Human flavin-containing monooxygenase 3: Structural mapping of gene polymorphisms and insights into molecular basis of drug binding

Human flavin-containing monooxygenase 3: Structural mapping of gene polymorphisms and insights into molecular basis of drug binding

Chongliang Gao, Gianluca Catucci, Giovanna Di Nardo, Gianfranco Gilardi and Sheila J. Sadeghi *

Department of Life Sciences and Systems Biology, University of Torino, via Accademia Albertina 13, 10123 Torino, Italy.

*Corresponding author: S.J. Sadeghi, Department of Life Sciences and Systems Biology, Via Accademia Albertina 13, 10123 Torino, Italy.

Phone: +39-011-6704528

Fax: +39-011-6704643

e-mail: sheila.sadeghi@unito.it

Abstract

Human hepatic flavin-containing monooxygenase 3 (hFMO3) catalyses the monooxygenation of carbon-bound reactive heteroatoms and plays an important role in the metabolism of drugs and xenobiotics. Although numerous hFMO3 allelic variants have been identified in patients and their biochemical properties well-characterised *in vitro*, the molecular mechanisms underlying loss-of-function mutations have still not been elucidated due to lack of detailed structural information of hFMO3. Therefore, in this work a 3D structural model of hFMO3 was generated by homology modeling, evaluated by a variety of different bioinformatics tools, refined by molecular dynamics simulations and further assessed based on *in vitro* biochemical data. The molecular dynamics simulation results highlighted 4 flexible regions of the protein with some of them overlapping the data from trypsin digest. This was followed by structural mapping of 12 critical polymorphic variants and molecular docking experiments with five different known substrates/drugs of hFMO3 namely, benzydamine, sulindac sulphide, tozasertib, methimazole and trimethylamine. Localisation of these mutations on the hFMO3 model provided a structural explanation for their observed biological effects and docked models of hFMO3-drug complexes gave insights into their binding mechanism demonstrating that nitrogen- and sulfur-containing substrates interact with the isoalloxazine ring through Pi-Cation interaction and Pi-Sulfur interactions, respectively. Finally, the data presented give insights into the drug binding mechanism of hFMO3 which could be valuable not only for screening of new chemical entities but more significantly for designing of novel inhibitors of this important Phase I drug metabolising enzyme.

Keywords: Monooxygenase, 3D model, molecular dynamics, fish odour syndrome, molecular docking, polymorphism

1. Introduction

The mammalian flavin-containing monooxygenases (FMO) belong to the oxidoreducatase subclass 1.14 and comprise a family of five functional enzymes that catalyse the oxygenation of a wide variety of nucleophilic heteroatom-containing compounds and have specific endogenous biological roles (Cashman and Zhang, 2006). In humans, FMO3 (hFMO3) has been described as playing a prominent role in the catabolism of drugs by selective oxygenation of nitrogen- and sulfur-containing xenobiotics resulting in their detoxification via excretion (Cashman and Zhang, 2006). The latter reactions are part of human hepatic Phase I drug metabolism mainly carried out by cytochromes P450. However, unlike cytochromes P450, the hFMO3 reaction produces few toxic metabolites and therefore poses considerable clinical advantages (Cashman, 2005; 2008). Moreover, in addition to its role in drug metabolism, FMO3 has also been shown more recently to act as an important modifier of human health in atherosclerosis and cholesterol/glucose/lipid metabolism (Leiser et al., 2015; Schugar et al., 2015).

Certain mutations within the hFMO3 gene can lead to the abnormal drug or chemical metabolism. One such genetic disease, trimethylaminuria (TMAU or fish odour syndrome), is due to genetic polymorphisms in the hFMO3 gene that cause defective trimethylamine (TMA) N-oxygenation (Dolphin et al., 1997). In addition, mutations can also affect the substrate selectivity with some abolishing or lowering both N- or S-oxygenation capacity of the enzyme, while other mutations only abolish the N-oxygenation activity (Zhou and Shephard, 2006).

In spite of growing knowledge of genetic polymorphisms in hFMO3, the molecular mechanisms underlying the loss of function or substrate specificity of genetic variants is still unclear due to lack of detailed structural information for hFMO3. However, the crystal structures of a eukaryotic FMO (yFMO) from *Schizosaccharomyces pombe* (Eswaramoorthy et al., 2006) and a prokaryotic FMO

(bFMO) from *Methylophaga* sp. Strain SK1 (Alfieri et al., 2008) have been solved recently providing useful templates for the construction of a hFMO3 model. In previous years our group (Sadeghi et al., 2010) and others have published 3D models of hFMO3 (Borbás et al., 2006; Yeung et al., 2007) however, in most cases the templates selected were different. The goal of the present study is to build a hFMO3 model used for structural mapping of critical polymorphic variants to provide a structural explanation for their biological effects. In addition, molecular docking is performed to investigate the possible molecular basis of binding of different drugs.

2. Materials and methods

2.1. Sequence analysis and molecular modeling

The search and the identification of templates suitable for homology modeling were performed by BLAST (http://blast.ncbi.nlm.nih.gov/). The hFMO3 sequence (NCBI CCDS ID: CCDS1292.1) was defined as the target sequence, and the crystal structures of bFMO (PDB ID: 2VQ7) (Alfieri et al., 2008) and yFMO (PDB ID: 2GV8) (Eswaramoorthy et al., 2006) served as templates. Sequence identities and homologies to hFMO3 are 23.0% and 40.4% for bFMO and 21.5% and 37.5% for yFMO, respectively. Multiple sequence alignment was performed with ClustalW (Chenna, 2003) and ESPript 3.0 (Gouet et al., 1999). This alignment was used for the model construction by YASARA Model package (Krieger and Vriend, 2002). A long insertion of 70 amino acids (D226-G295) together with the extended C-terminal portion which could not align to the templates were constructed by Phyre2 server (http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index) and integrated into the final model by YASARA (Krieger and Vriend, 2002). The quality of the obtained hFMO3 structural model was evaluated using PROCHECK (http://services.mbi.ucla.edu/PROCHECK/), ProSA (https://prosa.services.came.sbg.ac.at/prosa.php) and ERRAT (Colovos and Yeates, 1993;

Wiederstein and Sippl, 2007).

2.2. Molecular dynamics (MD) simulation

The molecular dynamics simulation was performed by YASARA structure package (Krieger and Vriend, 2002). This software package contains all the functions needed to predict and validate macromolecular structures, including ligand docking and highly accurate force fields with knowledge-based potentials. Prior to MD simulation, the hFMO3 model was first optimized for its hydrogen bonding network to achieve more stable trajectories. The optimized structure was then placed in a simulation cell that was 10 Å larger than the protein. The force Field AMBER 03 was assigned and the cell was then filled with water molecules and NaCl at the physiological concentration of 0.9%. Counter ions were also placed in the cell, predicting pKa values and assigning protonation states and the temperature was set to 298K. The simulation cell was subsequently rescaled to obtain a density for water residues of 0.997 g/L. The initial structure was subjected to energy minimization performing a simulated annealing minimization which moves the structure to a nearby stable energy minimum. Initial atom velocities were set according to a Boltzmann distribution and the molecular dynamics was then followed for 5 nanoseconds recording snapshots at regular time intervals of every 25 picoseconds. The AMBER 03 force field was employed throughout all simulations. The cutoff for van der Waals interactions was 7.86 Å and the Particle Mesh Ewald (PME) algorithm was applied to compute the long range electrostatic interactions. The results of the molecular dynamics simulation were analysed with YASARA using an embedded pre-written script that, on the basis of the time course, is able to analyse structure stability creating a per residue table with average RMSFs and RMSDs from the starting structure.

