From Experiments to a Fast Easy-to-Use Computational Methodology to Predict Human Aldehyde Oxidase Selectivity and Metabolic Reactions

(Article begins on next page)

26 April 2024

From experiments to a fast easy-to-use computational methodology to predict human aldehyde oxidase selectivity and metabolic reactions

Gabriele Cruciani,[1]* Nicolò Milani,[1] Paolo Benedetti,[1] Susan Lepri,[1] Lucia Cesarini,[1] Massimo Baroni,[2] Francesca Spyrakis,[3] Sara Tortorella,[4] Edoardo Mosconi[5] and Laura Goracci[1]

[1]Department of Chemistry, Biology and Biotechnology, University of Perugia, via Elce di Sotto 8, 06123 Perugia, Italy.

[2]Molecular Discovery Ltd, Centennial Park, Borehamwood, Hertfordshire, United Kingdom.

[3] Department of Drug Science and Technology, University of Turin, via P. Giuria 9, 10125 Turin, Italy.

[4]Molecular Horizon srl, via Montelino 32, 06084 Bettona, Italy.

[5] Computational Laboratory for Hybrid/Organic Photovoltaics, National Research Council–Institute of Molecular Science and Technologies, Via Elce di Sotto 8, I-06123 Perugia, Italy.

ABSTRACT

Aldehyde oxidase (AOX) is a molibdo-flavoenzyme that raised great interest in recent years, since its contribution in xenobiotics metabolism has not always been identified before clinical trials, with consequent negative effects on the fate of new potential drugs. The fundamental role of AOX in metabolizing xenobiotics is also due to the attempt of medicinal chemists to stabilize candidates toward cytochrome P450 activity, which increases the risk for new compounds to be susceptible to AOX nucleophile attack. Therefore, novel strategies to predict the potential liability of new entities towards AOX enzyme are urgently needed to increase effectiveness, reduce costs, and prioritize experimental studies. In the present work, we present the most up-to-date computational method to predict liability towards human AOX, for applications in drug design and pharmacokinetic optimization. The method was developed using a large dataset of homogeneous experimental data, which is also disclosed as supplementary material.

INTRODUCTION

Evaluation of the metabolism of a new drug or drug candidate in pharmaceutical companies is essential in measuring both safety and efficacy. The development of reliable screening systems has produced a 10% decrease in the rate of pharmaceutical drug candidate clinical trials interruption.[1]

At the same time, the development of *in silico* methods to predict the metabolism of xenobiotics has made the chemical discovery phase much more effective.[2-5] Nowadays, medicinal chemists have various working strategies available to overcome the issue of unstable candidates with respect to the P450 cytochrome (CYP) family. Once the site of metabolism (SoM) in a molecule has been identified using fast experimental (MetID via Mass-MetaSite)[6] or in silico methods,[2-4] the metabolism of the molecule can be modified by directly changing the SoM or different molecular regions. For instance, the addition of fluorine, halogens, or deuterium to the SoM, when possible,

can significantly modify the pharmacokinetic profile of the compound. The substitution of a known labile group with a more resistant one towards oxidation has a significant benefit. Secondly, the medicinal chemist must initially understand which region of the molecule largely contributes to the exposure of the molecular site where the reaction takes place so that it can be modified. In both cases, aryl groups or carbocyclic aromatics possibly present in the scaffold can be substituted by heteroaromatic rings that, in addition to a positive impact on the clearance, can also impact on solubility and physical-chemical properties such as LogD.

Nevertheless, if understanding the mechanisms responsible for molecules liability has allowed medical chemists to become more effective in planning compounds stable to CYP enzymes, it has also abundantly increased the number of compounds that can be potentially metabolized by non-CYP enzymes. This means that the contribution of non-CYP metabolism is now very significant in modern pharmaceutical research and is steadily increasing in importance.[7] In particular, studies of the metabolism mediated by aldehyde oxidase (AOX) have assumed a prominent role as demonstrated by the numerous papers on this subject (about 400 papers published in the last few years),[8] and by the various experimental strategies developed to determine the propensity of a compound to be an AOX substrate.[9] AOX is a cytosolic enzyme that catalyzes oxidation by a nucleophile mechanism using water as a source of oxygen that is incorporated in the metabolite. AOX also shows significant ability to hydrolyze the amide bonds that are generally metabolically stable and hardly modified by CYP enzymes.

