

Predicting Reactivity to Drug Metabolism: Beyond P450s – Modelling FMOs and UGTs

*Mario Öeren[†], Peter J. Walton^{† *}, Peter A. Hunt[†], David J. Ponting[‡], Matthew D. Segall[†]*

† Optibrium Limited, Cambridge Innovation Park, Denny End Road, Cambridge, CB25 9PB, UK

*** School of Chemistry, University of Nottingham, University Park, Nottingham, NG7 2RD, UK**

‡ Lhasa Limited, Granary Wharf House, 2 Canal Wharf, Leeds, LS11 5PS, UK

E-mail of the corresponding author: mario@optibrium.com

ORCID of the corresponding author: 0000-0003-4292-5557

Abstract

We present a study based on density functional theory calculations to explore the rate limiting steps of product formation for oxidation by Flavin-containing Monooxygenase (FMO) and glucuronidation by the UDP-glucuronosyltransferase (UGT) family of enzymes. FMOs are responsible for the modification phase of metabolism of a wide diversity of drugs, working in conjunction with Cytochrome P450 (CYP) family of enzymes, and UGTs are the most important class of drug conjugation enzymes. Reactivity calculations are important for prediction of metabolism by CYPs and reactivity alone explains around 70 – 85 per cent of the experimentally observed sites of metabolism within CYP substrates. In the current work we extend this approach to propose model systems which can be used to calculate the activation energies, i.e. reactivity, for the rate-limiting steps for both FMO oxidation and glucuronidation of potential sites of metabolism. These results are validated by comparison with the experimentally observed reaction rates and sites of metabolism, indicating that the presented models are suitable to provide the basis of a reactivity component within generalizable models to predict either FMO or UGT metabolism.

Keywords

DFT, FMO, Glucuronidation, Oxidation, Reactivity, UGT

Declarations

Funding: The research was funded by Optibrium Ltd. and Lhasa Ltd.

Conflicts of interest: MÖ, PAH and MDS are employees of Optibrium Ltd. DJP is an employee of Lhasa Ltd. PJW was an employee of Optibrium Ltd. for the duration of the FMO project.

Ethics approval: The research does not involve human participation or personal data.

Availability of data and material: The data used in the project is available through the referenced publications.

Availability of data and material: The supporting information includes Cartesian coordinates, Z-matrices and total DFT energy values of the presented structures. In addition, the general DFT input files are provided.

Code availability: The software used in the project is freely available through the referenced publications.

Introduction

One of the main functions of metabolism is the elimination of endo- and xenobiotics. Since organisms are continually exposed to a myriad of substrates which vary in physical and chemical properties, numerous pathways and enzymes are required for their elimination. Conventionally, metabolic elimination is divided into three phases – modification, conjugation and excretion. [1, 2] During the modification phase, a variety of enzymes catalyse the oxidation, reduction or hydrolysis of endo- and xenobiotics. The objective of these transformations is to increase the solubility of a compound or provide a site for conjugation reactions. [2] In the conjugation phase, enzymes known as transferases link a polar moiety such as glucuronic or sulfonic acid from a cofactor to a site of metabolism (SoM), further increasing the solubility of the substrate. [2, 3] The two phases, both of which are not always required, prepare the endo- and xenobiotic compounds for excretion, commonly, but not exclusively in urine. [1]

Studying modification and conjugation reactions within metabolic pathways helps to understand the metabolic fates and pharmacokinetics of molecules which, in turn, is valuable for the development of drugs, agrochemicals, nutritional supplements, and cosmetics. [4, 5] The most important enzymes are from the cytochrome P450 (CYP) family, which are responsible for the modification of around 75 per cent of hepatically cleared drugs. [6, 7] The properties of CYPs have been thoroughly studied, and predictive models of metabolic regioselectivity and isoform specificity have reached accuracies of 90 per cent. [5, 8, 9, 10, 11, 12] Enzymes which work in conjunction with CYPs in the modification phase or enzymes which catalyse the conjugation reactions have been far less extensively studied.

One of the most important auxiliary enzyme classes in the modification phase is called Flavin-containing Monooxygenase (FMO). FMOs are found in multiple tissues in the human body and they metabolise a wide range of endo- and xenobiotics. Eleven human isoforms of the FMO family have been identified. Five of these are functional active isoforms and are labelled FMO1–5 and the rest are non-functional pseudogenes (FMO6P–11P). [13] There is currently no crystal structure for human FMO isoforms as they are found in the membranes of the endoplasmic reticulum (ER), which makes them difficult to crystallise. [14] However, examples of bacterial and fungal crystal structures have been observed. [15]

An example of a significant family of conjugation enzymes is the uridine diphosphate-glucuronosyltransferases (UGT). As with FMOs, UGTs are found across the human body, but the total number of known isoforms is much greater with 22 active isoforms and 9 pseudogenes. Based on their sequence similarity, human UGTs are divided into four categories – UGT1, UGT2, UGT3 and UGT8. [16, 17, 18] The first two are mainly responsible for the transformation of both xenobiotics and endogenous compounds, UGT3 enzymes only have a minor role in the metabolism of xenobiotics, and the UGT8 family appears to have no role in metabolism (it catalyses the transfer of galactose from UDP-galactose to ceramide, important for the myelin sheath of nerve cells). [17, 19, 20] The names of the active isoforms are the following: 1A1, 1A3-10, 2A1-3, 2B4, 2B7, 2B10, 2B11, 2B15, 2B17, 2B28, 3A1, 3A2 and 8A1. [16, 20] Human UGTs are also found in the ER as well and the only reported crystal structure to date is the C-terminal domain of 2B7. [21, 22]

Recently, attention has been placed on understanding the metabolic behaviour of FMOs. While CYPs are the major contributors to oxidative metabolism (approximately 96 per cent of cases), the role of FMOs should not be underrated. The substrates for both enzyme classes overlap and, due to this, the contribution of FMOs to drug metabolism is often underestimated or ignored. [6, 13, 23] Furthermore, FMOs might also yield toxic metabolites such as sulfines and sulfenes. [24] Thus, modelling FMO activity would enable chemists to design drugs that are geared towards FMO-mediated metabolism, should CYP interaction be problematic (e.g. due to drug-drug interactions) or undesirable, while avoiding toxic metabolites. The number of studies regarding FMOs is slowly growing [7], and at least one QSAR model, using descriptors derived from quantum mechanics (e.g. Fukui reactivity indices) and circular fingerprints, has been published [25].

UGTs on the other hand are known to be responsible for the conjugation of around 15 per cent of current hepatically cleared drugs and around 40 per cent of conjugation reactions. [6, 7, 26, 27] Thus, UGTs are considered the second most important enzyme class when considering the metabolism of drug-like compounds. The number of studies regarding UGTs has steadily grown over the years [6, 7] and many quantitative structure-activity relationship (QSAR) models have been published using publicly available data, e.g. by Sorich et al. [28], Dang et al. [29] and Mazzolari et al. [30]. However, none of the referenced models explicitly utilize reactivity and accessibility of potential substrates in a similar fashion to some of the successful CYP models. [5, 8]

Aim of the Study

Successful CYP metabolism models which combine empirical data with mechanistic approaches, [5, 8] consider two factors that influence the SoM – reactivity and accessibility. Accessibility may be taken into account through docking within a model of the active site or two-dimensional SoM-specific (rooted) steric and orientation descriptors, whereas the reactivity of potential SOMs is calculated using quantum mechanical (QM) calculations. The QM calculations are used to estimate the activation energy (E_a) of product formation for each potential SoM, which describes their pure reactivity. In the CYP models, the E_a alone explains around 70 – 85 per cent of the experimentally observed SoMs within substrates. [5, 8]

In order to conduct the calculations within a reasonable time frame, the E_a is calculated using a simplified model of the enzyme system. While several simplifications are introduced, it is important to ensure that the simplified model correlates with the experimental observations – the experimental reaction rates and E_a values should be inversely correlated through the Arrhenius equation (Eq. 1) and the majority of observed sites would be expected to have the lowest calculated E_a within the substrate molecule. [5, 8]

$$\ln(k) = \ln(A) - \frac{E_a}{kT} \quad (\text{Eq. 1})$$

A detailed understanding of the energetics of the catalytic mechanism, especially the rate limiting step of product formation, of both FMOs and UGTs is the first stepping stone of generating accurate and transferable models of metabolism by these enzymes. In the following sections we will explore the roles of both FMOs and UGTs in the drug metabolism, review the details known about the reaction mechanisms and study the rate limiting steps of product formation using density functional theory (DFT).

Thus, the aim of the current study is to (a) determine simplified mechanisms of the product formation steps for both FMOs and UGTs and (b) validate the mechanisms by comparing the calculated E_a values to experimentally measured reaction rates and observed SoMs. The study will include *N*- and *S*-oxidation for FMOs and *N*- and *O*-glucuronidation for UGTs, since those four reactions are prevalent amongst the chosen enzyme systems. The results would give a strong foundation for generating a combined reactivity-accessibility model for both FMO and UGT metabolism in the future.

