



OPEN

## Molecular determinant of substrate binding and specificity of cytochrome P450 2J2

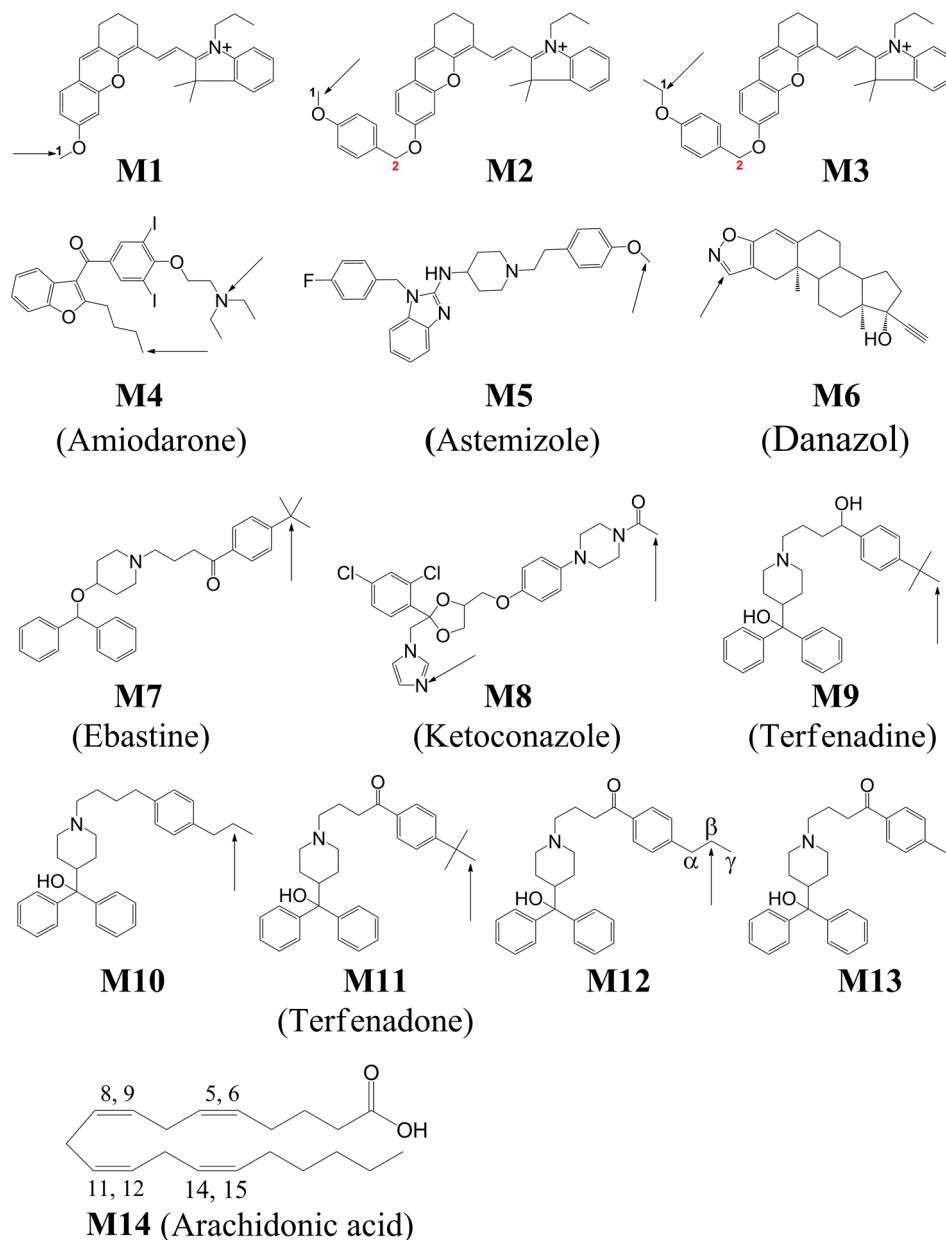
Liang Xu<sup>✉</sup> & Liao Y. Chen

Cytochrome P450 2J2 (CYP2J2) is responsible for the epoxidation of endogenous arachidonic acid, and is involved in the metabolism of exogenous drugs. To date, no crystal structure of CYP2J2 is available, and the proposed structural basis for the substrate recognition and specificity in CYP2J2 varies with the structural models developed using different computational protocols. In this study, we developed a new structural model of CYP2J2, and explored its sensitivity to substrate binding by molecular dynamics simulations of the interactions with chemically similar fluorescent probes. Our results showed that the induced-fit binding of these probes led to the preferred active poses ready for the catalysis by CYP2J2. Divergent conformational dynamics of CYP2J2 due to the binding of each probe were observed. However, a stable hydrophobic clamp composed of residues I127, F310, A311, V380, and I487 was identified to restrict any substrate access to the active site of CYP2J2. Molecular docking of a series of compounds including amiodarone, astemizole, danazol, ebastine, ketoconazole, terfenadine, terfenadone, and arachidonic acid to CYP2J2 confirmed the role of those residues in determining substrate binding and specificity of CYP2J2. In addition to the flexibility of CYP2J2, the present work also identified other factors such as electrostatic potential in the vicinity of the active site, and substrate strain energy and property that have implications for the interpretation of CYP2J2 metabolism.

Cytochrome P450 (CYP) enzymes constitute a diverse group of heme-containing proteins, which play pivotal roles in the metabolism of xenobiotics including a wide variety of drugs, and the conversion of polyunsaturated fatty acids (PUFAs) to biologically active molecules<sup>1,2</sup>. The crystal structures of a large number of CYPs reveal that CYPs share a similar overall protein topology but exhibit varying substrate selectivity<sup>2,3</sup>. The molecular basis for this unusual specificity of CYPs could be attributed to the plasticity of the substrate-binding pocket near the heme motif that enables a specific CYP to accommodate one or two substrates in its active site<sup>4</sup>. Among 57 CYPs identified in human<sup>5</sup>, the 2J2 isoform (CYP2J2) mediates lipid metabolism of PUFAs via epoxidation reactions, yielding the epoxyeicosatrienoic acids (EETs). For instance, in human heart, the CYP2J2 enzyme metabolizes arachidonic acid (ARA, **M14**, Fig. 1) predominantly via olefin epoxidation to all four regioisomeric EETs, with epoxidation occurring preferentially at the 14,15-olefin (37% of total EET products), followed by 11,12-EET (18%), 8,9-EET (24%), and 5,6-EET (21%)<sup>6</sup>. The metabolism of ARA to EETs seems to relate to the cardioprotective role of CYP2J2<sup>7</sup>, but could be influenced by the metabolism of drugs that competitively or noncompetitively inhibit the EET synthesis and consequently lead to cardiotoxicity<sup>8</sup>. In addition to cardiovascular disease, CYP2J2 is also implicated in cerebrovascular diseases, diabetes, and cancer<sup>9</sup>.

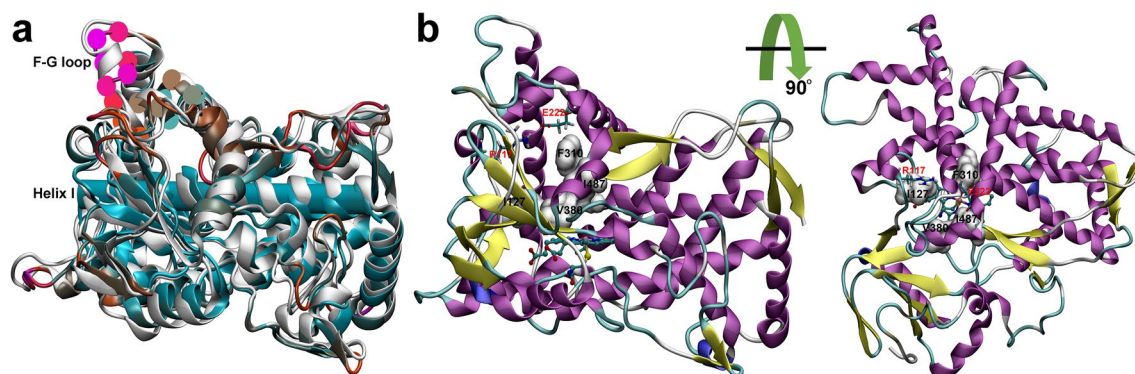
Besides endogenous substrates, CYP2J2 has been reported to metabolize some antihistamine drugs such as astemizole (**M5**, Fig. 1)<sup>10</sup>, ebastine (**M7**, Fig. 1)<sup>11</sup>, and terfenadine (**M9**, Fig. 1)<sup>12</sup>. Several new substrates including tamoxifen, albendazole, danazol (**M6**, Fig. 1), amiodarone (**M4**, Fig. 1), and ketoconazole (**M8**, Fig. 1) were also identified<sup>13–16</sup>. Interestingly, danazol was shown to strongly inhibit the CYP2J2-mediated metabolism of albendazole, astemizole, ebastine, and terfenadine in a substrate-independent manner<sup>17</sup>. Further studies suggested that the inhibitory effects of danazol depended on the human microsome used<sup>18</sup>, and terfenadone (**M11**, Fig. 1) was proposed as a general CYP2J2 inhibitor in both human liver and intestinal microsomes<sup>18</sup>. Despite its critical role in drug metabolism, the lack of crystal structure of CYP2J2 hampers a better understanding of its adaptation to chemically diverse compounds. The emerging computational modeling and simulations provided invaluable insights into the structure of CYP2J2 and corresponding substrate interactions. Several CYP2J2 homology models have been constructed based on single<sup>19–21</sup> or multiple templates<sup>12,13,22</sup>. These templates are primarily CYP gene in group 2 (CYP2) showing varying sequence identity (26–46%) with CYP2J2. A brief summary of these published

Department of Physics and Astronomy, University of Texas at San Antonio, One UTSA Circle, San Antonio, TX 78249, USA. ✉email: liang.xu@utsa.edu



**Figure 1.** Molecular structures of **M1–M14** used in this work. The  $C^1$  in **M1–M3** indicated the actual metabolic site catalyzed by CYP2J2 whereas the  $C^2$  in **M2** and **M3** was the metabolic site predicted by current molecular docking studies. The potential metabolic sites of **M4–M14** was indicated by an arrow (**M4–M13**) or labels (**M14**). The metabolic site of **M1–M3** was adopted from reference<sup>25</sup>; The metabolic site of **M4** was adopted from reference<sup>14</sup>; The metabolic site of **M5** was adopted from reference<sup>10</sup>; The metabolic site of **M6** was adopted from reference<sup>11</sup>; The metabolic site of **M7** was adopted from reference<sup>13</sup>; The metabolic site of **M8** was adopted from CYP3A4 in complex with two **M8** (PDB ID: 2J0C, see reference<sup>23</sup>) and reference<sup>26</sup>; The metabolic site of **M9–M13** was adopted from reference<sup>12</sup>; and The metabolic site of **M14** was adopted from reference<sup>6</sup>.