2.3. Protein expression and purification

E. coli JM109 cells transformed with pJL2-hFMO3 were used for the expression of hFMO3 protein as previously described (Catucci et al., 2012). The cells were harvested 22 h post-induction and the protein was extracted from membrane fractions using 1% IGEPAL and purified by Ni affinity chromatography. The purified protein was collected and analysed by 12.5% SDS-PAGE. Protein concentration was determined by spectroscopy with the peak absorbance at 450 nm and an extinction coefficient of 11900 $M^{-1}cm^{-1}$ as described previously (Catucci et al., 2012).

2.4. Trypsin digestion

Trypsin digestion was carried out to analyse the banding patterns of peptides generated (Coulter-Mackie and Lian, 2008). Briefly, 116 µg hFMO3 in 80 µl phosphate buffer (pH 7.4) was mixed with 60 µg trypsin and incubated for 2, 5, 10, 15, 30, 60, 120 min at 37 °C. At each time interval, 10 µl of the reaction mixture was removed and mixed with 10 µl SDS loading buffer (60 mM Tris pH 6.8/2% SDS/12% glycerol/0.001% bromophenol blue) and incubated at 98 °C for 10 min. All samples collected in this manner were subsequently subjected to SDS-PAGE analysis.

2.5. Molecular docking

Autodock version 4.0 (Goodsell and Morris, 1996) incorporated into the YASARA Structure package was used for molecular docking studies. Five different known substrates of hFMO3, namely benzydamine, methimazole, sulindac sulfide, tozasertib and trimethylamine were docked into the substrate binding pocket of the hFMO3 model. Each of these substrates (ligands) were initially built and optimized using Discovery Studio Visualizer (Dassault Systèmes, 2016) prior to the docking experiments. First a global docking experiment was performed by executing a total of one hundred runs of Global Docking centring a 15 × 15 × 15 Å simulation cell on the FAD group. In this

experiment, the substrate is originally outside the simulation box and is placed inside the cell by exploiting the Autodock algorithm, resulting in a series of binding modes classified by the binding energy outputs. Among these binding modes the complex protein-ligand bearing the highest binding energy, calculated by YASARA as the mechanical energy required for disassembling a whole into separate parts (where positive energies indicate stronger binding and negative energies equate to no binding), was selected and further refined by local docking. After global docking the best binding mode (pose) was selected based on the best binding energy. The complexes were then subjected to 999 runs of Local Docking yielding the final docked binding modes. In local docking experiments the ligand is within the simulation cell and the possible conformations are assayed around the starting pose. The results are again sorted by binding energy, and the most suitable binding mode with the highest binding energy was finally selected and optimized by energy minimization using YASARA.

3. Results and Discussion

3.1. Sequence alignment and homology modeling

As mentioned earlier, currently there are only two FMO proteins with known crystal structures: one bacterial (bFMO) (Alfieri et al., 2008), the other from yeast (yFMO) (Eswaramoorthy et al., 2006). When these two protein structures are superimposed (Cα backbone) in spite of only 23.4% sequence identity, they present a highly conserved overall structure. The primary amino acid sequence of hFMO3 was aligned with the sequences of these two proteins and the FMO-identifying sequence motif together with the two Rossmann-fold motifs (Fraaije et al., 2002) were identified and are shown in Fig. 1. The latter alignment also demonstrates the lack of alignment within the insert region (S218-I292) and C-terminal portion of hFMO3 (V501-T532). Therefore, structural models of the insert region and the missing C-terminal region were built separately by Phyre2 based on homology-

searching program PSI-BLAST (position-specific iterated basic local alignment search tool). The result of multiple sequence alignment for insertion region shows this part is built using insertion regions of Baeyer-Villiger monooxygenases (PDB ID: 3GWD/96% confidence, 1W4X/96% confidence, 4AP3/96% confidence, 3UCL/95% confidence, 3UOY/95% confidence) as templates. The resulting secondary structures obtained are shown in Fig. 2B with 100% of residues of the insert region modelled with >90% confidence. While for the C-terminal portion, although 0% residues modelled at more than 90% confidence, a helix conformation was thought to be reasonable as reported previously (Sadeghi et al., 2010). The overall model, shown in Fig. 2A, was thus built based on the sequence alignment and integration of the insertion and C-terminal regions by YASARA Model package. Ramachandran Plot analysis (Figure S1 in the supporting Information) determined by PROCHECK reveals that 87.6%, 11.7% and 0.7% of the residues of the hFMO3 model are located in the most favoured, generously allowed and disallowed regions, respectively. The overall G-factor is 0.01, indicating a good quality model (acceptable values of the G-factor in PROCHECK are between 0 and −0.5, with the best models displaying values close to zero). The results of ERRAT, a method for discriminating between correctly and incorrectly determined fragments of protein structures based on characteristic atomic interaction, and ProSA (Figure S2 in the supporting Information) suggest that the quality of the hFMO3 structural model is satisfactory and can be used for further analysis. A summary of the hFMO3 model quality data gathered through different bioinformatics tools is shown in Table 1.

Table 1. Summary of the main parameters obtained from validation of hFMO3 homology model through PROCHECK, ProSA and ERRAT servers.

Ramachandran plot	

Residues in most favoured and allowed regions (%)	87.6
Residues in generously allowed regions (%)	11.7
Residues in disallowed regions (%)	0.7
G-factor (dihedrals)	-0.21
M/c bond lengths	
within limits (%)	100
highlighted (%)	0
M/c bond angles	
within limits (%)	96.6
highlighted (%)	3.4
Planar groups	
within limits (%)	89.9
highlighted (%)	10.1
G-factor (overall)	0.01
ProSA z-score	-7.73
ERRAT score	96.367

The hFMO3 model was further refined by 5 ns of MD simulations and reached structural stability after only 2 ns as shown in Figure 3A. The variation in RMSD for each Cα atom of the model was determined during the MD simulations and revealed the most flexible regions of the protein to be grouped in four different regions. These highly flexible regions namely V110-H150, V220-A250, P400-Y430 and P450-T532 are located on the solvent-accessible surface of the hFMO3 model as

shown in Fig. 3B. The location of these flexible regions will be discussed later and put into context with the *in vitro* experimental data presented in the trypsin digestion section.

The final 3D generated model of hFMO3 showed two distinct domains, a smaller $NADP^+$-binding domain with a larger FAD-binding domain connected through a double linker, similar to the bacterial and yeast FMO structures ((Eswaramoorthy et al., 2006; Alfieri et al., 2008) with the exception of the extra insert region (S218-I292). The hFMO3 model also shows very similar active site to those of the two crystallised bacterial and yeast FMOs.