The Role and Mechanism of FMOs

Oxidation by FMOs works by transferring an oxygen atom from a hydroxyperoxyflavin, generated from the flavin adenine dinucleotide (FAD) cofactor, to the “soft-nucleophile” of the respective substrate as depicted in Figure 1. The commonly oxidised nucleophiles include nitrogen and sulfur atoms, but examples from various species include selenium, iodine, boron, phosphorus and even carbon oxidation. [13, 24] The reaction types observed within humans include *N*- and *S*-oxidation, *N*-demethylation and desulfuration, but *N*- and *S*-oxidation are by far the most prevalent. [13]

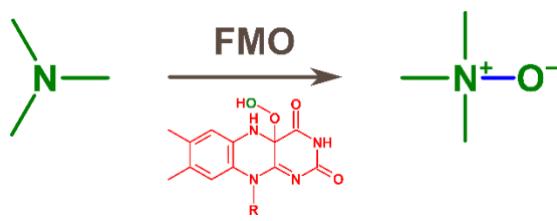


Fig. 1 Oxidation of trimethylamine (TMA), a known FMO substrate (the blue bond represents the newly formed bond). [13] The substituent “R” represents the adenine dinucleotide part of FAD.

The substrate specificity of FMOs overlaps with CYPs and although substrates tend to be metabolised by CYPs, there are molecules which are exclusively or predominately metabolised by FMOs – benzylamine [31], itopride [32], albendazole [33], cimetidine [34, 35, 36] to name a few. It should be noted that the metabolites of FMO are often less toxic than CYP-mediated metabolites. [13] The potential sites of metabolism include tertiary-, secondary- and primary alkyl and aryl amines, hydrazines, sulfides, thiols and disulfides, thiocarbamides and thioamides, and mercapto-purines, pyrimidines, and imidazoles. [24]

The initial studies with pig FMOs (FMO1) showed that FMOs are very promiscuous towards substrates and metabolise any compound containing a “soft-nucleophile” which can access the active site. Compounds with a single positive charge are excellent substrates while compounds with more than one positive charge are not. Zwitterions and compounds with a negative charge are generally not metabolised with few exceptions. [24, 37] More recent studies with human FMOs, however, show distinct differences in specificity between FMO isoforms based on shape of the active site, e.g. FMO1 can oxidise substrates with soft nucleophiles extending up to 5 Å from a bulky ring system (e.g. imipramine) while FMO2 cannot because of its buried active site. From the drug metabolism perspective, the most interesting isoforms are FMO1 and FMO3, since the active site of FMO2 appears to be the most restricted and not much is known about FMO4 and FMO5. [24]

The catalytic cycle of oxidation of substrates by FMOs, depicted on Figure 2, includes reduced nicotinamide adenine dinucleotide phosphate (NADPH), FAD and molecular oxygen (O_2). It starts by reducing the FAD cofactor to $FADH_2$ using NADPH. $FADH_2$ reacts with O_2 , forming a stable hydroperoxyflavin ($FAD-OOH$), which is then ready to react with any suitable nucleophile which can gain access to the active site. During the oxidation reaction, the distal oxygen atom of the hydroperoxide is transferred to the substrate, effectively oxidising the substrate and forming $FAD-OH$. The original FAD reaction centre is then regenerated by the elimination of water and $NADP^+$. [15, 24] Based on the bacterial FMO structures, the adenine dinucleotide tail of FAD is believed to be tightly bound to a groove in the large domain of the protein and only the tricyclic motif is exposed to the potential substrate. The NADPH cofactor is bound to another groove of the large domain and is held close to FAD and an asparagine residue (Asn-91), which aids the delivery of O_2 , to allow rapid progression through the catalytic cycle. [13]

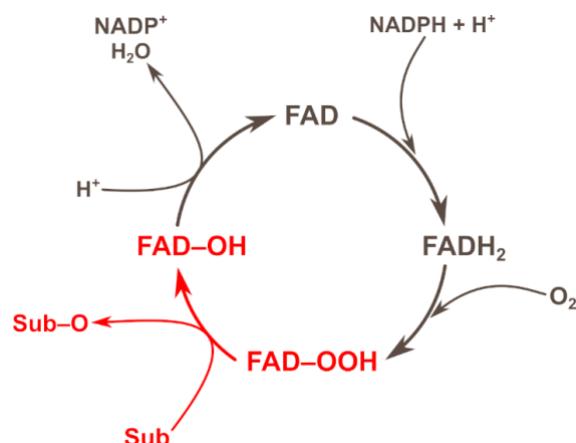


Fig. 2 The simplified catalytic cycle of FMOs. “Sub” stands for substrate and this work will focus on substrate oxidation, highlighted in red.

Previously published studies of FMO mediated *N*- and *S*-oxidation have proposed two potential reaction types for the rate limiting step: an S_N2 reaction with a heterolytic split of the peroxide bond [38] and a radical reaction with a homolytic split of the peroxide bond [39].

The Role and Mechanism of UGTs

UGTs are a sub-class of glycosyltransferases (GT), which are responsible for catalysing the formation of the glycosidic bonds to form glycosides as depicted in Figure 3. The prevalent sites of metabolism for UGTs are phenol groups, carboxylic acids and alcohols (*O*-glucuronidation) followed by amines, amides and *N*-heterocycles (*N*-glucuronidation [40]). *S*- and *C*-glucuronidation are known but rarely observed in humans. [41, 42] The substrate spectrum for UGTs is broad and structurally unrelated, for example, in humans, ethanol, [43] salicylic acid, [44] nicotine, [45] benzo(a)pyrenol [46] and bilirubin [47] are all glucuronidated.

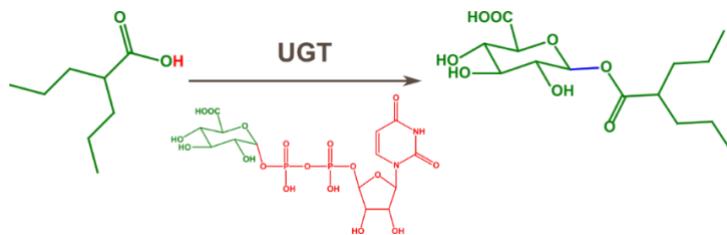


Fig. 3 Glucuronidation of valproic acid (the blue bond represents a glycosidic bond). [48]

The broad substrate specificity for the class is due to the number of different isoforms, and the overlap between which compounds they are able to metabolise. [49, 50, 51] Nonetheless, isoforms have their preferred substrate types; for example, 1A6 mainly glucuronidates small phenolic substances and 2B7 favours opioids. In addition, some isoforms have been tailored, but not limited, to specific tasks; for example, bilirubin is solely glucuronidated by 1A1 [52] and 1A4 can glucuronidate tertiary amines (a reaction that has been seen mostly in humans and higher primates). [40, 53] Due to their role in xenobiotic metabolism, most glucuronidation studies relate to substrate activity feature UGT1 and UGT2 enzymes.

In general, the glucuronosyl reactions follow a mechanism where the sugar donor and the substrate are bound sequentially, followed by the sugar transfer, inverting the configuration at the anomeric centre. The product is then released, followed by the release of the nucleotide moiety. Inverting GTs utilize a direct displacement S_N2 -like mechanism with an oxocarbenium ion-like transition state. [54, 55] While many GT's use cations (generally Mn^{2+} or Mg^{2+}) to stabilize the leaving group, there is no evidence of a bound metal ion in UGT's. [56, 57, 58] According to the docking studies, which used a homology model of 1A1, two histidine residues (supported by an aspartate residue) are the base (H39) and acid (H372) for the reaction, forming a “catalytic dyad” depicted in Figure 4. [56] The phosphates of the leaving group are stabilized by serine and glycine residues (S38 and G377 respectively) removing the need for a metal cation. [56, 57, 58] Other authors have suggested that either the D151 or H372 may instead act as a general acid for the UGT1 family, [57, 58] but irrespective of the details concerning residues, the general acid-base mechanism, where proton transfers accompany an S_N2 mechanism for UGTs has been generally accepted. [59]

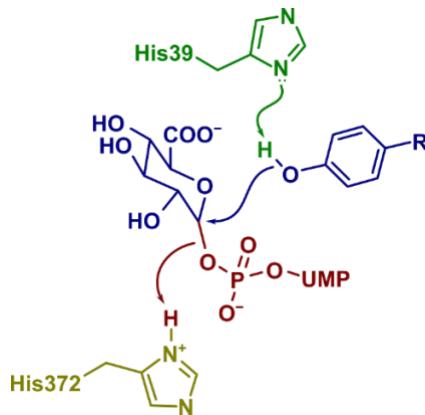


Fig. 4 The proposed catalytic dyad of glucuronidation (UGT1A1), where UMP stands for uridine monophosphate and R is a placeholder for an undefined substituent. [56]

No computational studies of the mechanisms of action of UGTs have been reported prior to this work.

Methods

For DFT calculations, which are computationally expensive in comparison with methods routinely applied to drug discovery and metabolism prediction, the system under study should be small, but retain its chemical characteristics. A series of simplifications for both FMOs and UGTs was tested to reduce the size of the system without significantly changing the reactivity of the reaction centre towards different SoMs. To take into account the long-range effects of the substrate structure and not only the local environment of the SoM, the E_a will be evaluated in the context of the whole substrate molecule without any fragmentation. The models will be validated using experimental data after an acceptable simplification of the system has been achieved, and a transition state geometry has been found.

