *J. Chem. Inf. Model.* **2010**, *51*, 69–82. [[CrossRef](#)]
91. Qian, H.; Duan, M.; Sun, X.; Chi, M.; Zhao, Y.; Liang, W.; Du, J.; Huang, J.; Li, B. The binding mechanism between azoles and FgCYP51B, sterol 14 $\alpha$ -demethylase of *Fusarium graminearum*. *Pest. Manag. Sci.* **2018**, *74*, 126–134. [[CrossRef](#)]
92. Fan, J.R.; Li, H.; Zhang, H.X.; Zheng, Q.C. Exploring the structure characteristics and major channels of cytochrome P450 2A6, 2A13, and 2E1 with pilocarpine. *Biopolymers* **2018**, *109*, e23108. [[CrossRef](#)]
93. Gaulton, A.; Bellis, L.J.; Bento, A.P.; Chambers, J.; Davies, M.; Hersey, A.; Light, Y.; McGlinchey, S.; Michalovich, D.; Al-Lazikani, B. ChEMBL: A large-scale bioactivity database for drug discovery. *Nucleic Acids Res.* **2012**, *40*, D1100–D1107. [[CrossRef](#)] [[PubMed](#)]
94. Dennington, R.; Keith, T.; Millam, J. *GaussView*, version 5; Semichem Inc.: Shawnee Mission, KS, USA, 2009.
95. Webb, B.; Sali, A. Comparative protein structure modeling using MODELLER. *Curr. Protoc. Bioinform.* **2014**, *47*, 5.6.1–5.6.32. [[CrossRef](#)] [[PubMed](#)]
96. Kapelyukh, Y.; Paine, M.J.; Maréchal, J.-D.; Sutcliffe, M.J.; Wolf, C.R.; Roberts, G. Multiple substrate binding by cytochrome P450 3A4: Estimation of the number of bound substrate molecules. *Drug Metab. Dispos.* **2008**, *36*, 2136–2144. [[CrossRef](#)] [[PubMed](#)]
97. Morris, G.M.; Goodsell, D.S.; Halliday, R.S.; Huey, R.; Hart, W.E.; Belew, R.K.; Olson, A.J. Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J. Comput. Chem.* **1998**, *19*, 1639–1662. [[CrossRef](#)]
98. Huey, R.; Morris, G.M.; Olson, A.J.; Goodsell, D.S. A semiempirical free energy force field with charge-based desolvation. *J. Comput. Chem.* **2007**, *28*, 1145–1152. [[CrossRef](#)]
99. Jorgensen, W.L.; Chandrasekhar, J.; Madura, J.D.; Impey, R.W.; Klein, M.L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926–935. [[CrossRef](#)]
100. Wang, J.; Wolf, R.M.; Caldwell, J.W.; Kollman, P.A.; Case, D.A. Development and testing of a general amber force field. *J. Comput. Chem.* **2004**, *25*, 1157–1174. [[CrossRef](#)]
101. Maier, J.A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K.E.; Simmerling, C. ff14SB: Improving the accuracy of protein side chain and backbone parameters from ff99SB. *J. Chem. Theory Comput.* **2015**, *11*, 3696. [[CrossRef](#)]
102. Shahrokh, K.; Orendt, A.; Yost, G.S.; Cheatham, T.E., III. Quantum mechanically derived AMBER-compatible heme parameters for various states of the cytochrome P450 catalytic cycle. *J. Comput. Chem.* **2012**, *33*, 119–133. [[CrossRef](#)]

103. Cornell, W.D.; Cieplak, P.; Bayly, C.I.; Gould, I.R.; Merz, K.M.; Ferguson, D.M.; Spellmeyer, D.C.; Fox, T.; Caldwell, J.W.; Kollman, P.A. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.* **1995**, *117*, 5179–5197. [[CrossRef](#)]
104. Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H.J. Numerical integration of the cartesian equations of motion of a system with constraints: Molecular dynamics of n-alkanes. *J. Comput. Phys.* **1977**, *23*, 327–341. [[CrossRef](#)]
105. Madhulatha, T.S. An overview on clustering methods. *IOSR J. Eng.* **2012**, *2*, 719–725. [[CrossRef](#)]
106. Metropolis, N.; Rosenbluth, A.W.; Rosenbluth, M.N.; Teller, A.H.; Teller, E. Equation of state calculations by fast computing machines. *J. Chem. Phys.* **1953**, *21*, 1087–1092. [[CrossRef](#)]
107. Tyka, M.D.; Sessions, R.B.; Clarke, A.R. Absolute free-energy calculations of liquids using a harmonic reference state. *J. Phys. Chem. B* **2007**, *111*, 9571–9580. [[CrossRef](#)] [[PubMed](#)]
108. Hamelberg, D.; McCammon, J.A. Standard free energy of releasing a localized water molecule from the binding pockets of proteins: Double-decoupling method. *J. Am. Chem. Soc.* **2004**, *126*, 7683–7689. [[CrossRef](#)] [[PubMed](#)]
109. Pearlman, D.A. A comparison of alternative approaches to free energy calculations. *J. Phys. Chem.* **1994**, *98*, 1487–1493. [[CrossRef](#)]
110. Woods, C.J.; Essex, J.W.; King, M.A. The development of replica-exchange-based free-energy methods. *J. Phys. Chem. B* **2003**, *107*, 13703–13710. [[CrossRef](#)]
111. Woods, C.J.; Essex, J.W.; King, M.A. Enhanced configurational sampling in binding free-energy calculations. *J. Phys. Chem. B* **2003**, *107*, 13711–13718. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).