3.2. Structural mapping of the critical hFMO3 allelic variants

To further evaluate the hFMO3 model and provide possible structural rationalization for the biological effects of the polymorphisms in hFMO3, twelve critical amino acids that have been previously well characterised (Dolphin et al., 2000; Zhang et al., 2003;Yamazaki et al., 2007; Yeung et al., 2007; Shimizu et al., 2012) (Table 2) were mapped onto the model and are shown in Fig. 2A. According to the hFMO3 model, residues: P70, P153, E158, R205, V257, E308, G421 and T488 are located at some distance from FAD and TMA binding pocket. Mutations at these sites may be expected to have little or no effect on K_m for TMA, which fits reasonably well with the *in vitro* results shown in Table 2. Besides, our model shows that P153 is located on the edge of substrate binding pocket and may affect ingress/egress of substrates (Fig. 2D). Therefore, the greatly reduced k_{cat} of P153L mutant can be readily rationalised. While for E32K which completely abolishes the catalytic activity of hFMO3 for TMA (Table 2), although not located in the catalytic pocket, E32 stabilizes FAD by hydrogen bonds (Fig. 2C). Therefore, the inactivity of this mutant is likely due to the poor retention of the FAD cofactor in the folded protein. As shown in Fig. 2D, V58, N61 and S195 are located in the catalytic pocket, but only N61 interacts directly with TMA through hydrogen bonding (Fig. 5D). This could

explain the greatly improved K_m of N61S mutant. On the other hand, V58 and S195 do not make any direct contact with TMA (Fig. 2D and Fig. 4D) and therefore, although located in the TMA binding pocket, V58I and S195L have little effect on the measured K_m value as summarised in Table 2. Furthermore, the model shows that S195 stabilizes bound $NADP^+$ through hydrogen bonding interactions (Fig. 2C). Since oxidized $NADP^+$ remains tightly bound and is involved in the stabilisation of the C4a-hydroperoxy-FAD intermediate that is crucial for substrate oxygenation (Alfieri et al., 2008; Orru et al., 2010), S195L mutation may destabilise the $NADP^+$ binding and accelerate the decay of the intermediate, which is in good agreement with the dramatically reduced k_{cat} for TMA (Table 2).

The above analyses demonstrate that the 3D generated model of hFMO3 can be reliable as the locations of the selected variants on the model fits reasonably well with the changes reported previously by *in vitro* measurements of their kinetic constants (Table 2).

Table 2. The ratios of kinetic constants of variants to the wild-type hFMO3

Variant	$K_{m(variant)}/K_{m(Wt)}$	$k_{cat(variant)}/k_{cat(Wt)}$	Reference
E32K	0	0	Zhang et al., 2003
V58I	0.92	0.66	Shimizu et al., 2012
N61S	6.61	0.24	Dolphin et al., 2000
P70L	0.92	0.51	Shimizu et al., 2012
P153L	0.97	0.09	Yeung et al., 2007
E158K	0.87	1.17	Yeung et al., 2007
S195L	1	0.06	Shimizu et al., 2012
R205C	1	0.26	Yamazaki et al., 2007
V257M	0.87	1.43	Yeung et al., 2007
E308G	0.94	0.78	Yeung et al., 2007
G421V	0.97	0.55	Shimizu et al., 2012
T488A	0.91	0.54	Shimizu et al., 2012

3.3. Trypsin digestion

In general, limited proteolysis is usually restricted to a few specific locations on the surface of any given protein. Previous studies on proteins with known crystal structures have demonstrated that the cleavage sites are mostly at the level of loops and not alpha helices (Fontana et al., 2004). Therefore, limited proteolysis experiments can be used to identify the regions with highest flexibility within a folded protein. Limited proteolysis can be carried out by trypsin digestion, as a useful measure of protein folding where the crystal structure is unknown. It can cleave proteins at carboxyl side of the trypsin-accessible lysine and arginine residues. Here, trypsin digestion was used to evaluate the homology model generated by analysing the peptide patterns generated over time. Fig. 4A shows the trypsin digestion results, in particular the presence of a new peptide band around 47.0 kDa together with several smaller molecular weight bands mainly ranging from 27 to 31 kDa. The possible lysine or arginine cleavage sites of trypsin together with their corresponding observed bands are summarised in Table 3. The cleavage sites resulting in the new band around 47.0 kDa may be formed by digestion at several contiguous residues including K415, K416, R417 or K418. While peptide bands between 27 kDa and 31 kDa may be the result of digestion at K265, R263, R279 and K280 which form peptides with 28.9 kDa, 28.7 kDa, 30.4 kDa and 30.5 kDa, respectively. The location of these lysine and arginine residues are shown in Fig. 4B-C which highlight the fact that these amino acids are on the surface, exposed to the solvent and readily accessible to trypsin. Therefore, these data give further experimental proof in support of the accuracy of the hFMO3 model.

Table 3. The possible trypsin cleavage sites and their corresponding digestion bands produced in SDS-PAGE.

Cleavage site	Digestion Bands	
R263	30.3 kDa	29.7 kDa

K265	29.9 kDa	29.1 kDa
R279	30.5 kDa	28.5 kDa
K280	31.6 kDa	28.4 kDa
K415-K418	~ 47 kDa	~ 13 kDa

Moreover, in order to better understand the connection between the surface exposure and the flexibility of the protein structure, the results of MD simulation were compared to the trypsin digestion patterns. Each residue of the model was evaluated in terms of root mean square fluctuation (RMSF) throughout the MD simulation. The analysis revealed the presence of four highly flexible regions (V110-H150, V220-A250, P400-Y430 and P450-T532) located on the solvent-accessible surface as shown in Fig. 3B. Of these, V220-A250 is located in the insert region and close to trypsin-accessible sites R279 and K280. The other three highly flexible regions are located in the FAD binding domain, and one of these, P400-Y430, contains cleavage sites K415-K418 of trypsin. The high flexibility and solvent-accessibility of K415-K418 make it an easy target for trypsin attack, which fits quite well with the results of trypsin digestion.

3.4. Binding mode analysis

To provide insights into the molecular basis of substrate binding, YASARA Structure package was used for the docking experiments and the results were analysed by DS Visualizer. Structure-based computer modelling of drug-enzyme interactions is one of the most important pillars of modern drug discovery (Charifson and Kuntz, 1997). The two main aims of these *in silico* binding studies are to firstly obtain reliable estimates of the affinities of the selected drugs to the protein and secondly to

analyse their mechanism of interaction (Hulme and Trevethick, 2010). To this end, the refined hFMO3 model served as receptor and several well-characterised drugs/substrates were used as ligands. The five ligands used for molecular docking were prepared in DS Visualizer and optimized using Clean Geometry function (Fig. 5A) and include benzydamine (Castrignanò et al., 2010), a nonsteroidal anti-inflammatory drug; methimazole, a drug used to treat hyperthyroidism; sulindac sulfide (Castrignanò et al., 2012), a non-steroidal anti-inflammatory drug that reduces tumour burden in human colorectal cancer; tozasertib (Catucci et al., 2013), a pan-Aurora kinase inhibitor for cancer treatment and, trimethylyamine, an organic compound found in dietary food with a "fishy" odour which if not metabolised by hFMO3 causes the fish odour syndrome (Dolphin et al., 1997). The best binding modes of each of these drugs with the highest binding energies were selected and optimized by energy minimization for further analysis. The docking results were analysed using non-bond interactions monitor function by DS Visualizer and are shown in Fig.5B-F. The calculated theoretical binding energies for each of the five substrates tested are also summarized in Table 4.