Computational Methods

All input molecular structures in this work were generated using the program Avogadro [60] or CORINA Classic [61] and pre-optimized using the semi-empirical method AM1 [62]. The same method was also used for scanning the potential energy surfaces of transition states of potential FMO and UGT substrates. The semi-empirical calculations were performed using the program package MOPAC 7. [63]

Further geometry optimizations were conducted with DFT using the B3LYP functional [64, 65, 66, 67] and the def2-SVP [68] basis set. B3LYP was chosen because the reaction mechanisms of oxidation and glucuronidation feature only organic molecules and, in previous cases, geometry optimizations, including transition states calculated using hybrid GGA functionals yielded similar results to the more expensive hybrid meta-GGA functionals. [69] The basis set for geometry optimization included the polarization functions for all atoms due to proton transfer and the potential for hydrogen bonding. The geometry optimisation results were checked by the frequency calculations to ensure that the chosen geometries were at either a local minimum or a first order saddle point, as appropriate. The geometry optimization and frequency calculations were followed by a single-point calculation using a B3LYP functional with a dispersion correction [70], along with the def2-TZVPD and def2-QZVPDD [68] basis set. All calculations were performed in vacuum in order to estimate the pure reactivity of each potential SoM and avoid any confounding factors. The density functional calculations were performed with the NWChem 6.8 program package. [71] The electronic wavefunctions were analysed using the Multiwfn program [72] and charges were calculated using Hirshfeld charges [73].

The Requirements of Experimental Data

While the E_a , calculated using DFT, is considered to be a good descriptor for lability, the accessibility cannot be ignored. In order to make a fair comparison between the measured reaction rates and the calculated E_a values, the variabilities in the protein environment and steric accessibility of the SoMs have to be minimized. Thus, the set of compounds with experimentally measured reaction rates should be structurally similar and measured for the same isoform. In addition, the compounds should be measured in the same laboratory using the same assay to exclude additional potential variabilities. Preferably the reaction rates are measured when the enzyme is fully saturated by the substrate (measuring the V_{max}). Unfortunately, for FMOs, we were unable to find a dataset which would satisfy all those requirements, thus instead of linear correlation, the ordering of potential SoMs will be measured. Nonetheless, the molecules in the dataset should be small, to minimise the effects rising from variations in binding orientation and steric accessibility. Thus, five smallest molecules from the data set were chosen which are known to be oxidised by FMO3 and have two or more potential SoMs.

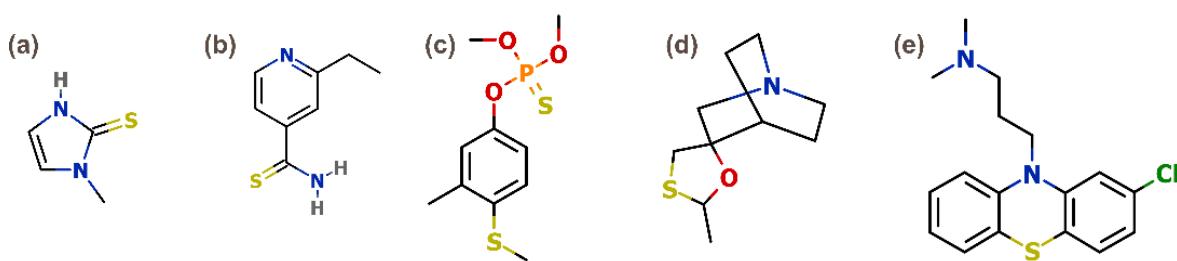


Fig. 5 The following compounds were used for validation: (a) methimazole; (b) ethionamide; (c) fenthion; (d) cevimeline; (e) chlorpromazine.

The molecules used in the FMO validation are: methimazole, [13] ethionamide, [13] fenthion, [74] cevimeline [13] and chlorpromazine [75] – Figure 5. Each molecule has at least two potential SoMs including both nitrogen and sulfur atoms (with the exception of fenthion, which has only sulfur atoms as potential SoMs). The test set has altogether five observed SoMs, thirteen potential SoMs with seven nitrogen atoms and six sulfur atoms.

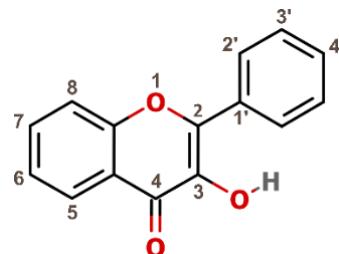


Fig. 6 The backbone of flavones used in the study.

The test set used for UGTs consists of six molecules and eleven observed sites of metabolism with measured V_{max} values for the UGT isoform 1A1. [76, 77] Each molecule in the set consists of a flavone backbone, depicted in Figure 6, with a varying number and location of hydroxy groups. Each potential SoM in the test set is a hydroxy group and undergoes O -glucuronidation. While the substrates and the SoM environments are very similar, it can be seen from Table 1 that the V_{max} values vary significantly, indicating that taking into account the long-range interactions during calculations is crucial.

Table 1 The experimental reaction rates (V_{max}) of the flavones used for validation of the model herein. Potential SoMs without a V_{max} (“–”) were not observed experimentally and were left out of the study. Data are means \pm standard error of three determinations. The observed sites of metabolism are indicated by an arrow.

Structure	Compound Name	Site	V_{max} (pmol min ⁻¹ mg)	Standard Error (pmol min ⁻¹ mg)
	3,7-dihydroxyflavone	3	1020	±540
		7	3040	±420
	Galangin	3	820	±140
		5	-	-
		7	4590	±460
	3,6,4'-Trihydroxyflavone	3	340	±50
		6	120	±10
		4'	340	±30
	Resokaempferol	3	680	±50
		7	1080	±130
		4'	-	-
	3,3',4'-Trihydroxyflavone	3	-	-
		3'	4810	±250
		4'	-	-
	Kaempferol	3	-	-
		5	-	-
		7	1380	±140
		4'	-	-

N-glucuronidation is understood to follow the same mechanism as *O*-glucuronidation since the C-terminal domain of UGTs, which is associated with the binding of the UDP-GA, is highly conserved and the H372 amino acid residue, seen in Figure 4, acts as a general acid in the “catalytic dyad” for all UGT1 isoforms. [20, 56, 58] Unfortunately, there are far fewer quantitative experimental data points for *N*-glucuronidation, and a set of structurally similar compounds cannot be compiled. Therefore, only one compound, trifluoperazine, will be used, with which to test the calculated reaction mechanism. The SoM which is observed to be glucuronidated by UGT1A4 is the tertiary amine seen in Figure 4. [78]

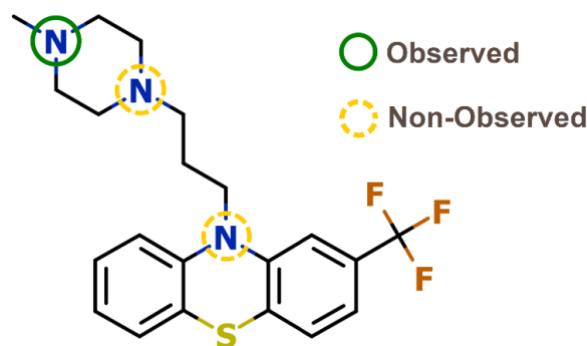


Fig. 7 *N*-glucuronidation of trifluoperazine with the potential SoMs shown in green and yellow (dashed) circles.

Results and Discussion

The experimentally observed SoM on a substrate is determined by the reactivity and accessibility of each potential SoM. In this study, the focus is on the reaction mechanism, and hence on the reactivity, as determined by the E_a . The interactions of the substrate with the protein binding site, which influences the orientation of the substrate and hence, the accessibility of each site will not be considered. Therefore, the initial simplification is to eliminate the protein and focus on the interaction between the cofactor and the substrate. It should be noted that the protein environment can be reintroduced via steric- and orientation descriptors as discussed earlier, and the approach has been proved to be fruitful for estimating the reactivity of the SoMs for CYPs. [5, 8] While the most significant contribution to calculation costs has been removed, additional simplifications for both FMOs and UGTs will be explored in the following sections.

FMOs

In addition to disregarding the effects raising from the protein environment, simplification of the FAD–OOH would also be beneficial to the calculation costs. The FAD structure can be divided into two parts: the adenine nucleotide and the flavin mononucleotide, which are connected through phosphate groups. According to experimental studies on bacterial crystal structures [15], the substrate interacts only with the flavin mononucleotide and the rest of the molecule, including the phosphate groups, is buried deep within a groove of the protein. This knowledge of the bacterial structure is assumed to be conserved in the human protein and therefore may permit a simplification by replacement of the adenine nucleotide and the bridging chain with a hydrogen atom (Figure 8, blue line). This is possible because the long-range electronic effects of the removed fragment are expected to be negligible during the dissociation of the peroxide bond. This was confirmed by calculation of the potential energy profile for peroxide bond dissociation, where the bond length was gradually increased from 1.45 to 2.0 Å, for both the full FAD–OOH and simplified system as seen in Figure 8. The system was then simplified further by investigating the contribution of the di-methyl benzene ring to the energy of the peroxide bond dissociation using DFT. It was found that the simplified structure, the red line in Figure 8, maintained indistinguishable peroxide bond dissociation behaviour when compared to the larger FAD system and to the full FAD–OOH as seen in the graph in Figure 8.