models was provided by Proietti et al.<sup>21</sup>. When used to interpret the protein-substrate interactions, it seems challenging to achieve consistent results with structural models generated using different protocols. For instance, based on the CYP2J2 homology model developed using multiple CYP2 template proteins (2A6, 2B4, 2C5, 2C8, and 2D6), one terfenadone derivative (**M12**, Fig. 1) was predicted to interact with Arg117 through its keto (C=O) group<sup>12</sup>, whereas such an interaction was not found in another study where a CYP2J2 model was constructed based on the crystal structure of CYP2C9<sup>19</sup>. The difficulty in the prediction of the preference of metabolic site of ARA was demonstrated in two studies of interactions between CYP2J2 and ARA. Based on the CYP2J2 model generated using the template of CYP2R1, ARA was predicted to form hydrogen bonds with Gly486 and Leu378 in its favorable binding mode<sup>20</sup>. However, these hydrogen bond interactions were not observed in another study using the CYP2J2 model built based on the crystal structure of rabbit CYP2B4, and the preferred binding pose



**Figure 2.** (a) Structural superimposition of initial (white color) and the representative conformations of CYP2J2. The helix I composed of residues 297–330 was also shown in Cartoon representation. The mobile F-G loop formed by residues 224–240 was highlighted by spheres. The flexibility increased from the cyan region (inflexible) to the orange and magenta regions (most flexible). (b) Representative conformation of CYP2J2 in side and top views, obtained from the 500-ns MD simulations of ligand-free CYP2J2. The key residues Ile127, Phe310, Ala311, Val380, and Ile487 were shown in white spheres (hydrophobic residues). The heme moiety was shown in CPK representation. Arg117 and Glu222, which formed a stable salt-bridge during simulations, were also highlighted in Licorice representation and labeled in red color. This figure was rendered using VMD<sup>27</sup> (<http://www.ks.uiuc.edu/Research/vmd/>).

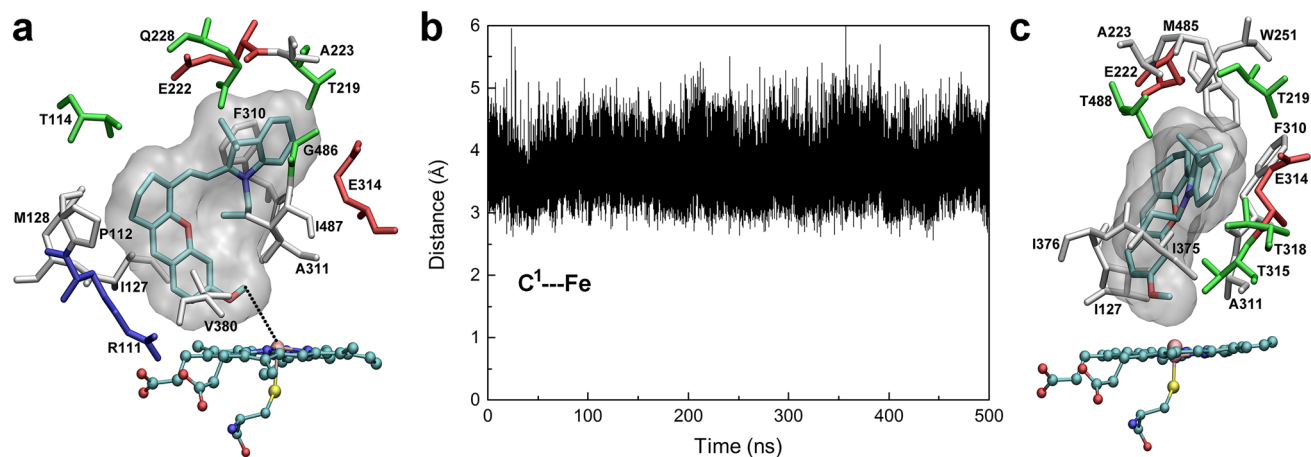
that yielded the metabolic product of 13,14-EET was not predicted<sup>21</sup>. In addition, the binding modes of the above-mentioned antihistamine drugs in the active site of those developed CYP2J2 models were not carefully examined. The multiple ligand binding modes and the conformational dynamics induced by ligand binding were observed in CYP3A4, indicating the structural flexibility in CYP enzymes<sup>23</sup>. On the other hand, CYP3A4 could regioselectively metabolize chemically diverse drugs, suggesting specific protein-substrate interactions in the active site of CYP enzymes<sup>24</sup>. The substrate promiscuity in CYP enzymes challenges the attempt to identify the molecular determinant of the substrate specificity in the active site of CYP2J2.

In this work, we revisited the structure of CYP2J2 by homology modeling based on multiple templates, and fully relaxed the model structure by performing 500-ns molecular dynamics (MD) simulations in aqueous solution (Fig. 2). Three newly reported fluorescent probes (**M1**–**M3**, Fig. 1)<sup>25</sup> were used to investigate the protein–ligand interactions and conformational dynamics of CYP2J2 upon ligand binding. The CYP2J2-mediated *O*-demethylation (*C*<sup>1</sup> in **M1**–**M3**, Fig. 1) and subsequent 1,6-elimination of *p*-hydroxybenzyl resulted in the probe that produced a strong near-infrared fluorescence signal<sup>25</sup>. These compounds are structurally similar and suitable to study the difference in the ligand binding modes caused by the subtle variations in the ligand structures. The present simulation results showed that these three probes displayed distinct interactions with CYP2J2 in the binding pocket. The probe **M2** showing the highest fluorescence response toward CYP2J2 was predicted to have the highest binding affinity for CYP2J2. The molecular determinant for the access of any substrate to the active center of CYP2J2 was identified based on our computational simulation and docking results. The structural information obtained from the present study sheds light on the substrate specificity mediated by interactions with CYP2J2, and may have implications for drugs metabolized by CYP2J2.

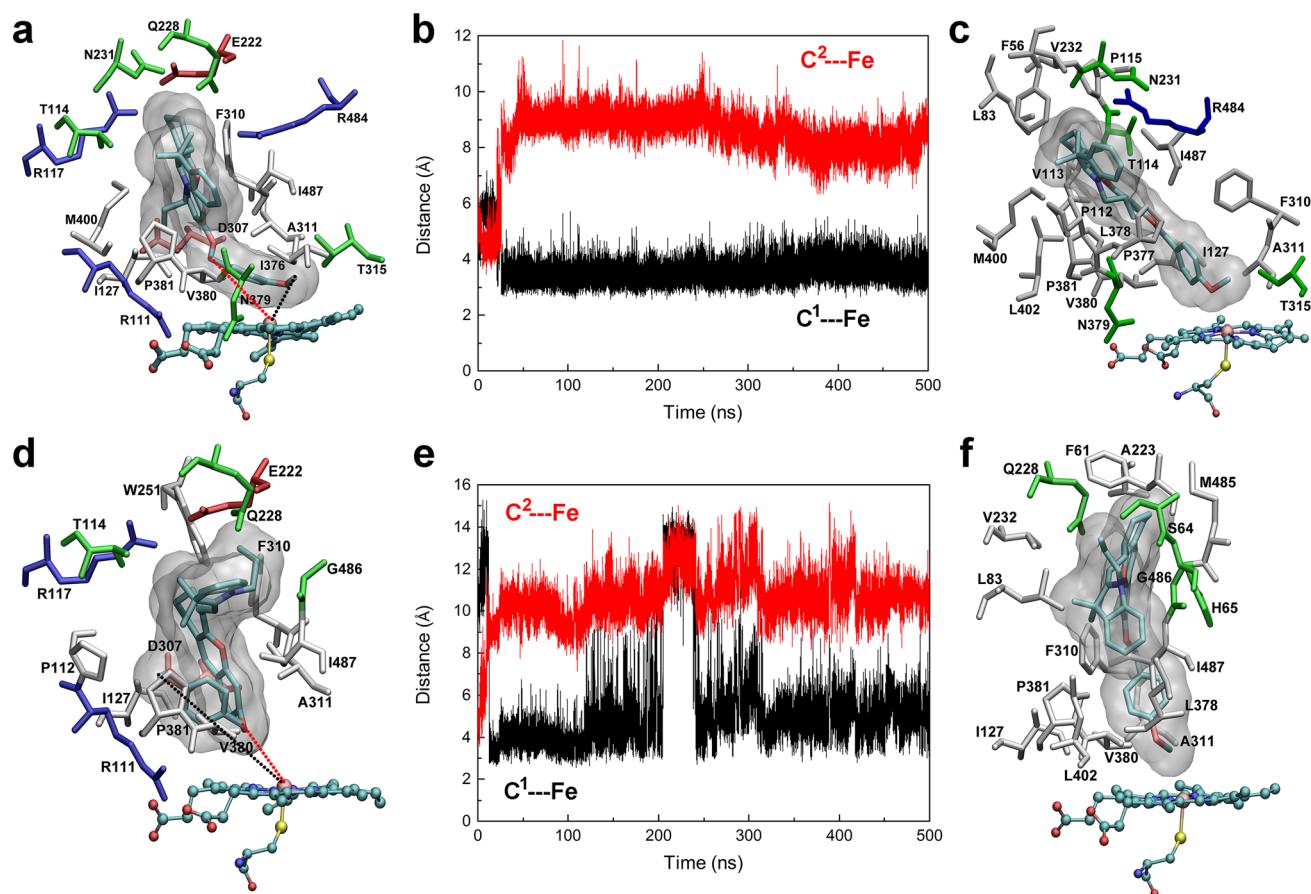
## Results

**Induced-fit binding of M1–M3 to CYP2J2 resulted in active conformers.** The stability and structural fluctuations of the CYP2J2 structure were demonstrated by the root-mean-square deviation (RMSD) and root-mean-square fluctuation (RMSF) values over the 500-ns simulations (Fig. S1 under Supplementary Information). After 300-ns dynamics, the structure seemed well equilibrated with the RMSDs below 2.0 Å and RMSFs below 2.5 Å. For CYP2J2 in complex with **M1**/**M2**/**M3**, the RMSDs were less than 2.5 Å. The large fluctuations in the RMSFs appeared in **M2**-bound CYP2J2. The binding of **M2** in the extended (**M2a**) and folded conformer (**M2b**) resulted in more flexible F-G loop composed of residues 224–240 and the N-terminus, respectively. The conserved helix I region composed of residues 290–330 remained relatively stable (Fig. 2 and Fig. S1). The volume of the binding pocket of the representative conformation of CYP2J2 was calculated using the MDpocket software<sup>28</sup>, and a value of 1499 Å<sup>3</sup> was obtained, which seemed large enough to accommodate **M1** and **M2** in their extended conformer. As shown in Figs. 3a and 4a, the initial binding pose of **M1** and **M2** displayed the optimal distance between the catalytic site (*C*<sup>1</sup> in **M1**–**M3**, Fig. 1) and the heme iron. However, for **M3**, such an extended conformer was not found in the current docking study, and a folded conformer with the minimum distance between another carbon atom (*C*<sup>2</sup> in **M3**, Fig. 1) and the heme iron was selected as the initial binding conformation (Fig. 5a). For comparison, a similar folded binding conformer of **M2** was also simulated (Fig. 4d). The overlapping of the initially folded binding conformers of **M2** and **M3** was shown in Fig. S2, indicating similar binding poses.