But why should medicinal chemists matter about AOX? This effort is actually encouraged by the numerous recent interruptions of clinical trials due to AOX metabolism. In particular, the clinical trials of candidates such as BIBX1382, FK3453, RO-1, and carbazeran have been suspended because of poor bioavailability and rapid elimination in humans,[10-13] while SGX-523 was interrupted due to renal insufficiency following aldehyde oxidase metabolism in humans.[14] In reply to the growing role assumed by AOX, the scientific community in pharmaceutical companies has

activated appropriate screening programs to evaluate the metabolic instability caused by new chemical entities in different biomatrices. However, a more global strategy is needed to increase efficacy, reduce costs, and to prioritize the experimental study of new candidates. In this context, *in silico* methods can offer a valuable strategy to model and predict the role of AOX towards potential new compounds.

In the present work, we present a novel computational method to predict liability towards human AOX. This new model combines the knowledge acquired from both experimental and *in silico* sources. Indeed, more than 600 compounds (acquired or synthesized to meet the needs of assessing structure-activity relationships) were experimentally tested in human liver cytosol (HLC) to generate the largest database of homogeneous experimental data publically available so far. In addition, our experience in modelling metabolism by phase I metabolic enzymes (CYP and Flavin-containing monooxygenase 3, or FMO3) in terms of chemical reactivity and exposure to the enzyme represented the background to develop a new model to predict human aldehyde oxidase (*h*AOX) metabolism. To better describe why we decided to address our efforts towards the AOX modelling, examples of ineffective optimization recently occurred are analyzed. Subsequently, correlation of AOX liability with lipophilicity (LogP and LogD) and electronic parameters (Parr's index) is discussed. Finally, our proposed model for AOX liability prediction and its validation are described.

RESULTS AND DISCUSSION

From a CYP substrate to an AOX substrate: examples of ineffective optimization

Here we report two real cases of metabolic shift from CYP to AOX recently occurred in our laboratories. First, when compound **1** (Figure 1-a), planned to interfere with several nuclear receptors, was incubated in human liver microsomes (HLM), it demonstrated notable metabolic instability. After 60 minutes the compound was totally degraded into three metabolites, with the

major one being reported in Figure 1a as **1M**. Incubation of **1** with recombinant CYP isoforms demonstrated the preeminent role of the enzyme CYP-3A4 (data not shown).
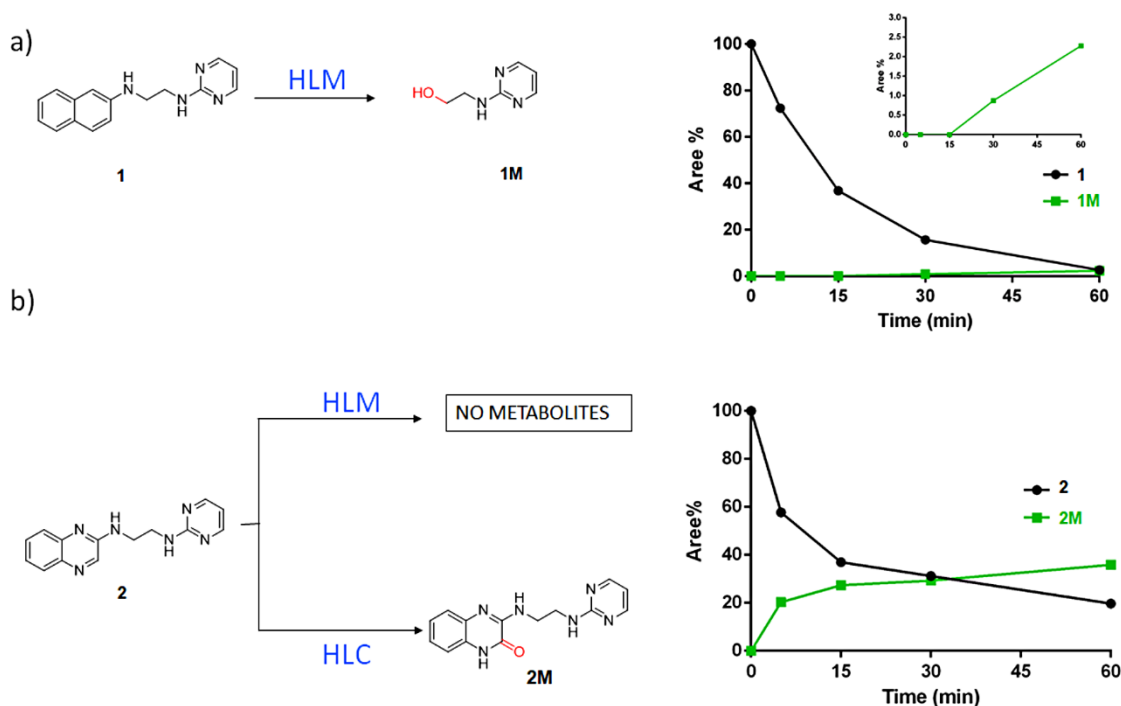