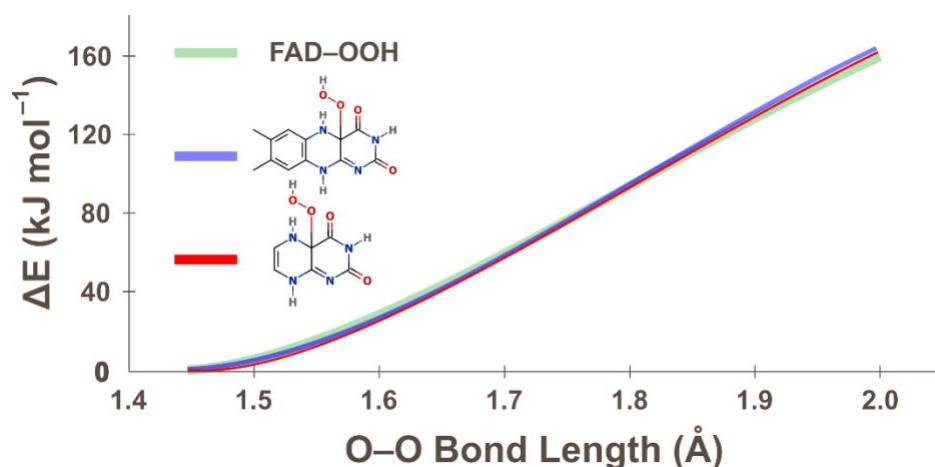


Fig. 8 The electronic behaviour of (a) the full FAD–OOH in green, (b) the FAD without the adenine nucleotide tail in blue and (c) the simplified FAD in red during the dissociation of the peroxide bond.

By building upon the work of previous studies [31, 39], we were able to locate the transition state for the oxidation of trimethylamine (TMA), one of the simplest known substrates for FMOs. However, despite multiple attempts, systematic potential energy scans and alterations to our approach (open-shell, restricted open-shell and various spin-multiplicities), a transition state for a radical mechanism could not be found. However, the same starting geometry, with closed-shell calculations, provided a transition state, representative of an S_N2 reaction (Figure 9). The current result provides evidence against the possibility of N - and S -oxidation occurring via a radical mechanism.

Of particular note in the transition state structure is the hydrogen of the peroxide. At the transition state the hydrogen is rotated towards and anchored by the carbonyl oxygen of the FAD ring system. This, in turn, lowers the E_a , assisting with the oxygen transfer. The reaction coordinate studies showed that after the transition state, the hydrogen is transferred directly to the proximal oxygen atom (although, in solution, it could be water-mediated).

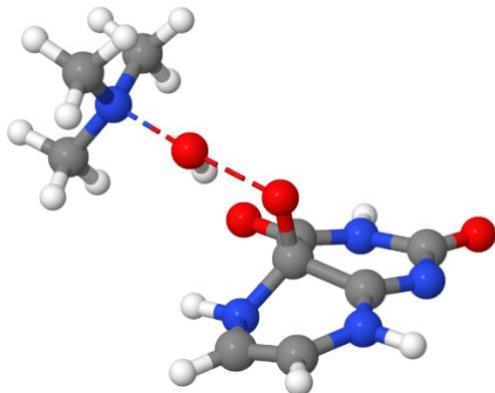


Fig. 9 Transition state structure for the closed-shell system.

By comparing the energy of the transition state with the sum of the energies of the isolated optimised reactants, an E_a of 61 kJ mol^{-1} was found using the SVP basis set. This energy was then probed further with the larger QZVPPD basis set to obtain an activation energy of 61 kJ mol^{-1} . Considering these calculations are performed at 0° K , the addition of thermal energy to the system suggests that the barrier height of this reaction is small enough to be overcome at internal body temperature. When comparing the energies of the isolated optimised reactants with isolated optimised products, products were found to be 32 kJ mol^{-1} lower in energy in the SVP basis set and 56 kJ mol^{-1} with the QZVPPD basis set. The difference in energy between basis sets can be explained by the better description of the zwitterionic trimethylamine oxide product. Despite this discrepancy there is still a large energy difference between the reactants and products which is the major driving force for the reaction. This investigation is summarised in Figure 10. Furthermore, as we are most interested in the E_a , this suggests that we do not need to use the more expensive basis set to get reasonable energies.

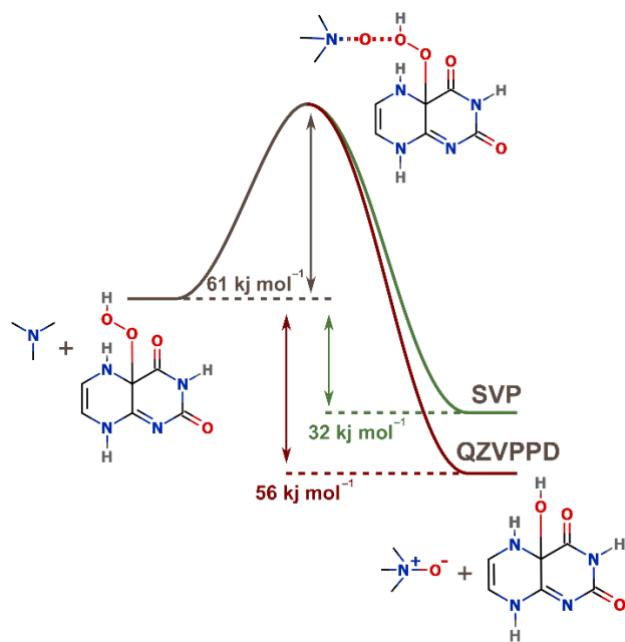


Fig. 10 The reaction coordinates of the oxidation of TMA.

Figure 11a shows that in the optimised reactant form, the charges present on the two oxygen atoms of the FAD–OOH are identical. However, at the transition state (Figure 11b), the charge on the proximal oxygen bonded to the ring system has more than doubled, with a charge equal to that seen on the carbonyl oxygen of the FAD ring. The charge on the distal oxygen that is being transferred to the nitrogen has kept a consistent charge to that which was seen in the reactant. The nitrogen of TMA, which would typically have a negative charge similar to what is seen on the labelled nitrogen of the FAD ring, has now become positive. These charges seen at the transition state indicate that a charge transfer is taking place, where an electron pair from the distal oxygen moves onto the proximal oxygen and an electron pair from the nitrogen atom of TMA moves onto the distal oxygen. This shows that the charge transfer occurring during oxidation is a “two-electron process”, consistent with an S_N2 mechanism.

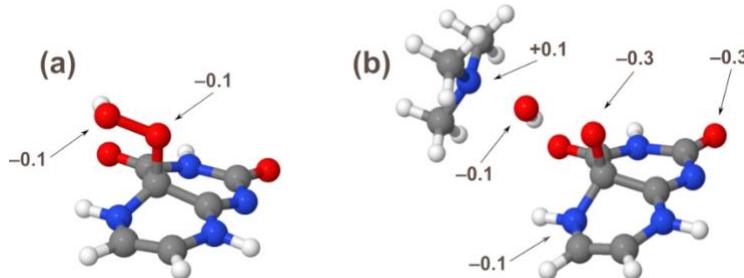


Fig. 11 Hirshfeld Charges present on specified atom in (a) isolated optimised FAD–OOH and (b) transition state.

The results of these investigations led to the following simplified curly-arrow mechanism to explain the initiation of the reaction (Figure 12): the approach of the lone pair towards the O–O σ^* antibonding orbital of FAD–OOH causes the peroxide bond to lengthen. The donation of both electrons from the lone pair of TMA causes the two electrons that were once shared between the oxygens to be pushed onto the proximal oxygen.

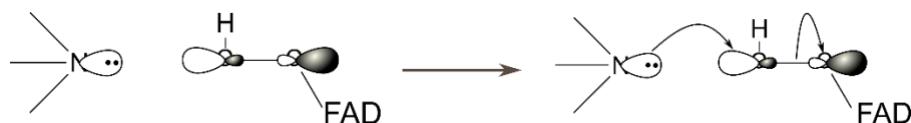


Fig. 12 Curly-arrow mechanism for the initiation of FMO-mediated oxidation.

After the thorough investigation into the S_N2 mechanism’s feasibility with TMA, the same modelling approach was applied to other known FMO substrates. Building the FAD–OOH reaction centre around possible SoMs for each substrate and then performing a transition state search enabled the determination of the E_a for each SoM.