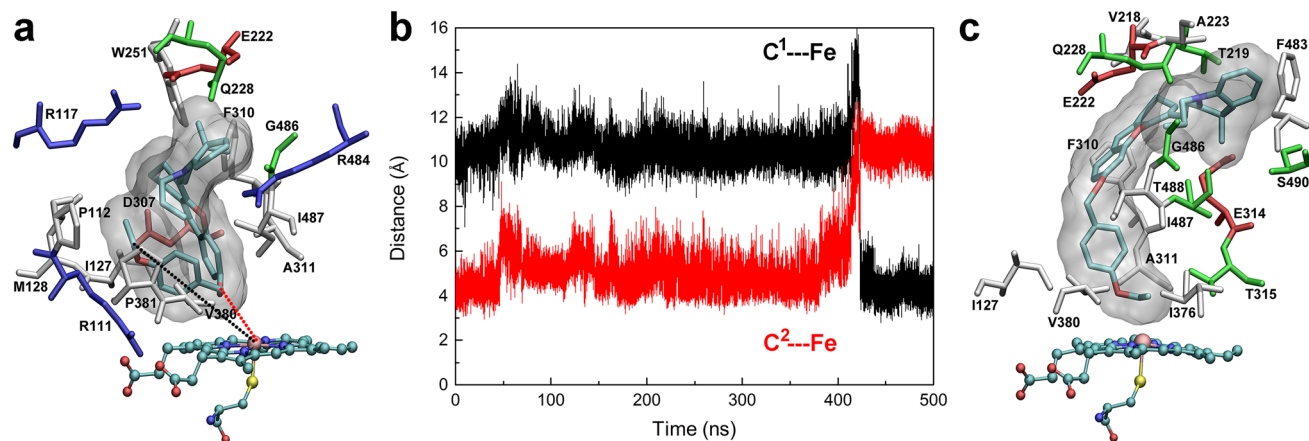
The distance between the catalytic site of **M1** and the heme iron fluctuated marginally from the initial 3.9 Å to a value of  $3.7 \pm 0.2$  Å block-averaged over the last 100-ns simulations (Fig. 3b), implying a relatively stable binding of **M1** to CYP2J2. Because of the induced-fit binding<sup>29</sup>, the interactions of **M1** with CYP2J2 altered



**Figure 3.** MD simulations of CYP2J2 in complex with M1. (a) The initial binding pose of M1 in the active site of CYP2J2; (b) The distance between the proposed metabolic site ( $C^1$  atom) and the heme iron during the simulations; and (c) The representative binding conformation of M1 in the active site of CYP2J2. Interacting residues with a contact frequency  $> 50\%$  calculated from the last 100-ns trajectory were shown and colored based on their types: blue for basic, red for acidic, green for polar, and white for nonpolar residues. A contact occurs if any atom of the ligand is within  $3 \text{ \AA}$  of any atom of protein residues. This figure was rendered using VMD<sup>27</sup> (<http://www.ks.uiuc.edu/Research/vmd/>).



**Figure 4.** MD simulations of CYP2J2 in complex with M2 in the extended (a–c) or folded conformation (d–f). (a, d) The initial binding pose of M2 in the active site of CYP2J2; (b, e) The distance between the proposed metabolic site ( $C^1$  atom) and the heme iron during the simulations. For comparison, the distance between another  $C^2$  atom and the heme iron was also monitored; and (c, f) The representative binding conformation of M2 in the active site of CYP2J2. This figure was rendered using VMD<sup>27</sup> (<http://www.ks.uiuc.edu/Research/vmd/>).



**Figure 5.** MD simulations of CYP2J2 in complex with **M3**. (a) The initial binding pose of **M3** in the active site of CYP2J2; (b) The distance between the proposed metabolic site ( $C^1$  atom) and the heme iron during the simulations. For comparison, the distance between another  $C^2$  atom and the heme iron was also monitored; and (c) The representative binding conformation of **M3** in the active site of CYP2J2. This figure was rendered using VMD<sup>27</sup> (<http://www.ks.uiuc.edu/Research/vmd/>).

the orientation of **M1** in the active site, which was demonstrated by the changes in the interacting residues. Residues including Arg111, Pro112, Thr114, Met128, Gln228, Val380, Gly486, and Ile487 that were observed in the initial binding conformation were not identified to interact with **M1** during the simulations. Several new residues such as Trp251, Thr315, Thr318, Ile375, Ile376, Met485, and Thr488 were found to be involved in the interactions with **M1**. And residues including Ile127, Thr219, Glu222, Ala223, Phe310, Ala311, and Glu314 persistently interacted with **M1**.

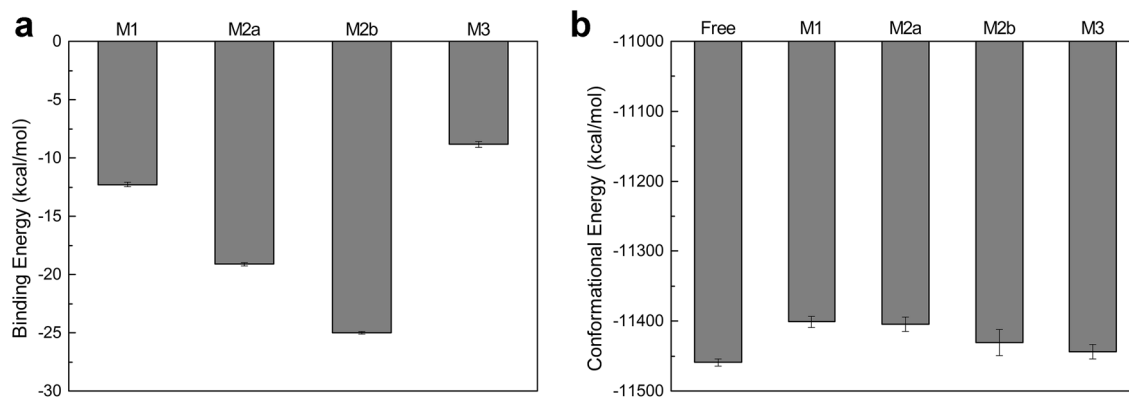
Two binding poses of **M2** in the active site of CYP2J2 were investigated, with **M2** in either extended (Fig. 4a–c) or folded conformer (Fig. 4d–f). Except for the large fluctuations during the first 50-ns simulations, the compound **M2** with an initially extended structure seemed adapted to the binding environment of CYP2J2, as the distance between the catalytic site and the heme iron was well maintained, and an average value of  $3.7 \pm 0.1$  Å was obtained based on the last 100-ns simulations. Similar to **M1**, the induced-fit binding of **M2** also mediated its interactions with CYP2J2. More preserved residues were identified including Thr114, Ile127, Asn231, Phe310, Ala311, Thr315, Asn379, Val380, Pro381, Met400, Arg484, and Ile487.

Interestingly, when **M2** bound to CYP2J2 in a folded state (Fig. 4d–f), it quickly changed its conformation from the initially deformed state to the extended state, probably because the deformed binding pose could cause large strain energy that may not be compensated by the binding interactions with the protein<sup>30</sup>. As a consequence, **M2** adopted a different binding pose that was induced by the interactions with CYP2J2. And the distance between the catalytic site and the heme iron decreased from the initial 10 Å to  $5.0 \pm 0.2$  Å (block-averaged over the last 100-ns simulations). Although **M2** underwent significant conformational changes, there were still eight interacting residues preserved, including Ile127, Gln228, Phe310, Ala311, Val380, Pro381, Gly486, and Ile487.

In the case of **M3** where it adopted a deformed conformation similar to that of **M2** initially (Fig. S2), it took **M3** a longer time (~200 ns) to adopt the extended conformation in the binding pocket of CYP2J2; then approach to the active site after 423 ns; and remain in the bound state (Fig. 5 and Fig. S3) till the end of the simulations. The calculated distance between the catalytic site and the heme iron was  $4.3 \pm 0.1$  Å (block-averaged over the last 50-ns simulations). The preserved interacting residues included Ile127, Glu222, Gln228, Phe310, Ala311, Val380, Gly486, and Ile487.

**Structural basis for the high binding affinity of M2 to CYP2J2.** The *O*-alkyl group in the structure of **M1**–**M3** was supposed to be the metabolic site. And **M1** showed rather weak fluorescence response, indicating that **M1** could hardly be catalyzed by CYP2J2<sup>25</sup>. With the introduction of the linker (*p*-hydroxybenzyl) to **M1**, **M2** displayed the highest fluorescence response toward CYP2J2. Replacing the *O*-methyl group in **M2** with *O*-ethyl group led to **M3**, which significantly reduced the CYP2J2-mediated *O*-demethylation<sup>25</sup>. From our simulations, it was found that these three probes were capable of binding to the catalytic site of CYP2J2, especially for **M2** and **M3** even when their initial binding poses deviated from the final active conformers, indicating a high binding preference of **M2** and **M3** to CYP2J2. The distance from the heme iron to the metabolic site of the substrate was assumed to be related to the rate of metabolism of CYP2J2<sup>12,31</sup>. Based on this assumption, **M1** ( $3.7 \pm 0.2$  Å) should exhibit a higher reactivity than **M3** ( $4.3 \pm 0.1$  Å), which seemed to contradict the experimental results.

The distance between the catalytic site of substrates to the heme iron varies in different CYP enzymes. For instance, a distance of 2.1 Å was observed in the crystal structure of CYP3A4 in complex with ketoconazole (PDB ID: 2J0C)<sup>23</sup>; a larger distance of 4.4 Å was observed in another crystal structure of CYP3A4 in complex with midazolam (PDB ID: 5TE8)<sup>24</sup>; and an even larger distance of 6.6 Å was found in the crystal structure of CYP2C9 (PDB ID: 5W0C) in complex with one inhibitor and water molecules in the active site<sup>32</sup>. In the system where the



**Figure 6.** The binding energy of **M1–M3** to CYP2J2 (**a**) and conformational energy of CYP2J2 in complex with **M1–M3** (**b**). For comparison, conformational energy for the free CYP2J2 was also calculated. **M2a** and **M2b** denote initially extended and folded conformation of **M2** in the binding pocket of CYP2J2, respectively.