Table 4. Substrate binding energy for hFMO3 model calculated by YASARA

Substrate	Binding energy (kcal/mol)
Tozasertib	5.74
Sulindac sulfide	5.61
Benzydamine	4.29
Methimazole	3.64
Trimethylamine	1.89

As can be seen in the table, the binding energies range between 1.89 to 5.74 kcal/mol and follow the order Tozasertib > sulindac sulfide > benzydamine > methimazole > trimethylamine. Unfortunately, these theoretical binding energies cannot be compared to experimental results due to lack of any such reported data within the literature.

The docking results demonstrate that hydrogen bonds, electrostatic and hydrophobic interactions all play important roles in substrate binding. The non-bond interactions are identified by DS Visualizer based on mainly interatomic distance and angle constraints. Fig. 5B-F shows how residues in the catalytic pocket stabilise the substrates by non-bond interactions. The docking result of benzydamine/hFMO3 complex (Fig. 5B) shows that N61, G376 and N194 interact with substrate through non classical hydrogen bonds (Carbon Hydrogen Bonds (Pierce et al., 2002) and Pi-Donor Hydrogen Bonds (Bissantz et al., 2010)). For the binding mode of tozasertib (Fig. 5C), the substrate is mainly stabilised by the conventional hydrogen bond interactions (Bissantz et al., 2010) with Y331 and L375. In the case of tozasertib, it also forms carbon hydrogen bonds (a kind of weaker hydrogen bond) with G330 and G376 and, a Pi-donor hydrogen bond with N194. Finally, for the trimethylamine/hFMO3 binding mode (Fig. 5D), the substrate only interacts with N61 and G376 through carbon hydrogen bonds.

The docking results also show the interactions of sulfur-containing substrates, methimazole and sulindac sulfide, with their surrounding amino acid residues. As shown in Fig. 5E, a Pi-donor hydrogen bond with N194 and conventional hydrogen bonds with N61 contribute to the stabilisation of methimazole in the catalytic pocket. In the case of sulindac sulfide (Fig. 5F), this substrate interacts with H149, G193 and N194 through Pi-Sigma (Ozawa et al., 2008), carbon hydrogen bond and Pi-donor hydrogen bond interactions, respectively.

Substrate binding modes also indicate that FAD plays a critical role in the stabilisation of substrates. Fig. 5B-D show that the N-oxidation sites of benzydamine, tozasertib and trimethylamine interact with the isoalloxazine ring through Pi-cation interactions (Gallivan and Dougherty, 1999), a kind of electrostatic interaction. In addition, benzydamine and trimethylamine also form Pi-Sigma interactions with the isoalloxazine ring. While for the S-oxidation sites of methimazole and sulindac

sulfide, the interaction with the isoalloxazine ring of FAD cofactor is mainly through Pi-sulfur interactions (Fig. 5E-F) (Ringer et al., 2007). In the case of methimazole, as seen in Fig. 5F, there is the additional Pi-stacking interaction (McGaughey et al., 1998) which further contributes to the stabilisation of the binding of this substrate.

Taken together, the five different substrates in this study interact with residues around the binding site mainly through weak non-bond interactions. The nitrogen-containing substrates interact with the isoalloxazine ring through Pi-cation interaction, while the sulfur-containing substrates form Pi-sulfur interactions with the isoalloxazine ring. In the case of methimazole binding to hFMO3, the molecular docking results presented in this work, with Pi-stacking interactions between this substrate and the FAD cofactor, are very similar to the crystallography data of the co-crystals of methimazole and yeast FMO (PDB ID: 2GVC) (Eswaramoorthy et al., 2006) and also indole/FAD interaction in the bacterial FMO (PDB ID:2XVJ) (Cho et al., 2011). The latter observation is in support of the suitability of the hFMO3 model for *in silico* molecular docking applications.

4. Conclusions

In this work, a three-dimensional model of hFMO3 was generated by homology modelling with the help of YASARA Model package and Phyre2 using the only two available crystal structures of the prokaryotic and yeast FMOs. The model was further refined by 5 ns of MD simulations and evaluated by a variety of different bioinformatics tools, structural mapping of some critical mutation sites and *in vitro* trypsin digestion. The *in vitro* trypsin digestion results were in good agreement with the location of the four surface exposed flexible regions highlighted by the MD simulations and structural mapping of polymorphic variants provide the structural rational for their biological effects. In addition, docking capability of the model was assessed by docking of five important drugs/substrates of this

human hepatic enzyme. The best binding pose of each of the latter substrates was in agreement with the enzymatic product i.e. either N- or S-oxide. The calculated theoretical binding energies could not be directly compared to experimental data due to lack of any such reported data within the literature. The predicted binding modes of nitrogen- and sulfur-containing substrates could be useful for better understanding of their binding mechanism. Finally, the model presented is reliable in terms of 3D structure and stereochemistry and can therefore be used in not only screening of new chemical entities but also in designing novel inhibitors of this enzyme which currently are not available.

Supplementary material

There are two supplementary figures.

Acknowledgments

Chongliang Gao is the recipient of a three-year PhD scholarship from University of Torino for international students.

Disclosure statement

The authors declare no conflict of interest.

References

Alfieri, A., Malito, E., Orru, R., Fraaije, M.W., and Mattevi A. (2008) Revealing the moonlighting role of NADP in the structure of a flavin-containing monooxygenase. *Proceedings of the National Academy of Sciences, 105,* 6572-6577.

Bissantz, C., Kuhn, B., Stahl, M. (2010) A Medicinal Chemist's Guide to Molecular Interactions. *Journal of Medicinal Chemistry, 53,* 5061-5084.

Borbás, T., Zhang, J., Cerny, M. A., Likó, I. and Cashman, J.R. (2006) Investigation of structure and function of a catalytically efficient variant of the human flavin-containing monooxygenase form 3. *Drug Metabolism Disposition, 34,* 1995-2002.

Cashman, J. R. (2005) Some distinctions between flavin-containing and cytochrome P450

monooxygenases. *Biochemical and Biophysical Research Communications, 338*, 599–604.

Cashman, J. R. and Zhang, J. (2006) Human Flavin-Containing Monooxygenases. *Annual Review of Pharmacology and Toxicology, 46,* 65-100.

Cashman, J. R. (2008) Role of flavin-containing monooxygenase in drug development. *Expert Opinion in Drug Metabolism and Toxicology, 4,* 1507−1521.