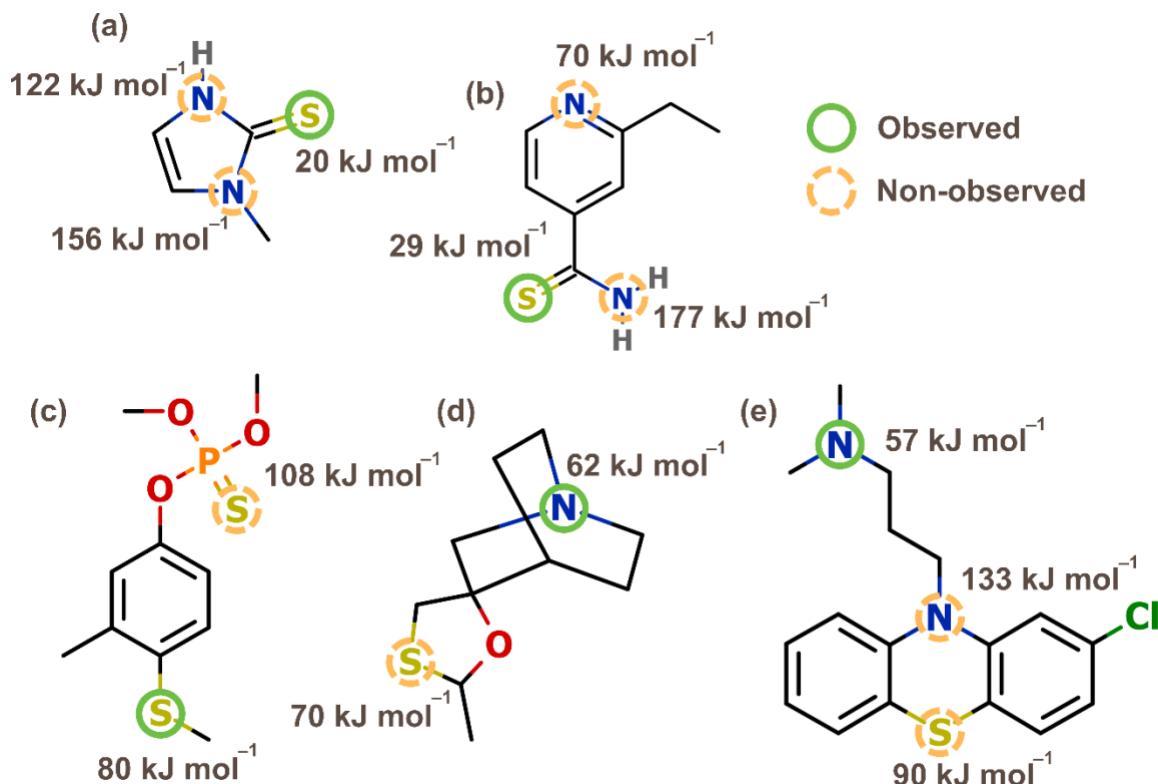


Fig. 13 E_a values of possible sites of metabolism for 5 known FMO substrates: (a) methimazole; (b) ethionamide; (c) fenthion; (d) cevimeline; (e) chlorpromazine. Green circles represent sites that are found to be oxidised in experiment and yellow (dashed) circles represent sites that are not.

The results of these calculations, shown in Figure 13, indicate that for each substrate, the site with the lowest activation energy is the site found to be oxidised experimentally. This is an indicator that activation energy plays an important role in determining which site within a molecule is the preferred SoM for FMOs. Another notable result of these calculations is that the S_N2 mechanism has been found to be applicable to larger systems with both nitrogen and sulfur environments. These results give an increased confidence that the selected simplified FAD–OOH and reaction mechanism are correct as it reproduces the site preferences of the real enzyme.

It should be noted that the SoMs of cevimeline cannot be explained only by reactivity due to the small difference in their E_a values and the preferred site is probably determined by additional steric and accessibility effects. Furthermore, here we are comparing the relative E_a values of potential sites within a molecule and the absolute E_a does not, itself determine if a site is metabolised. For example, the nitrogen atom in ethionamide has an E_a of 70 kJ mol⁻¹ but is not oxidised while the sulfur atom in fenthion, with an E_a of 80 kJ mol⁻¹, is metabolised. In this and other similar cases, the potential SoMs are prioritised by the E_a values within a molecule. Thus, a potential SoM with a low E_a value may not be metabolised due to competing SoMs with even lower E_a values in the same compound or due to steric and orientation effects.

UGTs

According to a homology model of UGT1A1, the UDP-GA is buried deep inside the protein and only a small part of the sugar ring, including the anomeric carbon, is accessible to a potential substrate. [79] Because the leaving group does not interact with potential substrates, while contributing approximately 70% to the mass of UDP-GA, the reaction mechanism could be simplified by replacing UDP with a smaller substituent, as we did with FMOs. Since the difference between E_a values of potential SoMs is of interest, rather than absolute E_a values, the systematic absence of resonance stabilization and long-range effects of UDP does not play a role.

In the context of simplifying the UGT reaction mechanism, a good candidate for a leaving group would be small and non-polar to minimise the number of hydrogen bonds it can form. Such prerequisites are necessary in order to minimize the calculation times and avoid spurious interactions between the leaving group and the substrate. In addition, the interactions would be unrealistic because, in practice, the protein would restrict the movement of both substrate and UGT-GA and, in the model, the E_a of transition states which include spurious hydrogen bonds would be artificially lowered. A good candidate would be a methoxy group (MeO), which fulfils the requirements.

The constrained energy profiles along the reaction coordinate of the dissociation of the UDP-GA and MeO-GA bonds are shown in Figure 14. The reaction coordinates were generated using the results from single-point energy calculations in which the bond length between the GA and leaving group was gradually increased from 1.4 to 6.0 Å. The comparison of the energetics along the reaction coordinates demonstrates that, compared to UDP, the MeO is a poorer leaving group. MeO is a weak acid (strong conjugate base), and it lacks the resonance stabilization of a phosphate group. However, the shape of the dissociation curves is very similar, and both plateau around 6.0 Å, MeO energy values being around 50 kJ mol⁻¹ higher on average. The difference in the leaving group will slightly alter the transition state structures as well, but the differences are taken into account by using the relative energy values. In addition, the gradient of the two curves is almost identical, suggesting that the MeO group could be used as a replacement for UDP, but its E_a value will always be systematically higher.

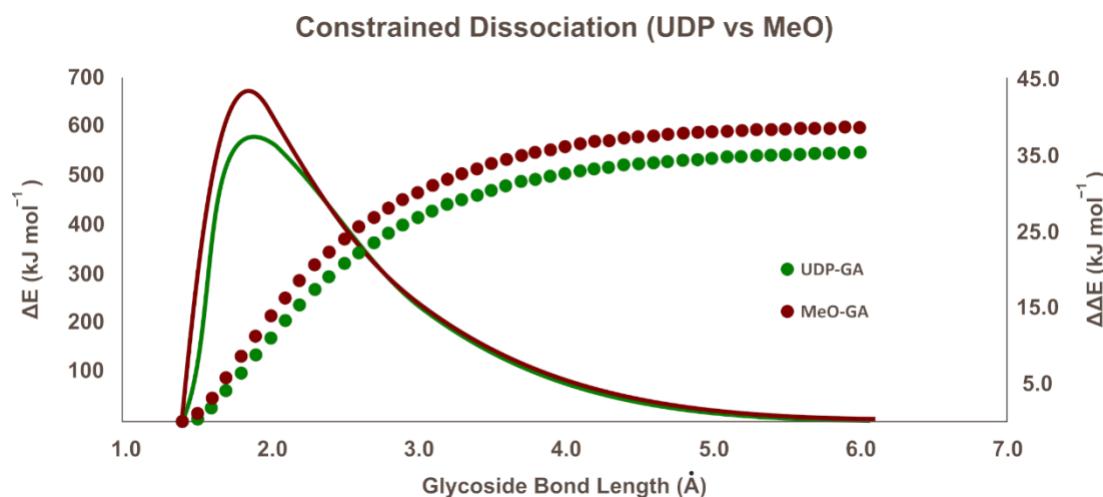


Fig. 14 The frozen energies (dots, y-axis on the left) and energy gradients along the reaction coordinates (lines, y-axis on the right) of the dissociation of UDP-GA and MeO-GA.

While it is not explicitly known which amino acid residue acts as an acid for the glucuronidation reaction for each isoform, based on the literature, it is clear that the reaction proceeds through an acid-base mechanism. There are three distinct possibilities to address the proton transfer, each one of them having their pros and cons: use two water molecules as an acid and base, use two amino acid residues (e.g. two histidine residues) or simplify the acid-base approach and use the deprotonated substrate.

From the experimental perspective, water molecules would be a balanced choice. On one hand, they are ideal for proton transfer and, theoretically, the active site of the protein will include water molecules. On the other hand, water molecules can form spurious hydrogen-bond networks in a simplified model. The amino acid residues, while being better at representing the glucuronidation, increase the calculation time considerably and also give rise to potential spurious hydrogen bonds. Simplifying the acid-base approach by deprotonating the substrate before the transition state calculations would yield the fastest calculation times but would not take the proton transfer into account explicitly.

We explored a series of systems with which to model the transition state as described in the Supporting Information. All of the model systems produced very similar transition state structures, therefore the deprotonated

system (shown in Fig. 15) it was chosen because it did not form any spurious hydrogen bonds during geometry optimisation and the calculations were quick to converge. In addition, as with other systems, it did take the proton transfer into account, although implicitly. In the following paragraphs, the reaction mechanism will be validated with experimental data, proving that the deprotonated substrate transition state model works.