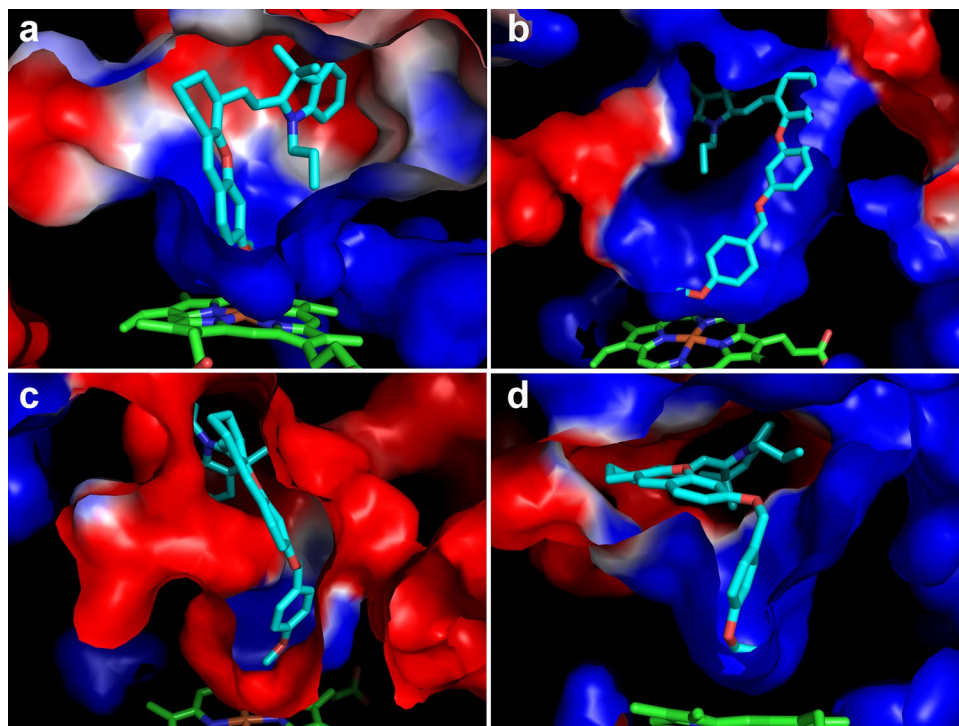
distance from **M2** to heme iron was  $5.0 \pm 0.2$  Å, water molecules were also found to transiently bind to heme, and may prevent **M2** from moving closer to the active site of CYP2J2 (Fig. S4). Taken together, it appeared that the distances between **M1–M3** and heme iron were in a reasonable range, and these three probes could be catalyzed by CYP2J2. However, the difference in their catalytic efficiency ( $\mathbf{M2} > \mathbf{M3} > \mathbf{M1}$ ) implied that the distance may not be the determining factor for substrate metabolism.

The estimated binding affinity of **M1–M3** to CYP2J2 was shown in Fig. 6. **M2** displayed the highest binding affinity to CYP2J2, regardless of their initially different binding poses. The finding that **M2** tended to adopt an extended conformation further suggested that little strain energy was induced when binding to CYP2J2. To test this, 100-ns MD simulations of **M1/M2/M3** in aqueous solution were performed, and it was found that these three molecules exclusively adopted an extended conformation in solution (Fig. S5). Compared to the unbound state in solvent, **M3** became deformed to some extent and the strain energy costs of restraining **M3** to the bound state was not taken into account in the binding energy calculations, which might result in substantial uncertainty in the calculated binding affinity of **M3** to CYP2J2.

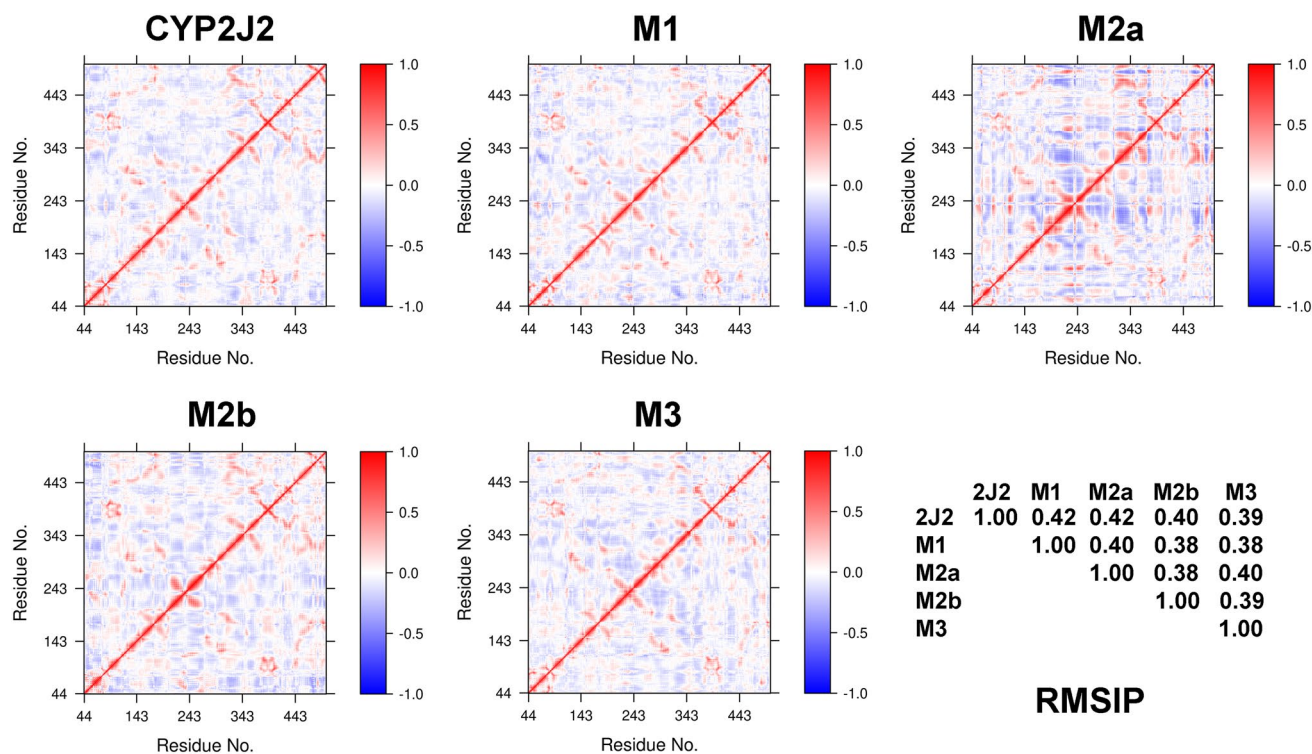
The difference in the interacting residues in the vicinity of the active site could contribute to the differing binding energy of **M2** in different binding poses. Note that **M1–M3** were positively charged, to further examine the effects of these interacting residues to the ligand binding, the electrostatic potential surface was calculated for the structure of ligand-bound CYP2J2 using the APBS tool<sup>34</sup> (Fig. 7). The presence of a significant area of negative electrostatic potential around **M2** provided another support for its highest binding affinity and catalytic efficiency (Fig. 7c). The presence of negative electrostatic potential was also observed in the vicinity of **M3**, especially around the positively charge moiety of **M3** (Fig. 7d). In contrast, for **M2** with a relatively lower binding affinity to CYP2J2, a large area of positive electrostatic potential in the vicinity of it was found (Fig. 7b), and the presence of both negative and positive potentials around **M1** was observed (Fig. 7a). The favorably negative potential was likely to play a critical role in stabilizing **M2** in the active site of CYP2J2. For comparison, we also calculated the electrostatic potential for the initial binding conformations (Fig. S6), none of them showed favorable potential for the binding of **M1–M3**, which may explain why **M1–M3** would undergo substantial changes in the binding poses during simulations. Because of the conformational dynamics of CYP2J2, we randomly selected another conformation for each system and calculated the electrostatic potential, and similar results were obtained (Fig. S7), confirming the presence of favorable electrostatic environment for the binding and catalysis of **M2** and **M3**. The present results were consistent with previous studies suggesting that in addition to the hydrophobic interactions, the electrostatic properties of the enzymes' surroundings also influenced the binding of both charged and nonpolar ligands<sup>35,36</sup>.

**Interactions with **M1–M3** induced divergent dynamics of CYP2J2.** The effects of induced-fit binding of **M1–M3** on the conformational dynamics of CYP2J2 were characterized in terms of the conformational energy and the cross-correlated network maps of C $\alpha$  atomic fluctuations using the R package Bio3D<sup>37,38</sup>. The calculated conformational energy suggested that the binding of **M1–M3** destabilized the conformations of CYP2J2 to different extents, compared to the free state of CYP2J2 (Fig. 6b). Of interest, less destabilization effects were observed in the two systems where **M2** and **M3** initially were deformed, implying the capability of CYP2J2 in the mediation of the interactions with the ligands. If **M2** bound to CYP2J2 with the extended conformation, it would become difficult for CYP2J2 to reshape the binding pocket and direct the binding of **M2** toward the conformations with less energy cost.

The dynamical cross-correlation maps (DCCMs) for the free and ligand-bound CYP2J2 were calculated from the MD simulations (Fig. 8). The cross-correlation coefficients vary from a value of 1 for completely correlated motions to  $-1$  for completely anticorrelated motions<sup>39</sup>. More anticorrelated motions were identified for the binding of **M2** with the initially extended conformation, consistent with its destabilization effect on the dynamics of CYP2J2. To quantitatively measure the amount of overlap between two conformations, the root mean square inner product (RMSIP) was calculated on the first 20 eigenvectors obtained from a principal component analysis of the C $\alpha$  covariance matrix of the atomic positional fluctuations<sup>40</sup>. The value of RMSIP ranges from 0



**Figure 7.** Electrostatic potential surface calculated for the representative conformation of CYP2J2 in complex with M1 (a), M2 in the initially extended (b) and folded (c) conformation, and M3 (d). Blue color indicates positive electrostatic potentials, and red color indicates negative electrostatic potentials. This figure was rendered using PyMOL<sup>33</sup>.



**Figure 8.** The dynamical cross-correlation maps (DCCMs) for the free and M1/M2/M3-bound CYP2J2. The root mean square inner product (RMSIP) between any two different systems was also calculated. M2a and M2b denote initially extended and folded conformation of M2, respectively. This figure was rendered using Bio3D<sup>37</sup>.

for mutually orthogonal (no correlation) subspaces to 1 for identical subspaces (full correlation). A value of 0.5 is considered fairly correlated subspaces<sup>41</sup>. The calculated RMSIP values between any two different systems varied between 0.38 and 0.42, indicating significant dissimilarity in the conformational spaces sampled by CYP2J2 bound to different ligands.