Castrignanò, S., Sadeghi, S. J. and Gilardi, G. (2010) Electro-catalysis by immobilized humanflavin-containing monooxygenase isoform 3 (hFMO3). *Analytical and Bioanalytical Chemistry, 398*, 1403-1409.

Castrignanò, S., Sadeghi, S. J. and Gilardi, G. (2012) Entrapment of human flavin-containing monooxygenase 3 in the presence of gold nanoparticles: TEM, FTIR and electrocatalysis, *Biochimica et Biophysica Acta- General Subjects, 1820*, 2072-2078.

Catucci, G., Gilardi, G., Jeuken L. and Sadeghi, S. J. (2012) In vitro drug metabolism by C-terminally truncated human flavin-containing monooxygenase 3. *Biochemical Pharmacology, 83,* 551-558.

Catucci, G., Occhipinti, A., Maffei, M., Gilardi, G. and Sadeghi, S. J. (2013) Effect of human flavin-containing monooxygenase 3 polymorphism on the metabolism of aurora kinase inhibitors. *International Journal of Molecular Sciences, 14,* 2707-2716.

Charifson, P. S. and Kuntz, I.D. (1997). Recent successes and continuing limitations in computer aided drug design. In: PS Charifson (Ed.), *Practical Application of Computer Aided Drug Design*. New York: Dekker.

Chenna, R. (2003) Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Research, 31,* 3497-3500.

Cho, H. J., Cho, H. Y., Kim, K. J., Kim, M. H, Kim, S. W. and Kang, B. S. (2011) Structural and functional analysis of bacterial flavin-containing monooxygenase reveals its ping-pong-type reaction mechanism. *Journal of Structural Biology, 175*, 39-48.

Colovos, C., and Yeates, T. O. (1993) Verification of protein structures: patterns of nonbonded atomic interactions. *Protein Science, 2,* 1511-1519.

Coulter-Mackie, M.B. and Lian, Q. (2008) Partial trypsin digestion as an indicator of mis-folding of mutant alanine:glyoxylate aminotransferase and chaperone effects of specific ligands. Study of a spectrum of missense mutants. *Molecular Genetics and Metabolism, 94*, 368-374.

Dassault Systèmes BIOVIA, Discovery Studio Modeling Environment, Release 2017, San Diego:

Dassault Systèmes, 2016.

Dolphin, C. T., Janmohamed, A., Smith, R.L., Shephard, E. A. and Phillips, I. R. (1997) Missense mutation in flavin-containing mono-oxygenase 3 gene, FMO3, underlies fish-odour syndrome. *Nature Genetics, 17,* 491-494.

Dolphin, C. T., Janmohamed, A., Smith, R. L., Shephard, E. A., and Phillips, I. R. (2000) Compound heterozygosity for missense mutations in the flavin-containing monooxygenase 3 (FMO3) gene in patients with fish-odour syndrome. *Pharmacogenetics, 10*, 799-807.

Eswaramoorthy, S., Bonanno, J. B., Burley, S. K., and Swaminathan, S. (2006) Mechanism of action of a flavin-containing monooxygenase. *Proceedings of the National Academy of Sciences, 103,* 9832-9837.

Fontana, A., Polverino de Laureto, P., Spolaore, B., Frare, E., Picotti, P., and Zambonin, M. (2004) Probing protein structure by limited proteolysis. *Acta Biochimica Polonica, 51*, 299-321.

Fraaije, M. W., Kamerbeek, N. M., van Berkel, W. J. H., and Janssen, D. B. (2002) Identification of a Baeyer-Villiger monooxygenase sequence motif. *FEBS Letters, 518*, 43–47.

Gallivan, J. P., and Dougherty, D. A. (1999) Cation-pi interactions in structural biology. *Proceedings of the National Academy of Sciences. U.S.A., 96,* 9459-9464.

Goodsell, D. S., Morris, G. M. and Olson, A. J. (1996) Automated docking of flexible ligands: applications of autodock. *Journal of Molecular Recognition, 9*, 1–5.

Gouet, P., Courcelle, E., Stuart, D. I., and Metoz, F. (1999) ESPript: multiple sequence alignments in PostScript. *Bioinformatics, 15*, 305–308.

Krieger, E., and Vriend, G. (2002) Model@Home: distributed computing in bioinformatics using a screensaver based approach. *Bioinformatics, 18*, 315–318.

Hulme, E. C. and Trevethick, M. A. (2010) Ligand binding assays at equilibrium: validation and interpretation. *British Journal of Pharmacology, 161*, 1219-1237.

Leiser, S. F., Miller, H., Rossner, R., Fletcher, M., Leonard, A., Primitivo, M., Rintala, N., Ramos, F. J., Miller, D. L. and Kaeberlein, M. (2015). Cell nonautonomous activation of flavin-containing monooxygenase promotes longevity and health span. *Science, 350*, 1375-1378.

McGaughey, G.B., Gagnes, M., and Rappe, A.K. (1998) pi-Stacking interactions. Alive and well in proteins. *Journal of Biological Chemistry, 273*, 15458-15463.

Orru, R., Pazmino, D. E .T., Fraaije, M. W., and Mattevi, A. (2010) Joint Functions of Protein

Residues and NADP(H) in Oxygen Activation by Flavin-containing Monooxygenase. *Journal of Biological Chemistry, 285,* 35021-35028.

Ozawa, T., Tsuji, E., Ozawa, M., Handa, C., Mukaiyama, H., Nishimura, T., Kobayashi, S., and Okazaki, K. (2008) The importance of CH/π hydrogen bonds in rational drug design: An ab initio fragment molecular orbital study to leukocyte-specific protein tyrosine (LCK) kinase. *Bioorganic & Medicinal Chemistry, 16*, 10311-10318.

Pierce, A. C., Sandretto, K. L., and Bemis, G. W. (2002) Kinase inhibitors and the case for CH?O hydrogen bonds in protein-ligand binding. *Proteins: Structure, Function, and Genetics, 49*, 567-576.

Ringer, A. L., Senenko, A., and Sherrill, C. D. (2007) Models of S/π interactions in protein structures: Comparison of the H2S-benzene complex with PDB data. *Protein Science, 16*, 2216-2223.

Sadeghi, S. J., Meirinhos, R., Catucci, G., Dodhia, V. R., Di Nardo, G., and Gilardi, G. (2010) Direct Electrochemistry of Drug Metabolizing Human Flavin-Containing Monooxygenase: Electrochemical Turnover of Benzydamine and Tamoxifen. *Journal of American Chemical Society, 132,* 458-459.

Schugar, R.C., and Brown, J. M. (2015) Emerging roles of flavin monooxygenase 3 in cholesterol metabolism and atherosclerosis. *Current Opinion in Lipidology, 26*, 426-431.

Shimizu, M., Kobayashi, Y., Hayashi, S., Aoki, Y., and Yamazaki, H. (2012) Variants in the flavin-containing monooxygenase 3 (FMO3) gene responsible for trimethylaminuria in a Japanese population. *Molecular Genetics and Metabolism, 107*, 330-334.