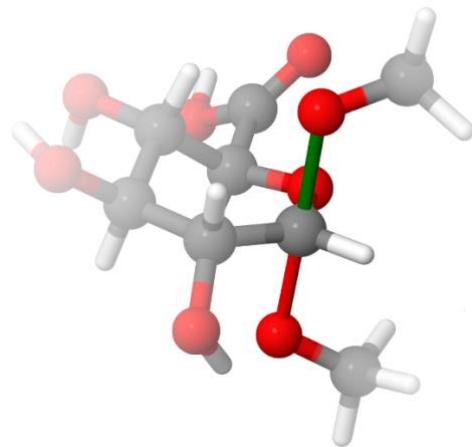


Fig. 15 The glucuronidation transition state with the deprotonated substrate. The green and red bonds illustrate the forming and breaking bonds, respectively.

The transition state energy profile for conjugation of a common substrate of UGTs, 7-hydroxyflavonol at the hydroxy group in the seventh position (Figure 6), is depicted in Figure 16. The E_a , calculated using the simplified methoxy model and a deprotonated oxygen, is 152 kJ mol^{-1} . Surprisingly, the energy of the products of the reaction is not considerably lower than that of the reactants. A larger sample size would be needed to confirm this, but glucuronides are known to be hydrolysed to form the original substrate and GA or to go through acyl migration, [80] which would be consistent with a small difference in the total energy values of the reactants and products. The bond lengths of the breaking and forming bonds of the anomeric carbon were very similar to the bond lengths observed in the transition state structures presented in the previous section.

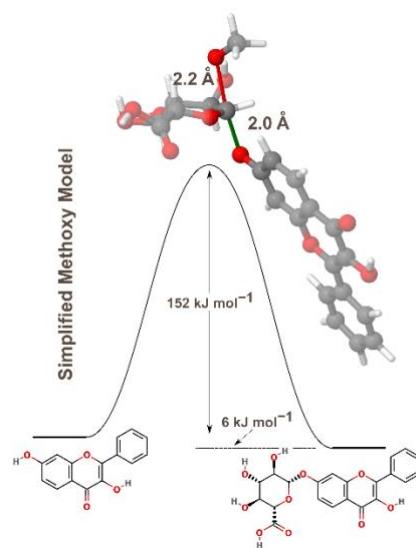


Fig. 16 The energy profile of the glucuronidation of 3,7-dihydroxyflavone.

To validate the simplified reaction mechanism for O-glucuronidation, experimentally measured reaction rates for eleven sites of metabolism within six compounds were used (Table 1). As stated before, the compounds are structurally very similar to minimize variations due to differences in accessibility of the sites of metabolism due to

variations in interactions with the enzyme active site; thus, it is anticipated that the majority of the differences in the observed rates will only be due to variations in the reactivity of the SoMs. In addition, the results have been measured by the same laboratory using the same assay conditions to minimize inter-laboratory variations, which could lead to different reaction rates.

The results, shown in Figure 10, demonstrate an approximate linear relationship between the logarithm of the reaction rates and E_a values. However, two outliers can be seen in Figure 16, which lower the squared Pearson correlation coefficient to 0.34. Interestingly, both of those points correspond to the 3,6,4'-flavone which differs from other flavones by the addition of a hydroxy group in the sixth position. In addition, the point with the worst correlation is the E_a value for the hydroxy group in the sixth position. One possible explanation is that the 3,6,4'-flavone forms stronger interactions in the active site of UGT1A1, resulting in a slower release of the formed glucuronide [81] or it could be an orientation effect due to an interaction of the hydroxy at the sixth position with the active site. Nonetheless, these results show a trend, which confirms that it is possible to differentiate between the more and less labile SoMs using the simplified reaction mechanism. The correlation coefficient excluding the outliers would be 0.68. Note that we would not expect a perfect correlation between the E_a and the observed rates of metabolism because these will also be influenced by steric accessibility effects and the trend suggests that the calculated activation energy would be a suitable descriptor in a reactivity-accessibility model for UGTs. The correlation value using the dispersion corrected functional with a higher basis set was 0.43. The correlation value using the higher level of theory, but excluding the outliers, would be 0.68 as well. The use of dispersion corrections and a higher basis set did not significantly improve the overall level of correlation, thus as with FMOs the use of a higher basis set is not necessary.

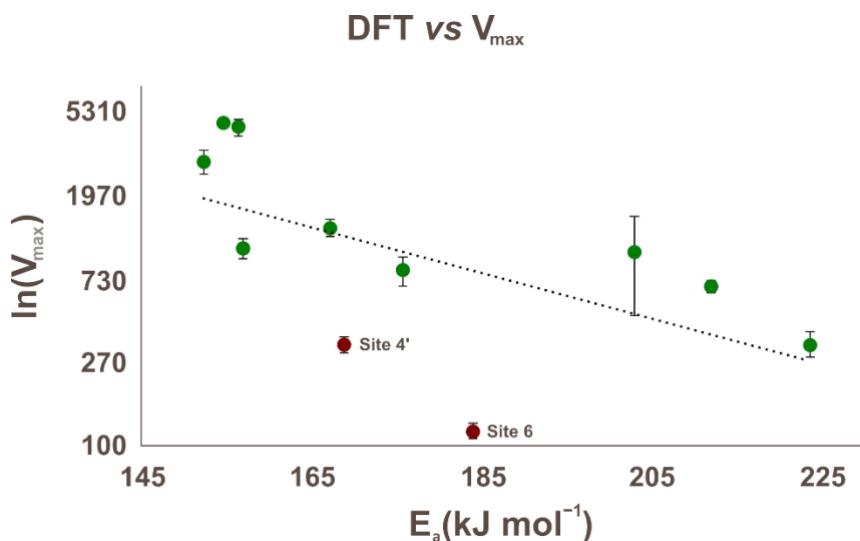


Fig. 17 The correlation between the activation energy and reaction rates for glucuronidation of flavonoids by UGT1A1. The outliers, potential SoMs 6 and 4' of 3,6,4'-flavone, have been marked with the colour red.

In the case of *N*-glucuronidation, trifluoperazine was studied, which is seen in Figure 7 and where the experimentally observed site also had the lowest calculated E_a value amongst all potential sites of metabolism with the E_a of 229 kJ mol⁻¹. The glucuronidation of the tertiary amine in the thiazine group had the E_a of 276 kJ mol⁻¹, and the glucuronidation of the other tertiary amine in the piperazine group had the E_a of 241 kJ mol⁻¹. It must be noted that the experimentally observed SoM depends on both reactivity and accessibility; thus, the site with the lowest E_a value is not always glucuronidated.

Conclusions

FMOs are important modification phase enzymes, assisting CYPs in the oxidation of both endo- and xenobiotic compounds. UGTs are considered the second most important enzyme class after CYPs for the general metabolism of

drug-like molecules. In this study we present simplified transition state systems for the S_N2 reaction mechanisms of both FMOs and UGTs. The E_a calculated by these mechanisms can be used to predict the lability of the potential SoMs for both enzyme classes. The use of DFT, as opposed to classical QSAR methods, provides generality and transferability, since the results are derived from fundamental physical principles. In addition, the use of a full substrate, rather than just a fragment, treats each molecule in its entirety when assessing the likelihood of metabolism; thus, each potential site of metabolism is considered in the context of the whole molecular environment.

For FMOs, the E_a proved to be an excellent descriptor and all experimentally observed sites of metabolism had the lowest E_a value within a molecule across the 5 molecules and 13 potential SoMs used in this study. For UGTs a data set of experimental V_{max} values was obtained and the calculated E_a correlated with the experimental data points. In the case of N-glucuronidation, the SoM was correctly classified, but it simply served as an example that the same reaction works with both N- and O-glucuronidation and is not statistically significant. The results show that the reactivity of N- and S-oxidation and N- and O-glucuronidation can be described with the simplified reaction mechanism to an acceptable level of correlation.

As previously said, most outstanding metabolism prediction models take into account both reactivity and accessibility. In order to successfully predict the metabolism for different isoforms of FMOs and UGTs, reactivity models, as presented in this work, should be combined with a method for assessing the accessibility of each potential SoM, such as a ligand-based approach. Such a model should be trained and tested using a larger number of compounds for which the experimentally observed SoMs have been determined. The current work resulted in an experimentally validated reactivity models, which give a strong foundation for generating combined reactivity-accessibility models (E_a combined with steric- and orientation descriptors) for FMO and UGT metabolism in the future.