One of interesting structural aspects of CYP450 is the substrate access channels that may modulate CYP450 function<sup>42</sup>. To demonstrate the impacts of the distinct conformational dynamics of CYP2J2 in complex with **M1–M3**, the potential substrate access channels were analyzed using CAVER<sup>43</sup>. The heme iron was chosen as the starting point to detect the potential access channels. For comparison, the substrate access channels in the human CYP2B6 in complex with two amlodipine molecules (PDB ID: 3UA5)<sup>44</sup> were also calculated. In this structure, in addition to one substrate in the active site, another substrate was trapped in access pathway, and thus the positions of the two substrates clearly implied the substrate access channel (Fig. S8). Interestingly, a similar substrate access channel was found in CYP2J2 regardless of the presence or absence of the ligand, suggesting that **M1–M3** could use the same channel to reach the active site. In addition to the access channels, other channels that may serve as substrate (or metabolic product) exit channels were detected when there was no or little overlap with the access channel. It was found that a distinct exit channel was identified in the ligand-free CYP2J2 when compared with the exit channel in CYP2B6. And this channel seemed well maintained only in the conformation of CYP2J2 in complex with **M2a**. For the conformations of **M2b-** and **M3-bound** CYP2J2, all other detected channels overlapped with the access channel to different extents and therefore the exit channels seemed to be largely buried in the protein interior. The above results suggested that the conformational dynamics induced by substrates significantly influenced the channel networks, especially the potential substrate exit channels.

**Molecular determinant for the access of substrate to the active center of CYP2J2.** Although **M1–M3** adopted distinct binding poses in the substrate pocket of CYP2J2, five interacting residues (Ile127, Phe310, Ala311, Val380, and Ile487) were identified to control the substrate access to the heme motif. The distances between these residues were calculated and shown in Fig. S9, and no significant deviations from the free CYP2J2 were observed, suggesting that these residues appeared to form a bottleneck that allowed the access of only one substrate to the heme motif. The binding of **M1**, however, resulted in notable increase of the distances associated with Ile487, but no interactions with Ile487 were found for **M1** (Fig. 3b, the actual contact frequency is 0.02%), suggesting that the binding of **M1** might distort this bottleneck. It remains to be elucidated if the catalytic rate of CYP2J2 would be related to the stability of these residues, but the maintenance of their positions in **M2** that exhibited the highest catalytic rate for CYP2J2 might indicate their important role in the metabolism of substrates.

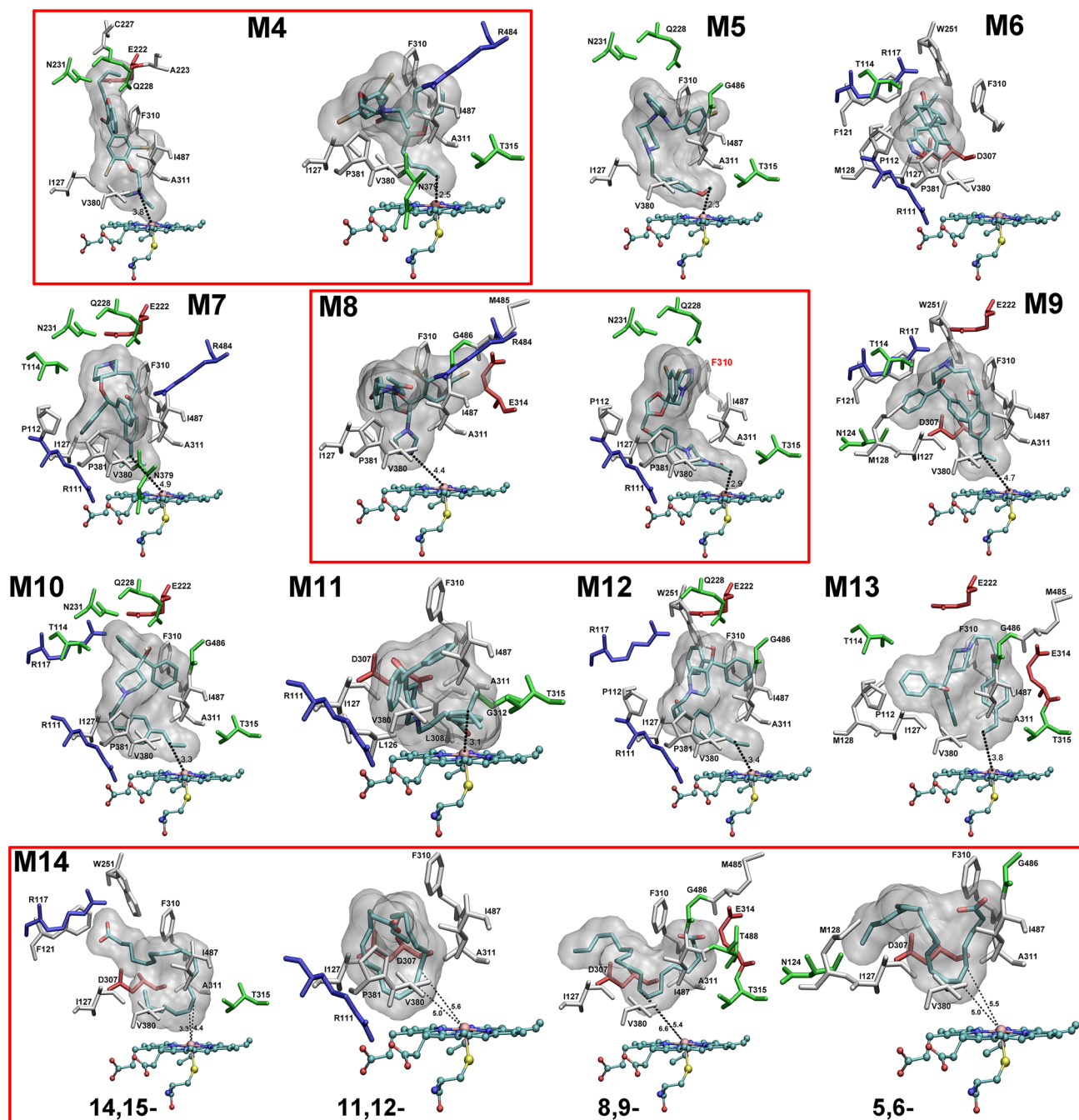
To further assess the availability of the above five residues for substrate binding to CYP2J2, molecular docking was performed to search the optimal binding pose of **M4–M14** (Fig. 1) in CYP2J2. For simplicity, we selected those binding conformers with the minimum distance between the proposed metabolic site(s) and the heme iron, and the results were shown in Fig. 9. Except for danazol (**M6**) where no “preferred” binding pose was found, and ketoconazole (**M8**) where Phe310 was not interacting residue in one of two binding poses, the five hydrophobic residues (Ile127, Phe310, Ala311, Val380, and Ile487) were involved in positioning all the other substrates near the active site of CYP2J2. In the following, we selectively discussed the docking results for some substrates, and compared with available experimental and simulation results.

In the present docking study, danazol (**M6**) was found to occupy the position far away from the active site. The distance between the proposed metabolic site to heme iron is 12.6 Å. Interestingly, only one binding pose as shown in Fig. 9 was identified, and either repeating the docking procedure or increasing the iterations to search more potential conformers led to the same binding pose (data not shown). Experimental data showed that danazol strongly inhibited CYP2J2-mediated astemizole (**M5**), ebastine (**M7**), and terfenadine (**M9**) metabolism in a substrate-independent manner<sup>17</sup>. The “preferred” binding pose identified from our docking study appeared to block any other substrates to enter the active site, which may have implications for its inhibitory mechanism.

For terfenadine (**M9**), experimental results showed that two variants of CYP2J2, P351L and P115L, had similar enzyme activity for *t*-butyl methyl hydroxylation to the wild-type CYP2J2<sup>45</sup>. Based on the present docking result, P351 and P115 were not involved in the interactions with terfenadine, and thus could have little effect on the metabolism of terfenadine. In addition, the G312R variant was implied to destabilize the structure of CYP2J2 by altering the heme environment<sup>45</sup>, and cause the loss of the catalytic activity of CYP2J2<sup>46</sup>. The same effect was also observed in the present CYP2J2 model, as Gly312 was near the active site and the introduction of Arg would significantly change the electrostatic potential around the heme, and consequently impair the enzyme catalytic activity and structural stability.

For the terfenadone derivatives **M10** and **M12** in Fig. 1, CYP2J2 exhibited regioselectivity of hydroxylation of the C $\beta$  atom, and Arg117 was assumed to play a critical role<sup>12,31</sup>. Based on the homology model of CYP2J2, a narrow channel of access to the heme was identified, which included Ile127, Phe310, Ile376, and Val380<sup>12</sup>. In the present CYP2J2 model, Ile376 was found to interact with **M1** and **M3**, but not with **M4–M14**. And Ile487 was found to be closer to the heme than Ile376 in the present CYP2J2 model (Fig. S10). Moreover, Arg117 was found to establish hydrogen bonds with the keto group of **M12** in the previous study<sup>12</sup>. In the current docking study, Arg117 was in contact with **M12**, but no hydrogen bond was formed between **M12** and the contacting residues of CYP2J2. In the present CYP2J2 model, Arg117 was found to stably interact with Glu222 by forming a salt-bridge, with an average distance between the oxygen atom of the carboxylic acid group of Glu222 and the nitrogen atom of the guanidine moiety of Arg117 being  $3.7 \pm 0.1$  Å. However, the distance from the C $\beta$  to the heme iron (3.4 Å) was larger than the distance from the C $\gamma$  atom to the heme iron (2.8 Å), which seemed to contradict the hydroxylation preference of C $\beta$  in **M12** by CYP2J2 according to the distance criterion. It should be noted that C $\gamma$  could still move freely around C $\beta$ , which may blur their distance difference relative to the heme





**Figure 9.** Summary of the preferred binding pose for **M4–M14** in the active site of CYP2J2. For **M4**, **M8**, and **M14**, multiple binding poses corresponding to different metabolic pathways were shown. Residues in contact with the substrate were colored based on their types: blue for basic, red for acidic, green for polar, and white for nonpolar residues. The unit for the distance is Å. This figure was rendered using VMD<sup>27</sup> (<http://www.ks.uiuc.edu/Research/vmd/>).

iron. Thus, we proposed that, apart from the distance measurement, the chemical property of **M12** should also be considered. Because of the delocalization effect of benzene moiety, the C-H bond strength increases from Ca to Cy. The calculated bond order for different Ca/β/γ-H bonds in terms of the natural bond order analysis with Gaussian 16<sup>47,48</sup> showed that the average bond order was 0.9781, 0.9833, and 0.9918 for Ca-H, Cβ-H, and Cγ-H bond, respectively. Therefore, although Cγ was closer to the heme iron, the Cγ-H bond was stronger than the Cβ-H bond. Taken together, it appeared that Cβ achieved a desirable compromise between the distance and bond strength, which could contribute to its regioselectivity of the hydroxylation by CYP2J2.