Figure 1. Metabolic stability assessment of compounds **1** and **2**. a) Metabolic stability assessment of compound **1** in HLM. The compound was rapidly metabolized in the ethylendiamine portion to give compound **1M**. b) Metabolic stability assessment of compound **2** in HLM (no metabolites observed) and HLC (rapid formation of metabolite **2M**). In the kinetic profiles, substrates and metabolites are colored in black and green, respectively.

An analysis of the reactive pose of compound **1** in CYP-3A4 performed with MetaSite[3,15,16] demonstrated the fundamental role of the naphthyl group in molecular exposure. Indeed, one of the two aromatic rings, interacting with Ile301, Phe304 and Ala305, stabilizes the docking pose and promotes the oxidation and *N*-demethylation reactions in the ethylendiamine part of the substrate (Figure 2).
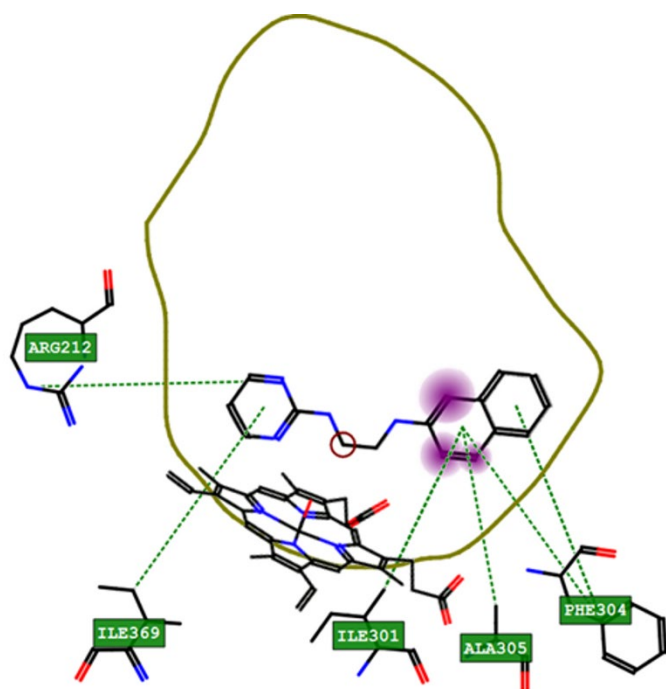
Figure 2. Docking pose of compound **1** in CYP-3A4 cavity, according to MetaSite prediction.

To make the compound metabolically more resistant, the naphthyl group was substituted with a quinoxalin group as in compound **2** (Figure 1b) that notably stabilized the substrate in HLM biological matrix. However, an analysis of the newly synthesized compound **2** in human cytosol HLC demonstrated the rapid degradation of the compound to give the **2M** metabolite (Figure 1b). Moreover, attempting to make the compound more stable to the biotransformation induced by the cytochrome favored the attack by the AOX enzyme. Therefore, the problem has not been solved but only shifted and in some ways, dangerously hidden.

A similar situation might easily occur with another fundamental organic group to the medicinal chemist, that is, the amide group. Amide groups have typically been inserted in molecules of interest to improve chemical stability, rigidity, and bond directionality, and to modulate polarity and bioavailability. However, it is now clear that amide groups can be rapidly cleaved *in vivo* by AOX enzyme.[17,18]