Wiederstein, M., and Sippl, M.J. (2007) ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Research, 35*, W407-W410.

Yamazaki, H., Fujita, H., Gunji, T., Zhang, J., Kamataki, T., Cashman, J. R., and Shimizu, M. (2007) Stop codon mutations in the flavin-containing monooxygenase 3 (FMO3) gene responsible for trimethylaminuria in a Japanese population. *Molecular Genetics and Metabolism, 90*, 58-63.

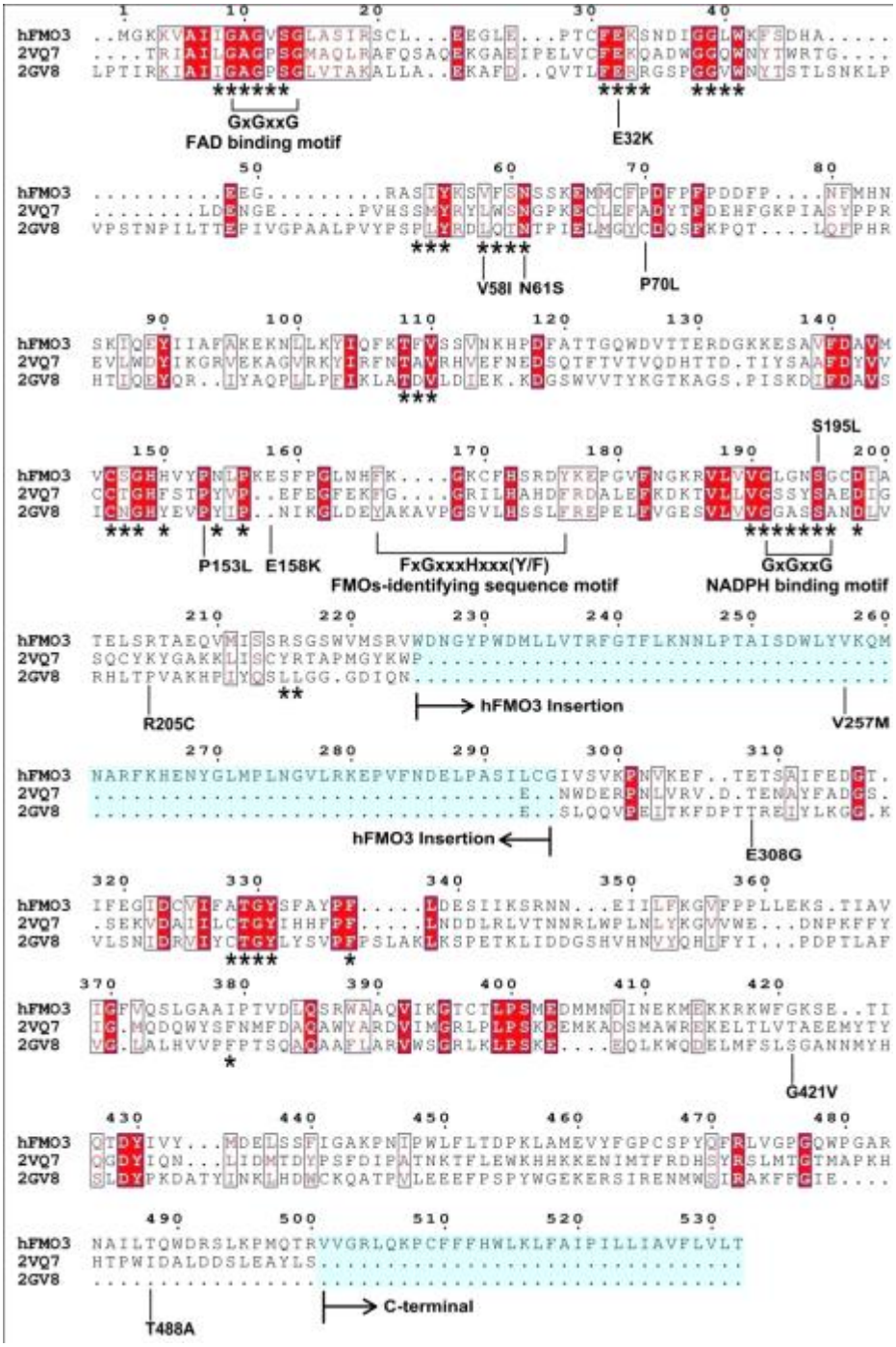
Yeung, C. K., Adman, E. T., and Rettie, A. E. (2007) Functional characterization of genetic variants of human FMO3 associated with trimethylaminuria. *Archives of Biochemistry and Biophysics, 464,* 251-259.

Zhang, J., Tran, Q., Lattard, V., and Cashman, J. R. (2003) Deleterious mutations in the flavin-containing monooxygenase 3 (FMO3) gene causing trimethylaminuria. *Pharmacogenetics, 13,*

495–500.

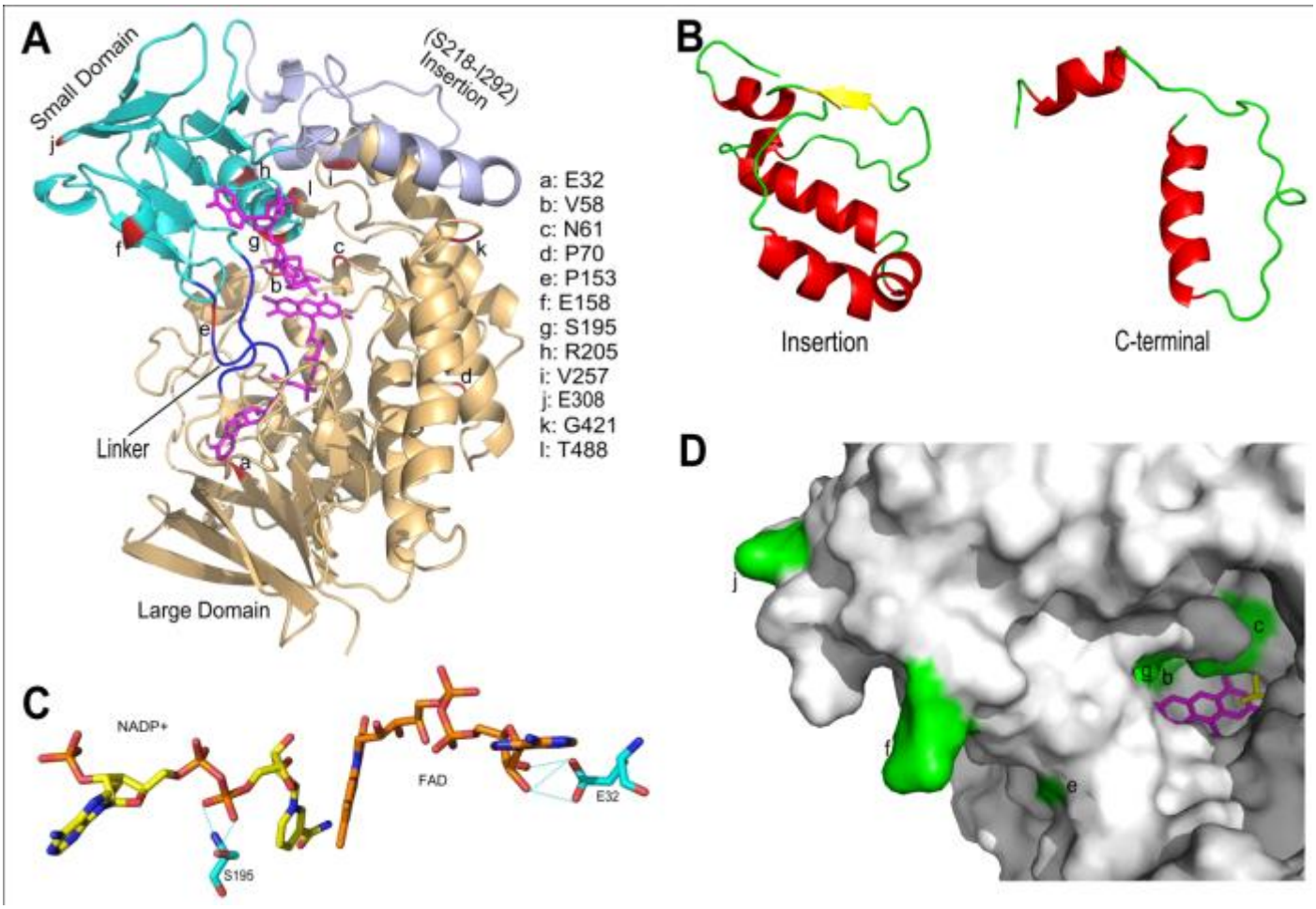
Zhou, J., and Shephard, E. A. (2006) Mutation, polymorphism and perspectives for the future of human flavin-containing monooxygenase 3. *Mutation Research/Reviews in Mutation Research, 612*, 165-171.