The last substrate discussed is the ARA (**M14**) whose primary epoxidation site seemed difficult to predict with the homology modeled structure of CYP2J2<sup>21</sup>. The present docking study demonstrated that the distance between the preferred metabolic site (14, 15-) and the heme iron was shorter than other metabolic sites, which

was in line with the experiment finding that the 14,15-olefin (37% of total EET products) was the preferential product metabolized by CYP2J2<sup>6</sup>. However, it remained difficult to directly relate the preference of other products including 8,9-EET (24%), 5,6-EET (21%), and 11,12-EET (18%)<sup>6</sup> with the corresponding distances of (5.4 Å, 6.6 Å), (5.0 Å, 5.5 Å), and (5.0 Å, 5.6 Å). Other factors such as the interacting residues and the conformations of **M14** in the active site might also contribute to the difference in the ARA regioselectivity mediated by CYP2J2, which needs further investigation.

## Discussion

In this work, a new structural model of CYP2J2 was developed based on the multiple-template homology modeling approach. Note that the sequence identity between the templates and CYP2J2 was only about 40% (Fig. S11), a long-time relaxation under more realistic conditions seemed necessary. Represented with the improved CHARMM36m force field parameters<sup>49</sup>, the CYP2J2 model was fully relaxed in explicit solvent by running 500-ns MD simulations. To assess the sensitivity of this model to substrates, three chemically similar probes **M1–M3** were docked to the representative conformation of CYP2J2, and the obtained initial binding poses were subjected to 500-ns MD simulations. The simulation results suggested that **M1–M3** tended to adopt extended conformations in the binding pocket of CYP2J2, similar to their behavior in solvent. The induced-fit binding of these probes enabled **M2** and **M3** to change their conformations until a favorable binding pose was adopted by each of them. The highest binding affinity predicted for **M2** may have implications for its strongest fluorescence response to CYP2J2. Besides the distance measurement to dictate the catalytic reactivity, the present results highlighted the important role of electrostatic potential around the substrates in predicting the catalytic efficiency of CYP2J2. We also showed that structurally similar **M1–M3** resulted in divergent conformational dynamics of CYP2J2, and such conformational diversity may contribute to its capability of accommodating and metabolizing substrates with varying sizes and shapes.

From the interactions of **M1–M3** with CYP2J2, five hydrophobic residues (Ile127, Phe310, Ala311, Val380, and Ile487), were identified to restrict **M1–M3** access to the heme. Compared to the conformation of free CYP2J2, the hydrophobic clamp formed by these residues seemed more stable in the interactions of **M2** with CYP2J2, suggesting that maintenance of this topology may be related to the catalytic property of CYP2J2. To confirm that these residues acted as molecular determinant for substrate binding and specificity, a benchmark testing was carried out for a set of 11 substrates of CYP2J2 (**M4–M14**). Except for danazol (**M6**) and one of binding poses of ketoconazole (**M8**) where Phe310 was not involved, the five key residues were all in contact with the bound substrates. Although the present docking studies predicted the preferred binding pose(s) for these substrates, it remained challenging to predict the regioselectivity of substrates such as terfenadone derivative (**M12**) and ARA (**M14**) based on docking with single conformation, as well the distance criterion used to evaluate the catalytic reactivity. The physico-chemical property of a substrate obtained by quantum calculations could also be useful to elucidate the substrate metabolism.

The present CYP2J2 model provided a structural basis for the investigation of the metabolism of other endogenous or exogenous compounds. It should be noted that most reported CYPs crystal structures and structural models of CYP2J2 lack the N-terminal transmembrane helix, whereas the truncated forms preserve the catalytic activity of the enzymes<sup>50</sup>. In addition to function in a soluble form, CYPs are able to attach to membrane with its N-terminal, and interact with its redox partner CYP reductase in the membrane environment, which are essential for electron transfer in the catalytic cycle<sup>51</sup>. Interactions with membrane have been observed to modulate the conformational dynamics of human aromatase (CYP19A1) and alter the substrate/inhibitor access channels<sup>52</sup>. Computational simulations and experimental data have also revealed that interactions of CYP enzymes with partner reductase in the membrane could affect the substrate binding<sup>53,54</sup>. With the addition of the N-terminus of CYP2J2, the present structural model of CYP2J2 could be applied to study its interactions with membranes, as well as interactions with its partner in membranes.

## Methods

**Homology modeling and model relaxation.** The homology model of CYP2J2 was constructed based on the following templates: CYP2D6 (PDB ID: 2F9Q)<sup>55</sup>, CYP2A6 (PDB ID: 1Z10)<sup>56</sup>, CYP2B4 (PDB ID: 2BDM)<sup>57</sup>, CYP2C9 (PDB ID: 1OG5)<sup>58</sup>, and CYP2C8 (PDB ID: 1PQ2)<sup>59</sup>. These CYP2 enzymes display similar sequence identity (40.9%–42.6%) with that of CYP2J2 (Fig. S11). The multiple sequence alignment was carried out using the Clustal Omega web server<sup>60</sup>, and the result was shown in Fig. S12. The homology modeling was performed using the Modeller software (v9.17)<sup>61,62</sup>, and the model built with the lowest DOPE score was selected for further relaxation.

The model structure of CYP2J2 was fully relaxed by running 500-ns MD simulations in aqueous solution. The intramolecular interactions were represented with the CHARMM36m force field parameters<sup>49</sup>. The protein was solvated by the TIP3P water model in a rectangular box with the minimum distance between the protein and the box boundary being 15 Å. The salt concentration (NaCl) of the system was 0.15 M. The particle mesh Ewald method was used to treat long-range electrostatic interactions<sup>63</sup>. The system was minimized for 5000 steps, and equilibrated for 50 ps with the backbone and sidechain atoms restrained with a force constant of 400 kJ/(mol·nm<sup>2</sup>) and 40 kJ/(mol·nm<sup>2</sup>), respectively. The system temperature was kept at 303 K using the Nosé–Hoover thermostat<sup>64</sup>. The 500-ns production simulation was performed in the NPT ensemble at 303 K and 1 bar without restraints. The time-step was 2 fs and the pressure was maintained at 1 bar using the Parrinello–Rahman method<sup>65</sup>. MD simulations were performed using the GROMACS program (version 2018.4)<sup>66,67</sup>. The CHARMM-GUI web server was used to generate the input files for all simulations<sup>68,69</sup>.

**Molecular docking and MD simulation of protein–ligand interactions.** The representative conformation of CYP2J2 was selected from the last 100-ns trajectory. The average structure of CYP2J2 was first calculated from the last-100 ns trajectory. As this average conformation may not represent the real conformation of CYP2J2, we used this average structure as a reference to calculate the RMSDs for the conformations from the last 100-ns trajectory, and selected the conformation with the minimum RMSD. If two conformations showed the same RMSD, the conformation with the lower energy (see conformational energy calculation below) was selected and used for molecular docking. The representative conformation of CYP2J2 was shown in Fig. 2. The PROCHECK program<sup>70</sup> was used to assess the stereochemical quality of the protein structure. Nearly all residues (99.5%) of the CYP2J2 model were found in the allowed regions in Ramchandran plot, and 369 residues (91.1%) whose  $\Phi$ - $\Psi$  angles were in the most favored regions (Fig. S14), suggesting a good quality model.

The molecular structure of **M1/M2/M3** was represented with the CHARMM General Force Field (CGenFF)<sup>71</sup> along with the calculated RESP charges (Fig. S13)<sup>72</sup>. The structure of CYP2J2 was represented with the CHARMM36m force field parameters<sup>49</sup>. Molecular docking was performed to search the favorable binding poses of **M1/M2/M3** in the structure of CYP2J2. Autodock 4.2 program<sup>73,74</sup> was applied to dock **M1/M2/M3** to the ligand binding pocket above the heme motif. In the docking process, the CYP2J2 structure was modeled as rigid while the ligand structure was modeled as flexible. Lamarckian genetic algorithm<sup>75</sup> was applied to search a total of 100 potential binding poses in CYP2J2. The optimal ligand conformer was selected based on the following criteria: (1) The distance between the metabolic site of the ligand and the heme iron atom should be as short as possible; and (2) and if there were multiple conformers showing the same minimum distance, the conformer with the lower binding energy estimated in terms of the docking score was selected. Molecular docking of **M4–M14** to the binding site of CYP2J2 was also performed.

The initial binding conformers of **M1–M3** were shown in Figs. 3–5, respectively. For **M2**, a second binding conformer that was similar to **M1** was also chosen. For **M3**, the present docking study was unable to find the preferred binding pose that displayed the minimum distance between the metabolic site and the heme iron, which was probably because the binding pocket in the representative conformation of CYP2J2 was not large enough to accommodate **M3** in the extended conformation. Therefore, we selected a binding conformer of **M3** that was similar to the second binding pose of **M2** (Fig. S2). MD simulations of each protein–ligand complex were performed following the same procedure as used in the simulations of CYP2J2. The production simulations lasted for 500 ns for each system, and the last 100- or 50-ns trajectory was used for further analyses. Note that in both docking and MD simulations, we used the same CHARMM force field parameters<sup>49</sup> for protein and CGenFF parameters<sup>71</sup> (with the RESP partial charges) for the ligands.

**Evaluation of protein stability and protein–ligand binding affinity.** The relative stability of CYP2J2 in complex with **M1/M2/M3** was evaluated by the conformational energy, which was calculated using the generalized Born using molecular volume (GBMV) implicit solvent model implemented in the CHARMM (v44b1) program<sup>76</sup>. The single point energy was calculated after a 200-step minimization of each conformation using the GBMV II algorithm<sup>77,78</sup>. Other energy terms including bonded energy, van der Waals energy, electrostatic energy, and solvation energy were also obtained with the GB implicit solvent model. The block average method was used to estimate the mean values and standard deviations.

The binding free energy of **M1/M2/M3** to CYP2J2 was calculated using the *g\_mmpbsa* program<sup>79</sup> which implements the molecular mechanics Poisson–Boltzmann surface area (MM-PBSA) method<sup>80,81</sup> to predict the binding affinity for protein–ligand complex. The entropy contribution was not included in the current binding energy calculations. The default bootstrap method was used to estimate the mean values and standard deviations.

## Data availability

The datasets generated and analyzed during the current study are available from the corresponding author on reasonable request. Structures of the modeled CYP2J2 (in pdb format) and **M1–M14** (in mol2 format) are available with the published manuscript.