References

- [1] Xu C, Li CY-T, Kong A-N T (2005) Induction of phase I, II and III drug metabolism/transport by xenobiotics. *Arch Pharm Res* 28:249–68
- [2] Liston HL, Markowitz JS, DeVane CL (2001) Drug glucuronidation in clinical psychopharmacology. *J Clin Psychopharmacol* 21(5):500–515
- [3] Miners JO, Mackenzie PI (1991) Drug glucuronidation in humans. *Pharmacol Ther* 51(3):347–369
- [4] Miners JO, Smith PA, Sorich MJ, Mckinnon RA, Mackenzie PI (2004) Predicting human drug glucuronidation parameters: application of in vitro and in silico modelling approaches. *Annu Rev Pharmacol Toxicol* 44(1):1–25
- [5] Tyzack JD, Hunt PA, Segall MD (2016) Predicting Regioselectivity and Lability of Cytochrome P450 Metabolism Using Quantum Mechanical Simulations. *J Chem Inf and Modelling* 56(11):2180–2193
- [6] Guengerich FP (2006) Cytochrome P450s and other enzymes in drug metabolism and toxicity. *AAPS Journal* 8(1):101–111
- [7] Dixit VA, Lal LA, Agrawal SR (2017) Recent advances in the prediction of non-CYP450-mediated drug metabolism. *WIREs Comput Mol Sci* e1323:1–43
- [8] Olsen L, Montefiori M, Tran KP, Jørgensen FS (2019) SMARTCyp 3.0: enhanced cytochrome P450 site-of-metabolism prediction server. *Bioinformatics* 35(17):3174–3175
- [9] Cruciani G, Carosati E, De BoeckB, Ethirajulu K, Mackie C, Howe T, Vianello R (2005) MetaSite: Understanding Metabolism in Human Cytochromes from the Perspective of the Chemist, *Journal of Medicinal Chemistry*, 48(22), 6970–6979
- [10] Hennemann M, Friedl A, Lobell M, Keldenich J, Hillisch A, Clark T, Göller A (2009) CypScore: Quantitative Prediction of Reactivity toward Cytochromes P450 Based on Semiempirical Molecular Orbital Theory, *ChemMedChem*, 4(4):657–669
- [11] Zaretzki J, Matlock M, Swamidass S J (2013) XenoSite: Accurately Predicting CYP-Mediated Sites of Metabolism with Neural Networks, *Journal of Chemical Information and Modeling*, 53(12):3373–3383

- [12] Šícho M, de Bruyn Kops C, Stork C, Svozil D, Kirchmair J (2017) FAME 2: Simple and Effective Machine Learning Model of Cytochrome P450 Regioselectivity. *Journal of Chemical Information and Modelling*, 57(8):1832–1846
- [13] Phillips I, Shephard E (2016) Drug metabolism by flavin-containing monooxygenases of human and mouse. *Expert Opinion on Drug Metabolism & Toxicology* 13(2):167–181
- [14] Alfieri A, Malito E, Orru R, Fraaije M, 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(18):6572–6577
- [15] Eswaramoorthy S, Bonanno J, Burley S, Swaminathan S (2006) Mechanism of action of a flavin-containing monooxygenase. *Proceedings of the National Academy of Sciences* 103(26), 9832–9837
- [16] Mackenzie PI, Owens IS, Burchell B, Bock KW, Bairoch A, Bélanger A, Fournel-Gigleux S, Green M, Hum DW, Iyanagi T, Lancet D, Louisot P, Magdalou J, Chowdhury JR, Ritter JK, Schachter H, Tephly TR, Tipton KF, Nebert DW (1997) The UDP glycosyltransferase gene superfamily: recommended nomenclature update based on evolutionary divergence. *Pharmacogenetics* 7(4):255–269
- [17] Nair PC, Meech R, Mackenzie PI, McKinnon RA, Miners JO (2015) Insights into the UDP-sugar selectivities of human UDP-glycosyltransferases (UGT): a molecular modelling perspective. *Drug Metab Rev* 47(3):335–345
- [18] Bock KW (2016) The UDP-glycosyltransferase (UGT) superfamily expressed in humans, insects and plants: Animal–plant arms-race and co-evolution. *Biochemical Pharmacology* 99:11–17
- [19] Mackenzie P, Rogers A, Treloar J, Jorgensen BR, Miners JO, Meech R (2008) Identification of UDP Glycosyltransferase 3A1 as a UDP N-Acetylglucosaminyltransferase. *J of Biological Chemistry* 283(52):36205–36210
- [20] Bosio A, Binczek E, Le Beau MM, Fernald AA, Stoffel W (1996) The Human Gene CGT Encoding the UDP-Galactose Ceramide Galactosyl Transferase (Cerebroside Synthase): Cloning, Characterization, and Assignment to Human Chromosome 4, Band q26. *Genomics* 34(1):69–75
- [21] Banerjee R, Pennington MW, Garza A, Owens IS (2008) Mapping the UDP-glucuronic acid binding site in UDP-glucuronosyltransferase-1A10 by homology-based modelling: confirmation with biochemical evidence. *Biochemistry* 47(28):7385–7392
- [22] Miley MJ, Zielinska AK, Keenan JE, Bratton SM, Radominska-Pandya A, Redinbo MR (2007) Crystal structure of the cofactor-binding domain of the human phase II drug-metabolism enzyme UDP-glucuronosyltransferase 2B7. *J Mol Biol* 369(2):498–511
- [23] Rendic S, Guengerich FP (2014) Survey of Human Oxidoreductases and Cytochrome P450 Enzymes Involved in the Metabolism of Xenobiotic and Natural Chemicals. *Chemical Research in Toxicology* 28(1):38–42
- [24] Krueger SK, Williams DE (2005) Mammalian flavin-containing monooxygenases: structure/function, genetic polymorphisms and role in drug metabolism. *Pharmacol Ther* 106(3):357–387
- [25] Fu C-W, Lin T-H (2017) Predicting the Metabolic Sites by Flavin-Containing Monooxygenase on Drug Molecules Using SVM Classification on Computed Quantum Mechanics and Circular Fingerprints Molecular Descriptors. *PLoS One* 12(1): e0169910
- [26] Guengerich FP (2008) Cytochrome p450 and chemical toxicology. *Chem Res Toxicol* 21(1):70–83
- [27] Williams JA, Hyland R, Jones BC, Smith DA, Hurst S, Goosen TC, Peterkin V, Koup JR, Ball SE (2004) Drug-drug interactions for UDP-glucuronosyltransferase substrates: a pharmacokinetic explanation for typically observed low exposure (AUC_i/AUC) ratios. *Drug Metab Dispos* 32(11):1201–1208
- [28] Sorich MJ, McKinnon RA, Miners JO, Smith PA (2006) The importance of local chemical structure for chemical metabolism by human uridine 5'-diphosphate–glucuronosyltransferase. *J Chem Inf Model* 46(6):2692–2697
- [29] Dang NL, Hughes TB, Krishnamurthy V, Swamidass SJ (2016) A simple model predicts UGT-mediated

metabolism. *Bioinformatics* 32(20):1–7

- [30] Mazzolari A, Afzal AM, Pedretti A, Testa B, Vistoli G, Bender A (2019) Prediction of UGT-mediated Metabolism Using the Manually Curated MetaQSAR Database. *ACS Medical Chemistry Letters* 10(4):633–638
- [31] Lang DH, Rettie AE (2000) In vitro evaluation of potential in vivo probes for human flavin-containing monooxygenase (FMO): metabolism of benzydamine and caffeine by FMO and P450 isoforms. *Br J Clin Pharmacol* 50(4):311–314
- [32] Mushiroda T, Douya R, Takahara E, Nagata O (2000) The involvement of flavin-containing monooxygenase but not CYP3A4 in metabolism of itopride hydrochloride, a gastroprotective agent: comparison with cisapride and mosapride citrate. *Drug Metab Dispos* 28(10):1231–1237
- [33] Rawden HC, Kokwaro GO, Ward SA, Edwards G (2000) Relative contribution of cytochromes P-450 and flavin-containing monooxygenases to the metabolism of albendazole by human liver microsomes. *Br J Clin Pharmacol* 49(4):313–322
- [34] Overby LH, Carver GC, Philpot RM (1997) Quantitation and kinetic properties of hepatic microsomal and recombinant flavin-containing monooxygenases 3 and 5 from humans. *Chem Biol Interact* 106(1):29–45
- [35] Hai X, Adams E, Hoogmartens J, Van Schepdael A (2009) Enantioselective in-line and off-line CE methods for the kinetic study on cimetidine and its chiral metabolites with reference to flavin-containing monooxygenase genetic isoforms. *Electrophoresis* 30(7):1248–57
- [36] Cashman JR, Park SB, Yang ZC (1993) Chemical, enzymatic, and human enantioselective S-oxygenation of cimetidine. *Drug Metab Dispos* 21(4):587–597
- [37] Taylor KH, Ziegler DM (1987) Studies on substrate specificity of the hog liver flavin-containing monooxygenase: anionic organic sulfur compounds. *Biochem Pharmacol* 36(1):141–146
- [38] Ottolina G, Gonzalo G, Carrea G (2005) Theoretical studies of oxygen atom transfer from flavin to electron-rich substrates. *Journal of Molecular Structure: THEOCHEM* 757(1-3):175–181
- [39] Bach R (2011) Role of the Somersault Rearrangement in the Oxidation Step for Flavin Monooxygenases (FMO). A Comparison between FMO and Conventional Xenobiotic Oxidation with Hydroperoxides. *The Journal of Physical Chemistry A* 115(40):11087–11100
- [40] Hawes EM (1998) *N*⁺-Glucuronidation, a Common Pathway in Human Metabolism of Drugs with a Tertiary Amine Group: 1996 ASPET N-Glucuronidation of Xenobiotics Symposium. *Drug Metab Dispos* 26(9):830–837
- [41] Nishiyama T, Kobori T, Arai K, Ogura K, Ohnuma T, Ishii K, Hayashi K, Hiratsuka A (2006) Identification of human UDP-glucuronosyltransferase isoform (s) responsible for the C-glucuronidation of phenylbutazone. *Archives of Biochemistry and Biophysics* 454(1):72–79
- [42] Buchheit D, Schmitt El, Bischoff D, Ebner T, Bureik M (2011) S-Glucuronidation of 7-mercaptop-4-methylcoumarin by human UDP glycosyltransferases in genetically engineered fission yeast cells. *Biol Chem* 392(12):1089–1095
- [43] Foti RS, Fisher MB (2005) Assessment of UDP-glucuronosyltransferase catalyzed formation of ethyl glucuronide in human liver microsomes and recombinant UGTs. *Forensic Science International* 153(2–3):109–116
- [44] Kuehl GE, Bigler J, Potter JD, Lampe JW (2006) Glucuronidation of the aspirin metabolite salicylic acid by expressed UDP-glucuronosyltransferases and human liver microsomes. *Drug Metabolism and Disposition* 34(2):199–202
- [45] Kuehl GE, Murphy SE (2003) *N*-glucuronidation of nicotine and cotinine by human liver microsomes and heterologously expressed UDP-glucuronosyltransferases. *Drug Metabolism and Disposition* 31(11):1361–1368
- [46] Jin C-J, Miners JO, Burchell B, Mackenzie PI (1993) The glucuronidation of hydroxylated metabolites of benzo [α] pyrene and 2-acetylaminofluorene by cDNA-expressed human UDP-