Figure 1



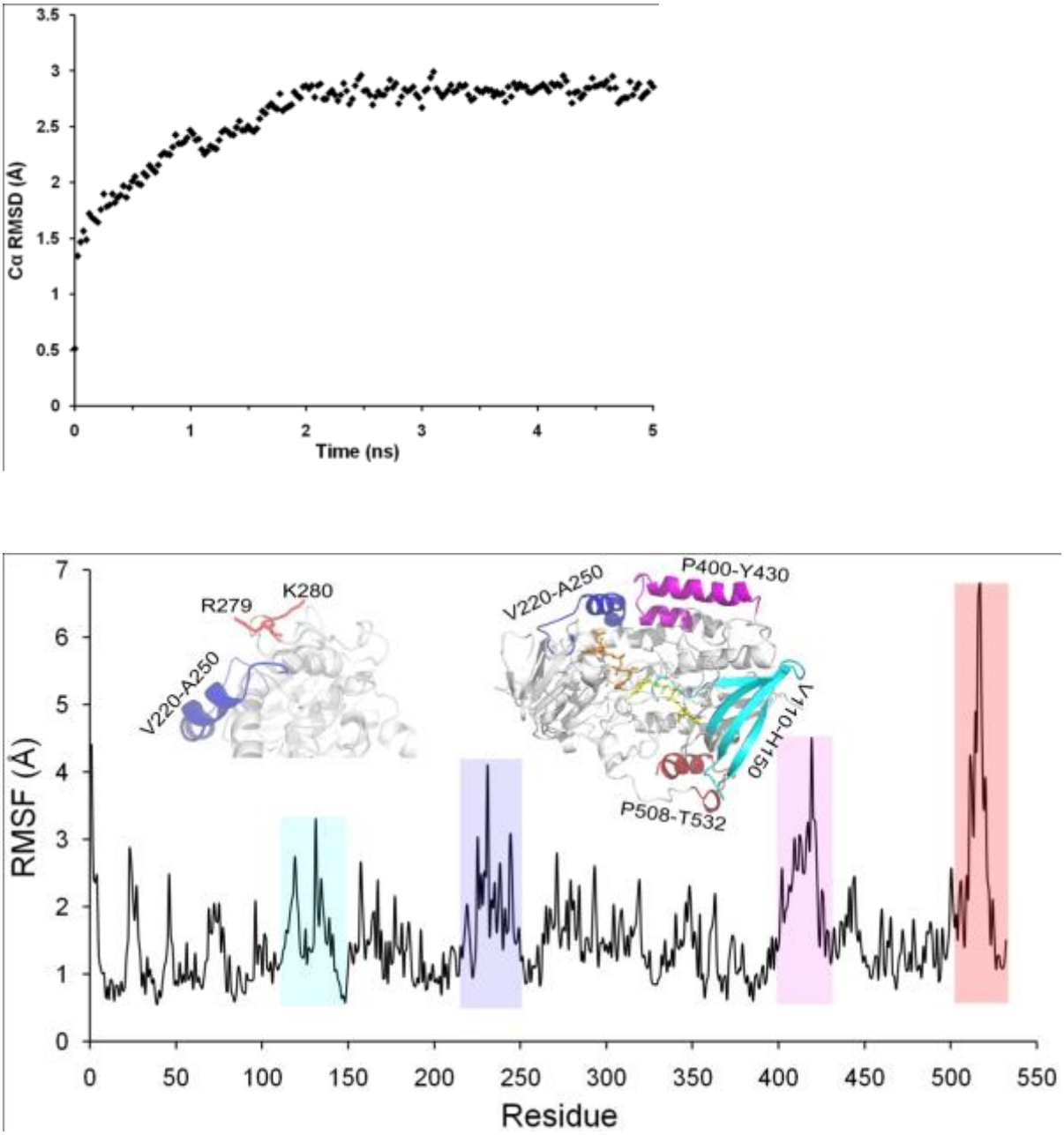
Alignment of the sequences of hFMO3, bFMO and yFMO. Conserved residues are colored in red, and boxed residues are semi-conserved. The residues around cofactors ($NADP^+$ and FAD) in bFMO and yFMO are starred underneath. Motifs and mutated residues in patients affected by TMAU are also indicated. The insertion and C-terminal regions are colored in cyan. The multiple sequence alignment was performed with ClustalW and prepared using ESPript 3.0.

Figure 2



(A) Structural model of hFMO3. NADP-binding domain is light orange with FAD-binding domain in cyan. The insert region is colored in light blue with the linker in dark blue. The mutation sites are labeled in red on the model. (B) The structural models of insert region and C-terminal portion of hFMO3 built by Phyre2. (C) Hydrogen bonds involving S195 and E32 residues with $NADP^+$ and FAD, respectively. (D) The surface representation of hFMO3 docked with trimethylamine (colored in yellow). The FAD is shown in magenta with mutation sites in green. Figures were generated with PyMOL.