Received: 13 October 2020; Accepted: 7 December 2020

Published online: 17 December 2020

## References

- Denisov, I. G., Makris, T. M., Sligar, S. G. & Schlichting, I. Structure and chemistry of cytochrome P450. *Chem. Rev.* **105**, 2253–2278. <https://doi.org/10.1021/cr0307143> (2005).
- Gay, S. C., Roberts, A. G. & Halpert, J. R. Structural features of cytochromes P450 and ligands that affect drug metabolism as revealed by X-ray crystallography and NMR. *Future Med. Chem.* **2**, 1451–1468. <https://doi.org/10.4155/fmc.10.229> (2010).
- Guengerich, F. P., Waterman, M. R. & Egli, M. Recent structural insights into cytochrome P450 function. *Trends Pharmacol. Sci.* **37**, 625–640. <https://doi.org/10.1016/j.tips.2016.05.006> (2016).
- Lamb, D. C. & Waterman, M. R. Unusual properties of the cytochrome P450 superfamily. *Philos. Trans. R. Soc. B: Biol. Sci.* **368**, 20120434. <https://doi.org/10.1098/rstb.2012.0434> (2013).
- Nebert, D. W. & Russell, D. W. Clinical importance of the cytochromes P450. *Lancet* **360**, 1155–1162. [https://doi.org/10.1016/S0140-6736\(02\)11203-7](https://doi.org/10.1016/S0140-6736(02)11203-7) (2002).
- Wu, S., Moomaw, C. R., Tomer, K. B., Falck, J. R. & Zeldin, D. C. Molecular cloning and expression of CYP2J2, a human cytochrome P450 arachidonic acid epoxygenase highly expressed in heart. *J. Biol. Chem.* **271**, 3460–3468. <https://doi.org/10.1074/jbc.271.7.3460> (1996).
- Aliwarga, T., Evangelista, E., Sotoodehnia, N., Lemaitre, R. & Totah, R. Regulation of CYP2J2 and EET levels in cardiac disease and diabetes. *Int. J. Mol. Sci.* **19**, 1916. <https://doi.org/10.3390/ijms19071916> (2018).
- Solanki, M., Pointon, A., Jones, B. & Herbert, K. Cytochrome P450 2J2: Potential role in drug metabolism and cardiotoxicity. *Drug Metab. Dispos.* **46**, 1053–1065. <https://doi.org/10.1124/dmd.117.078964> (2018).

9. Das, A. *et al.* CYP2J2 molecular recognition: A new axis for therapeutic design. *Pharmacol. Therap.* **215**, 107601. <https://doi.org/10.1016/j.pharmthera.2020.107601> (2020).
10. Matsumoto, S., Hirama, T., Matsubara, T., Nagata, K. & Yamazoe, Y. Involvement of CYP2J2 on the intestinal first-pass metabolism of antihistamine drug, Astemizole. *Drug Metab. Dispos.* **30**, 1240–1245. <https://doi.org/10.1124/dmd.30.11.1240> (2002).
11. Liu, K.-H. *et al.* Characterization of ebastine, hydroxyebastine, and carebastine metabolism by human liver microsomes and expressed cytochrome P450 enzymes: Major roles for CYP2J2 and CYP3A. *Drug Metab. Dispos.* **34**, 1793–1797. <https://doi.org/10.1124/dmd.106.010488> (2006).
12. Lafite, P., André, F., Zeldin, D. C., Dansette, P. M. & Mansuy, D. Unusual regioselectivity and active site topology of human cytochrome P450 2J2. *Biochemistry* **46**, 10237–10247. <https://doi.org/10.1021/bi700876a> (2007).
13. Lee, C. A. *et al.* Identification of novel substrates for human cytochrome P450 2J2. *Drug Metab. Dispos.* **38**, 347–356. <https://doi.org/10.1124/dmd.109.030270> (2010).
14. Karkhanis, A. *et al.* Multiple modes of inhibition of human cytochrome P450 2J2 by dronedarone, amiodarone and their active metabolites. *Biochem. Pharmacol.* **107**, 67–80. <https://doi.org/10.1016/j.bcp.2016.03.005> (2016).
15. Wu, Z. *et al.* CYP2J2 and CYP2C19 are the major enzymes responsible for metabolism of alendazole and fenbendazole in human liver microsomes and recombinant P450 assay systems. *Antimicrob. Agents Chemother.* **57**, 5448–5456. <https://doi.org/10.1128/aac.00843-13> (2013).
16. Evangelista, E. A., Kaspera, R., Mokadam, N. A., Jones, J. P. & Totah, R. A. Activity, inhibition, and induction of cytochrome P450 2J2 in adult human primary cardiomyocytes. *Drug Metab. Dispos.* **41**, 2087–2094. <https://doi.org/10.1124/dmd.113.053389> (2013).
17. Lee, E., Wu, Z., Shon, J. C. & Liu, K.-H. Danazol inhibits cytochrome P450 2J2 activity in a substrate-independent manner. *Drug Metab. Dispos.* **43**, 1250–1253. <https://doi.org/10.1124/dmd.115.064345> (2015).
18. Lee, E. *et al.* Terfenadone is a strong inhibitor of CYP2J2 present in the human liver and intestinal microsomes. *Drug Metab. Pharmacokinet.* **33**, 159–163. <https://doi.org/10.1016/j.dmpk.2018.02.001> (2018).
19. Li, W. *et al.* Probing ligand binding modes of human cytochrome P450 2J2 by homology modeling, molecular dynamics simulation, and flexible molecular docking. *Proteins* **71**, 938–949. <https://doi.org/10.1002/prot.21778> (2008).
20. Cong, S., Ma, X.-T., Li, Y.-X. & Wang, J.-F. Structural basis for the mutation-induced dysfunction of human CYP2J2: A computational study. *J. Chem. Inf. Model.* **53**, 1350–1357. <https://doi.org/10.1021/ci400003p> (2013).
21. Proietti, G., Abelak, K. K., Bishop-Bailey, D., Macchiarulo, A. & Nobeli, I. Computational modelling of the binding of arachidonic acid to the human monooxygenase CYP2J2. *J. Mol. Model.* **22**, 279. <https://doi.org/10.1007/s00894-016-3134-6> (2016).
22. Xia, X.-L., Fa, B.-T., Cong, S., Wang, J.-F. & Chou, K.-C. Research/review: Insights into the mutation-induced dysfunction of arachidonic acid metabolism from modeling of human CYP2J2. *Curr. Drug Metab.* **15**, 502–513. <https://doi.org/10.2174/1389200215666140908095647> (2014).
23. Ekroos, M. & Sjogren, T. Structural basis for ligand promiscuity in cytochrome P450 3A4. *Proc. Natl. Acad. Sci. USA* **103**, 13682–13687. <https://doi.org/10.1073/pnas.0603236103> (2006).
24. Sevrioukova, I. F. & Poulos, T. L. Structural basis for regiospecific midazolam oxidation by human cytochrome P450 3A4. *Proc. Natl. Acad. Sci. USA* **114**, 486–491. <https://doi.org/10.1073/pnas.1616198114> (2017).
25. Ning, J. *et al.* Molecular design strategy to construct the near-infrared fluorescent probe for selectively sensing human cytochrome P450 2J2. *J. Am. Chem. Soc.* **141**, 1126–1134. <https://doi.org/10.1021/jacs.8b12136> (2019).
26. Kim, J.-H., Choi, W.-G., Lee, S. & Lee, H. Revisiting the metabolism and bioactivation of ketoconazole in human and mouse using liquid chromatography-mass spectrometry-based metabolomics. *Int. J. Mol. Sci.* **18**, 621. <https://doi.org/10.3390/ijms18030621> (2017).
27. Humphrey, W., Dalke, A. & Schulten, K. V. M. D. Visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38. [https://doi.org/10.1016/0263-7855\(96\)00018-5](https://doi.org/10.1016/0263-7855(96)00018-5) (1996).
28. Schmidtke, P., Bidon-Chanal, A., Luque, F. J. & Barril, X. MDpocket: Open-source cavity detection and characterization on molecular dynamics trajectories. *Bioinformatics* **27**, 3276–3285. <https://doi.org/10.1093/bioinformatics/btr550> (2011).
29. Csermely, P., Palotai, R. & Nussinov, R. Induced fit, conformational selection and independent dynamic segments: an extended view of binding events. *Trends Biochem. Sci.* **35**, 539–546. <https://doi.org/10.1016/j.tibs.2010.04.009> (2010).
30. Mobley, D. L. & Dill, K. A. Binding of small-molecule ligands to proteins: “What You See” is not always “What You Get”. *Structure* **17**, 489–498. <https://doi.org/10.1016/j.str.2009.02.010> (2009).
31. Lafite, P. *et al.* Role of Arginine 117 in substrate recognition by human cytochrome P450 2J2. *Int. J. Mol. Sci.* **19**, 2066. <https://doi.org/10.3390/ijms19072066> (2018).
32. Liu, R. *et al.* Determinants of the inhibition of DprE1 and CYP2C9 by antitubercular thiophenes. *Angew. Chem. Int. Ed.* **56**, 13011–13015. <https://doi.org/10.1002/anie.201707324> (2017).
33. The PyMOL Molecular Graphics System, Version 1.8.x, Schrödinger, LLC.
34. Baker, N. A., Sept, D., Joseph, S., Holst, M. J. & McCammon, J. A. Electrostatics of nanosystems: Application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. USA* **98**, 10037–10041. <https://doi.org/10.1073/pnas.181342398> (2001).
35. Wade, R. C., Gabdouline, R. R., Ludemann, S. K. & Lounnas, V. Electrostatic steering and ionic tethering in enzyme-ligand binding: Insights from simulations. *Proc. Natl. Acad. Sci. USA* **95**, 5942–5949. <https://doi.org/10.1073/pnas.95.11.5942> (1998).
36. Bauer, M. R. & Mackey, M. D. Electrostatic complementarity as a fast and effective tool to optimize binding and selectivity of protein–ligand complexes. *J. Med. Chem.* **62**, 3036–3050. <https://doi.org/10.1021/acs.jmedchem.8b01925> (2019).
37. Grant, B. J., Rodrigues, A. P. C., ElSawy, K. M., McCammon, J. A. & Caves, L. S. D. Bio3d: an R package for the comparative analysis of protein structures. *Bioinformatics* **22**, 2695–2696. <https://doi.org/10.1093/bioinformatics/bt461> (2006).
38. Skjærven, L., Yao, X.-Q., Scarabelli, G. & Grant, B. J. Integrating protein structural dynamics and evolutionary analysis with Bio3D. *BMC Bioinform.* **15**, 399. <https://doi.org/10.1186/s12859-014-0399-6> (2014).
39. Hünenberger, P. H., Mark, A. E. & van Gunsteren, W. F. Fluctuation and cross-correlation analysis of protein motions observed in nanosecond molecular dynamics simulations. *J. Mol. Biol.* **252**, 492–503. <https://doi.org/10.1006/jmbi.1995.0514> (1995).
40. Amadei, A., Ceruso, M. A. & Di Nola, A. On the convergence of the conformational coordinates basis set obtained by the essential dynamics analysis of proteins’ molecular dynamics simulations. *Proteins* **36**, 419–424. [https://doi.org/10.1002/\(sici\)1097-0134\(19990901\)36:4%3c419::aid-prot5%3e3.0.co;2-u](https://doi.org/10.1002/(sici)1097-0134(19990901)36:4%3c419::aid-prot5%3e3.0.co;2-u) (1999).
41. David, C. C. & Jacobs, D. J. Characterizing protein motions from structure. *J. Mol. Graph. Model.* **31**, 41–56. <https://doi.org/10.1016/j.jmgm.2011.08.004> (2011).
42. Urban, P., Lautier, T., Pompon, D. & Truan, G. Ligand access channels in cytochrome P450 enzymes: A review. *Int. J. Mol. Sci.* **19**, 1617. <https://doi.org/10.3390/ijms19061617> (2018).
43. Petřek, M. *et al.* CAVER: A new tool to explore routes from protein clefts, pockets and cavities. *BMC Bioinform.* **7**, 316. <https://doi.org/10.1186/1471-2105-7-316> (2006).
44. Shah, M. B. *et al.* Conformational adaptation of human cytochrome P450 2B6 and rabbit cytochrome P450 2B4 revealed upon binding multiple amlodipine molecules. *Biochemistry* **51**, 7225–7238. <https://doi.org/10.1021/bi300894z> (2012).
45. Jeong, D. *et al.* Terfenadine metabolism of human cytochrome P450 2J2 containing genetic variations (G312R, P351L and P115L). *Drug Metab. Pharmacokinet.* **33**, 61–66. <https://doi.org/10.1016/j.dmpk.2017.10.004> (2018).
46. Lee, S. S. *et al.* Identification and functional characterization of novel CYP2J2 variants: G312R variant causes loss of enzyme catalytic activity. *Pharmacogenet. Genomics* **15**, 105–113. <https://doi.org/10.1097/01213011-200502000-00006> (2005).
47. Frisch, M. J. *et al.* *Gaussian 16* (R. A. Gaussian Inc, Wallingford CT, 2016).