glucuronosyltransferases. *Carcinogenesis* 14(12):2637–2639

- [47] Senafi SB, Clarke DJ, Burchell B (1994) Investigation of the substrate specificity of a cloned expressed human bilirubin UDP-glucuronosyltransferase: UDP-sugar specificity and involvement in steroid and xenobiotic glucuronidation. *Biochem J* 303(1):233–240
- [48] Jin C, Miners JO, Lillywhite KJ, Mackenzie PI (1993) Complementary deoxyribonucleic acid cloning and expression of a human liver uridine diphosphate-glucuronosyltransferase glucuronidating carboxylic acid-containing drugs. *The Journal of Pharmacology and Experimental Therapeutics* 264(1):475–479
- [49] Yang N, Sun R, Liao X, Aa J, Wang G (2017) UDP-glucuronosyltransferases (UGTs) and their related metabolic cross-talk with internal homeostasis: A systematic review of UGT isoforms for precision medicine. *Pharmacological Research* 121:169–183
- [50] Tukey RH, Strassburg CP (2000) Human UDP-glucuronosyltransferases: metabolism, expression, and disease. *Annu Rev Pharmacol Toxicol* 40:581–616
- [51] Li C, Wu Q (2007) Adaptive evolution of multiple-variable exons and structural diversity of drug-metabolizing enzymes. *BMC Evol Biol* 7(69):1–20
- [52] Bosma PJ, Seppen J, Goldhoorn B, Bakker C, Oude Elferink RP, Chowdhury JR, Chowdhury NR, Jansen PLM (1994) Bilirubin UDP-glucuronosyltransferase 1 is the only relevant bilirubin glucuronidating isoform in man. *J Biol Chem* 269:17960–17964
- [53] Fujiwara R, Yokoi T, Nakajima M (2016) Structure and protein–protein interactions of human UDP-glucuronosyltransferases. *Front Pharmacol* 7:1–15
- [54] Lairson LL, Henrissat B, Davies GJ, Withers SG (2008) Glycosyltransferases: structures, functions, and mechanisms. *Annu Rev Biochem* 77:521–555
- [55] Liang D-M, Liu J-H, Wu H, Wang B-B, Zhu H-J, Qiao J-J (2015) Glycosyltransferases: mechanisms and applications in natural product development. *Chem Soc Rev* 44(22):8350–8374
- [56] Locuson CW, Tracy TS (2007) Comparative modelling of the human UDP-glucuronosyltransferases: insights into structure and mechanism. *Xenobiotica* 37(2):155–168
- [57] Ouzzine M, Antonio L, Burchell B, Netter P, Fournel-Gigleux S, Magdalou J (2000) Importance of histidine residues for the function of the human liver UDP-glucuronosyltransferase UGT1A6: evidence for the catalytic role of histidine 370. *Mol. Pharmacol* 58(6):1609–1615
- [58] Li D, Fournel-Gigleux S, Barré L, Mulliert G, Netter P, Magdalou J, Ouzzine M (2007) Identification of aspartic acid and histidine residues mediating the reaction mechanism and the substrate specificity of the human UDP-glucuronosyltransferases 1A. *J Biol Chem* 282:36514–36524
- [59] Radominska-Pandya A, Czernik PJ, Little JM, Battaglia E, Mackenzie PI (1999) Structural and functional studies of UDP-glucuronosyltransferases. *Drug Metab Rev* 31(4):817–899 [41]
- [60] Hanwell MD, Curtis DE, Lonié DC, Vandermeersch T, Zurek E, Hutchison GR (2012) Avogadro: an advanced semantic chemical editor, visualization, and analysis platform. *J Cheminf* 4:1–17
- [61] Sadowski J, Gasteiger J (1993) From atoms and bonds to three-dimensional atomic coordinates: automatic model builders. *J Chem Rev* 93(7):2567–2581
- [62] Dewar MJS, Zoebisch EG, Healyand EF, Stewart JJP (1985) Development and use of quantum mechanical molecular models. 76. AM1: a new general purpose quantum mechanical molecular model. *J Am Chem Soc* 107(13):3902–3909
- [63] Stewart JJP (1990) MOPAC: A general molecular orbital package. *Quant Chem Prog Exch* 10:86
- [64] Becke AD (1993) Density-functional thermochemistry. III. The role of exact exchange. *JChemPhys* 98:5648–5652
- [65] Lee C, Yang W, Parr RG (1988) Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys Rev B* 37:785–789
- [66] Vosko SH, Wilk L, Nusair M (1980) Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis. *Can J Phys* 58:1200–1211

- [67] Stephens PJ, Devlin FJ, Chabalowski CF, Frisch MJ (1994) Ab initio calculation of vibrational absorption and circular dichroism spectra using density functional force fields. *J Phys Chem* 98(45):11623–11627
- [68] Weigend F, Ahlrichs R (2005) Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Phys Chem Chem Phys* 7(18):3297–3305
- [69] Simón L, Goodman JM (2011) How reliable are DFT transition structures? Comparison of GGA, hybrid-meta-GGA and meta-GGA functionals. *Organic and Biomolecular Chemistry* 9(3):689–700
- [70] Grimme S, Antony J, Ehrlich S, Krieg H (2010) A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *The J of Chem Phys* 132(15):154104–154119
- [71] Valiev M, Bylaska EJ, Govind N, Kowalski K, Straatsma TP, van Dam HJJ, Wang D, Nieplocha J, Apra E, Windus TL, de Jong WA (2010) NWChem: A comprehensive and scalable open-source solution for large scale molecular simulations. *Comput Phys Commun* 181(9):1477–1489
- [72] Lu T, Chen F (2012) Multiwfn: a multifunctional wavefunction analyser *Journal of Computational Chemistry* 33:580–592
- [73] Hirshfeld FL (1977) Bonded-atom fragments for describing molecular charge densities. *Theoretica chimica acta* 44:129–138
- [74] Leoni C, Buratti F, Testai E (2008) The participation of human hepatic P450 isoforms, flavin-containing monooxygenases and aldehyde oxidase in the biotransformation of the insecticide fenthion. *Toxicology and Applied Pharmacology* 233(2):343–352
- [75] Kim Y, Ziegler D (2000) Size Limits of Thiocarbamides Accepted as Substrates by Human Flavin-Containing Monooxygenase 1. *Drug Metabolism and Disposition* 28(8):1003–1006
- [76] Wu B, Xu B, Hu M (2011) Regioselective glucuronidation of flavonols by six human UGT1A isoforms. *Pharm Res* 28:1905–1918
- [77] Wu B, Zhang S, Hu M (2011) Evaluation of 3,3',4'-Trihydroxyflavone and 3,6,4'-Trihydroxyflavone (4'-O-Glucuronidation) as the in Vitro Functional Markers for Hepatic UGT1A1. *Mol Pharm* 8(6):2379–2389
- [78] Kubota T, Lewis BC, Elliot DJ, Mackenzie PI, Miners JO (2007) Critical roles of residues 36 and 40 in the phenol and tertiary amine aglycone substrate selectivities of UDP-glucuronosyltransferases 1A3 and 1A4. *Molecular Pharmacology* 72(4):1054–1062
- [79] Laakkonen L, Finel M (2010) A molecular model of the human UDP-glucuronosyltransferase 1A1, its membrane orientation, and the interactions between different parts of the enzyme. *Mol Pharmacol* 77(6):931–939
- [80] Lassila T, Hokkanen J, Aatsinki S-M, Mattila S, Turpeinen M, Tolonen A (2015) Toxicity of carboxylic acid-containing drugs: the role of acyl migration and CoA conjugation investigated. *Chemical Research in Toxicology* 28(12):2292–2303
- [81] Juovonen R, Rauhamäki S, Kortet S, Niinivehmas S, Troberg J, Petsalo A, Huuskonen J, Raunio H, Finel M, Pentikäinen OT (2018) Molecular docking-based design and development of a highly selective probe substrate for UDP-glucuronosyltransferase 1A10. *Molecular Pharmaceutics* 15(3):923–933