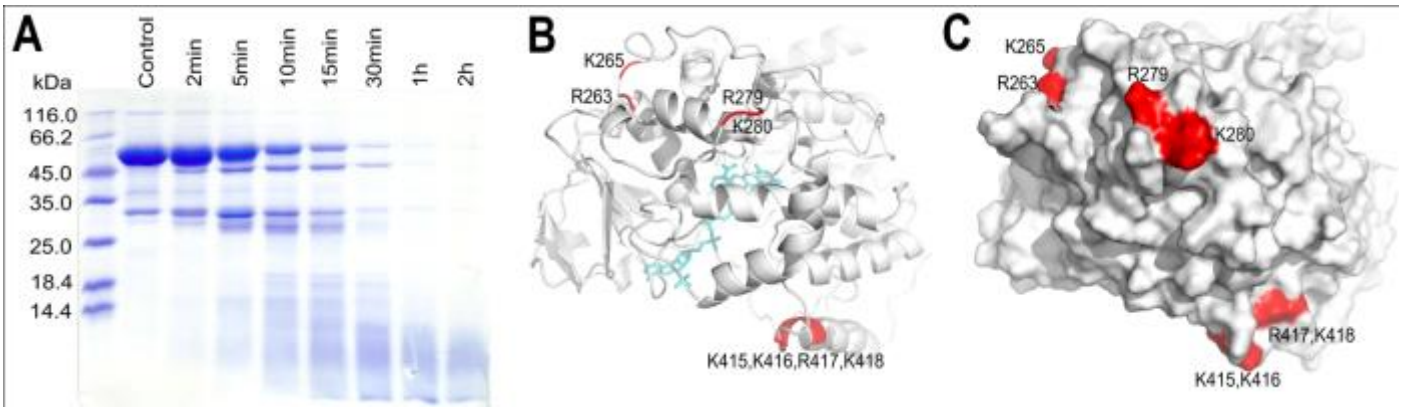
Figure 3



(A) Graphical representation of RMSD of the simulation trajectory of Cα atoms showing the structural stability of hFMO3 after 2 ns of MD simulation.

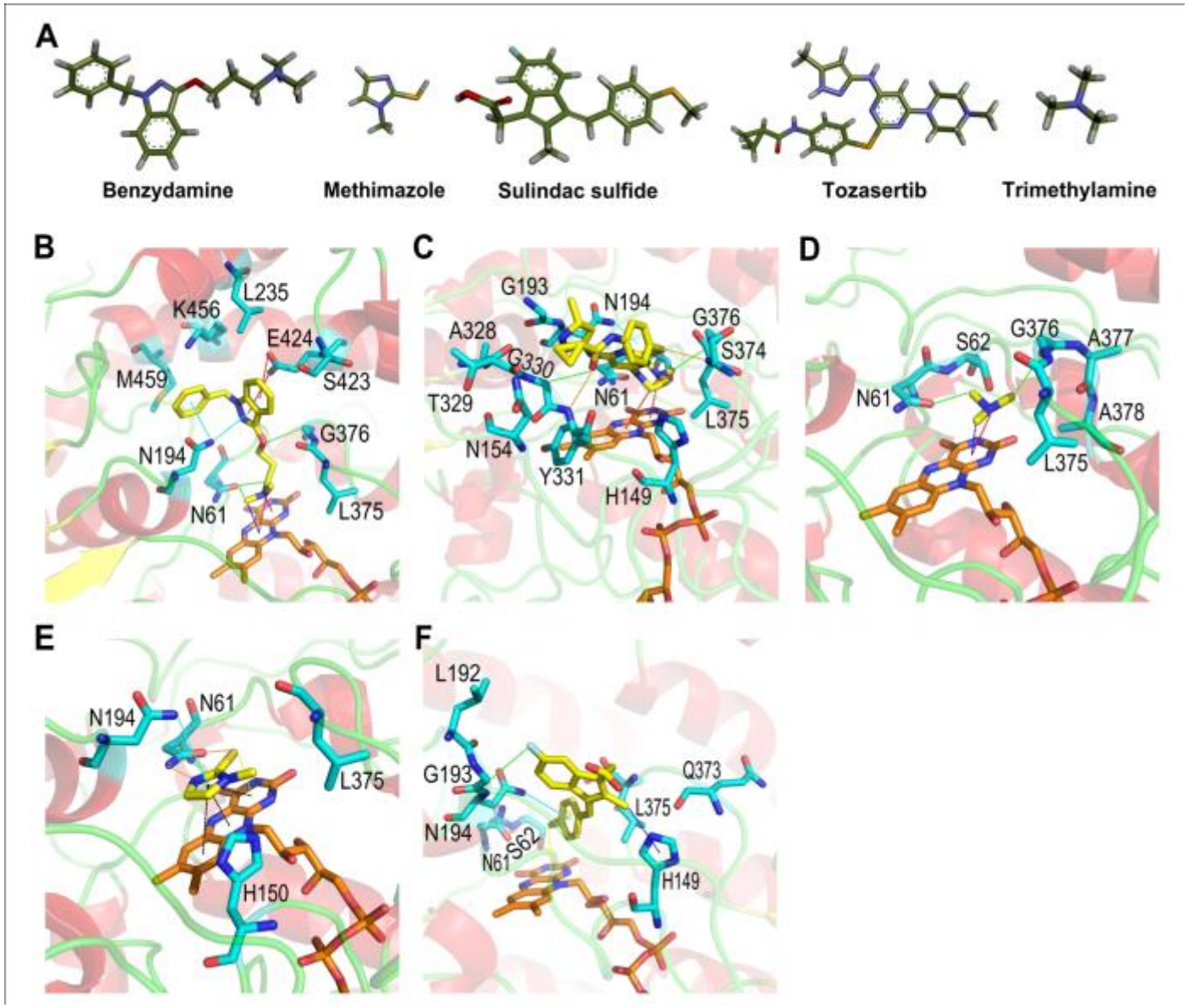
(B) RMSF per residue plot for hFMO3 with four highly flexible regions; V110-H150 (cyan), V220-A250 (blue), P400-Y430 (magenta) and P450-T532 (red). Figure inserts show these regions are located on the surface of the structure and exposed to the solvent.

Figure 4



(A) Time course of trypsin digestion of purified hFMO3. Samples were analysed by 12.5% SDS-PAGE. (B) Cartoon representation and (C) surface representation of the possible enzyme digestion sites based on trypsin digestion banding pattern are labeled on hFMO3 model. Figures were prepared by PyMOL.

Figure 5



(A) Substrates used for docking experiments were built and optimized using DS Visualizer. Binding modes of hFMO3-substrate complexes performed by YASARA Structure package.

Molecular docking experiments with hFMO3-Benzydamine (B); hFMO3-Tozasertib (C); hFMO3-Trimethylamine (D); hFMO3-Methimazole (E); hFMO3-Sulindac sulfide (F). Substrate binding site residues are shown as sticks and colored in cyan. Substrates are colored in yellow, and FAD is shown in orange. Pi-stacking, Pi-sigma, Pi-charge and Pi-sulfur interactions are shown as dotted black, blue, red and yellow lines, respectively. The conventional hydrogen bond, carbon hydrogen bond and

Pi-donor hydrogen bonds are shown in dotted orange, green and cyan lines, respectively. The interactions were monitored by DS Visualizer, and figures were generated with PyMOL.