48. Glendening, E. D., Reed, A. E., Carpenter, J. E. & Weinhold, F., NBO (version 3.1).
49. Huang, J. *et al.* CHARMM36m: An improved force field for folded and intrinsically disordered proteins. *Nat. Methods* **14**, 71–73. <https://doi.org/10.1038/nmeth.4067> (2016).
50. Šrejber, M. *et al.* Membrane-attached mammalian cytochromes P450: An overview of the membrane's effects on structure, drug binding, and interactions with redox partners. *J. Inorg. Biochem.* **183**, 117–136. <https://doi.org/10.1016/j.jinorgbio.2018.03.002> (2018).
51. Meling, D. D., McDougle, D. R. & Das, A. CYP2J2 epoxigenase membrane anchor plays an important role in facilitating electron transfer from CPR. *J. Inorg. Biochem.* **142**, 47–53. <https://doi.org/10.1016/j.jinorgbio.2014.09.016> (2015).
52. Di Nardo, G. *et al.* Working at the membrane interface: Ligand-induced changes in dynamic conformation and oligomeric structure in human aromatase. *Biotechnol. Appl. Biochem.* **65**, 46–53. <https://doi.org/10.1002/bab.1613> (2018).
53. Ritacco, I., Saltalamacchia, A., Spinello, A., Ippoliti, E. & Magistrato, A. All-atom simulations disclose how cytochrome reductase reshapes the substrate access/egress routes of its partner CYP450s. *J. Phys. Chem. Lett.* **11**, 1189–1193. <https://doi.org/10.1021/acs.jpcllett.9b03798> (2020).
54. Zhang, C., Catucci, G., Di Nardo, G. & Gilardi, G. Effector role of cytochrome P450 reductase for androstenedione binding to human aromatase. *Int. J. Biol. Macromol.* **164**, 510–517. <https://doi.org/10.1016/j.ijbiomac.2020.07.163> (2020).
55. Rowland, P. *et al.* Crystal structure of human cytochrome P450 2D6. *J. Biol. Chem.* **281**, 7614–7622. <https://doi.org/10.1074/jbc.M511232200> (2006).
56. Yano, J. K., Hsu, M.-H., Griffin, K. J., Stout, C. D. & Johnson, E. F. Structures of human microsomal cytochrome P450 2A6 complexed with coumarin and methoxsalen. *Nat. Struct. Mol. Biol.* **12**, 822–823. <https://doi.org/10.1038/nsmb971> (2005).
57. Zhao, Y. *et al.* Structure of microsomal cytochrome P450 2B4 complexed with the antifungal drug bifonazole. *J. Biol. Chem.* **281**, 5973–5981. <https://doi.org/10.1074/jbc.M511464200> (2006).
58. Williams, P. A. *et al.* Crystal structure of human cytochrome P450 2C9 with bound warfarin. *Nature* **424**, 464–468. <https://doi.org/10.1038/nature01862> (2003).
59. Schoch, G. A. *et al.* Structure of human microsomal cytochrome P450 2C8. *J. Biol. Chem.* **279**, 9497–9503. <https://doi.org/10.1074/jbc.M312516200> (2004).
60. Madeira, F. *et al.* The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res.* **47**, W636–W641. <https://doi.org/10.1093/nar/gkz268> (2019).
61. Webb, B. & Sali, A. Comparative protein structure modeling using MODELLER. *Curr. Protoc. Bioinform.* **54**, 5.6.1–5.6.37. <https://doi.org/10.1002/cpbi.3> (2016).
62. Marti-Renom, M. A. *et al.* Comparative protein structure modeling of genes and genomes. *Annu. Rev. Biophys. Biomol. Struct.* **29**, 291–325. <https://doi.org/10.1146/annurev.biophys.29.1.291> (2000).
63. Darden, T., York, D. & Pedersen, L. Particle mesh Ewald: An N<sup>3</sup>·log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089–10092. <https://doi.org/10.1063/1.464397> (1993).
64. Hoover, W. G. Canonical dynamics: Equilibrium phase-space distributions. *Phys. Rev. A* **31**, 1695–1697. <https://doi.org/10.1103/PhysRevA.31.1695> (1985).
65. Parrinello, M. & Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **52**, 7182–7190. <https://doi.org/10.1063/1.328693> (1981).
66. Abraham, M. J. *et al.* GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1–2**, 19–25. <https://doi.org/10.1016/j.softx.2015.06.001> (2015).
67. Bjelkmar, P., Larsson, P., Cuendet, M. A., Hess, B. & Lindahl, E. Implementation of the CHARMM Force Field in GROMACS: Analysis of protein stability effects from correction maps, virtual interaction sites, and water models. *J. Chem. Theory Comput.* **6**, 459–466. <https://doi.org/10.1021/ct900549r> (2010).
68. Lee, J. *et al.* CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field. *J. Chem. Theory Comput.* **12**, 405–413. <https://doi.org/10.1021/acs.jctc.5b00935> (2015).
69. Jo, S., Kim, T., Iyer, V. G. & Im, W. CHARMM-GUI: A web-based graphical user interface for CHARMM. *J. Comput. Chem.* **29**, 1859–1865. <https://doi.org/10.1002/jcc.20945> (2008).
70. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. PROCHECK: A program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283–291. <https://doi.org/10.1107/s0021889892009944> (1993).
71. Vanommeslaeghe, K. *et al.* CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J. Comput. Chem.* **31**, 671–690. <https://doi.org/10.1002/jcc.21367> (2010).
72. Wang, J., Cieplak, P. & Kollman, P. A. How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J. Comput. Chem.* **21**, 1049–1074. [https://doi.org/10.1002/1096-987x\(200009\)21:12%3c1049::aid-jcc3%3e3.0.co;2-f](https://doi.org/10.1002/1096-987x(200009)21:12%3c1049::aid-jcc3%3e3.0.co;2-f) (2000).
73. Forli, S. *et al.* Computational protein–ligand docking and virtual drug screening with the AutoDock suite. *Nat. Protocols* **11**, 905–919. <https://doi.org/10.1038/nprot.2016.051> (2016).
74. Morris, G. M. *et al.* AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J. Comput. Chem.* **30**, 2785–2791. <https://doi.org/10.1002/jcc.21256> (2009).
75. Morris, G. M. *et al.* Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J. Comput. Chem.* **19**, 1639–1662. [https://doi.org/10.1002/\(sici\)1096-987x\(199811\)19:14%3c1639::aid-jcc10%3e3.0.co;2-b](https://doi.org/10.1002/(sici)1096-987x(199811)19:14%3c1639::aid-jcc10%3e3.0.co;2-b) (1998).
76. Brooks, B. R. *et al.* CHARMM: The biomolecular simulation program. *J. Comput. Chem.* **30**, 1545–1614. <https://doi.org/10.1002/jcc.21287> (2009).
77. Lee, M. S., Salisbury, F. R. & Brooks, C. L. Novel generalized Born methods. *J. Chem. Phys.* **116**, 10606–10614. <https://doi.org/10.1063/1.1480013> (2002).
78. Feig, M. *et al.* Performance comparison of generalized born and Poisson methods in the calculation of electrostatic solvation energies for protein structures. *J. Comput. Chem.* **25**, 265–284. <https://doi.org/10.1002/jcc.10378> (2004).
79. Kumari, R., Kumar, R. & Lynn, A. g\_mmpbsa—A GROMACS Tool for high-throughput MM-PBSA calculations. *J. Chem. Inf. Model.* **54**, 1951–1962. <https://doi.org/10.1021/ci500020m> (2014).
80. Kollman, P. Free energy calculations: Applications to chemical and biochemical phenomena. *Chem. Rev.* **93**, 2395–2417. <https://doi.org/10.1021/cr00023a004> (1993).
81. Kollman, P. A. *et al.* Calculating structures and free energies of complex molecules: Combining molecular mechanics and continuum models. *Acc. Chem. Res.* **33**, 889–897. <https://doi.org/10.1021/ar000033j> (2000).

## Acknowledgements

The authors acknowledge support from the NIH (grant #GM121275), and the Texas Advanced Computing Center (TACC) at the University of Texas at Austin for providing HPC resources (Frontera and Longhorn) that have contributed to the research results reported within this paper.

### Author contributions

L.X. and L.C. designed the research; L.X. performed the simulations and analyzed the data; L.X. wrote the manuscript, and all authors revised the manuscript.

### Competing interests

The authors declare no competing interests.

### Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-020-79284-0>.

**Correspondence** and requests for materials should be addressed to L.X.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020