In this context, the second case discussed here concerns our attempts to metabolically stabilize compound **3** (Figure 3), active in signal transduction and metabolism cascades exerting a putative anti-cancer effect (data not shown). To this aim, we decided to substitute the labile 1*H*-indol-2yl-methyl group with the 1*H*-indole-2 carboxamide group.
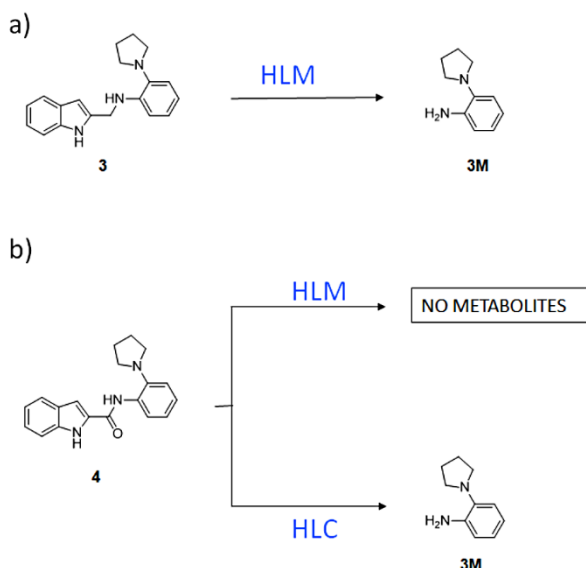

Figure 3. Metabolic stability assessment of compound **3** and **4**. a) Compound **3** was rapidly metabolized to give metabolite **3M**. b) Compound **4** was stable in HLM, nevertheless it was hydrolyzed by AOX in HLC to give, again. metabolite **3M**.

Again, although our strategy reduced the CYP-3A4-mediated metabolism, it favored the hydrolysis reaction promoted by AOX. It should be noted that this example (phenotypic shift) is even more relevant than the previous one, as the amide bond is notoriously difficult to break and is always used by medical chemists to stabilize pharmaceutical drug candidates.

These examples emphasize the fundamental importance of fully understanding the mechanisms of metabolic clearance, and the relevant differences of species when pharmaceutical drug candidates are clinically evaluated.

AOX liability prediction: which correlations are possible?

A certain number of *in silico* approaches were used in the past to predict the liability of a molecule to AOX and to predict its potential metabolic site. While the physical-chemical properties of molecules such as LogP or LogD are significant determinants of the metabolic clearance mediated by CYP,[19-22] the substrates of AOX did not demonstrated any correlations with these parameters.[23-25] To confirm this point, we calculated LogP and LogD$_{7.5}$ values for each chemical structure in our internal database, which contains more than 600 compounds experimentally tested to assess *h*AOX liability (the entire database is provided in the Supporting Information). As clearly reported in Figure 4, molecular lipophilicity has not proved to be discriminating for *h*AOX selectivity in either reaction mechanism.

Figure 4. LogP values distribution for experimentally determined AOX substrates (green) and non-substrates (red), for the oxidation (left) and hydrolysis reaction (right). LogP and LogD$_{7.5}$ values were obtained from VolSurf+.[26,27] For chemical structures see Supplementary Information.

Since the AOX-catalyzed oxidation of a heterocyclic ring initially leads to a nucleophile attack by the hydroxyl group, susceptibility to AOX can reasonably depend on the electronic deficiency of the reaction site. Consequently, descriptors of reactivity or parameters that determine electrophilicity at the metabolic site have been tested to predict susceptibility to AOX. The energy of the lowest unoccupied molecular orbital (E_{LUMO}) and the electrostatic potential of the carbon adjacent to the nitrogen atom in the heteroaromatic ring are two descriptors able to properly predict the occurrence of a nucleophile attack.[28] Ghafourian and Rashidi used these parameters to explore the quantitative structure - activity relationships (QSAR) of a group of compounds that contain a phthalazine and a quinazoline ring.[29] Recently, Dalvie et al. compared the E_{LUMO} of zoniporide, a molecule with great affinity as a substrate for AOX, and of various analogues to test the capability of these parameters in predicting AOX-substrate properties.[30] Torres et al. demonstrated how density functional theory (DFT) methodologies can be used to predict AOX metabolism sites in heteroaromatic rings.[31] Pryde et al. and Dalvie et al. used this strategy to predict the AOX lability of newly designed compounds.[25,30] More recently, the Parr's index was used to estimate molecule electrophilicity and establish a possible relationship with the observed extent of biotransformation.[32,33] However, as demonstrated in Figure 5, the Parr's electrophilicity index alone, calculated for each compound in our database, did not allow a good discrimination in substrate hAOX selectivity for both oxidation and hydrolysis reactions (see section S1 in Supporting Information for computational details).
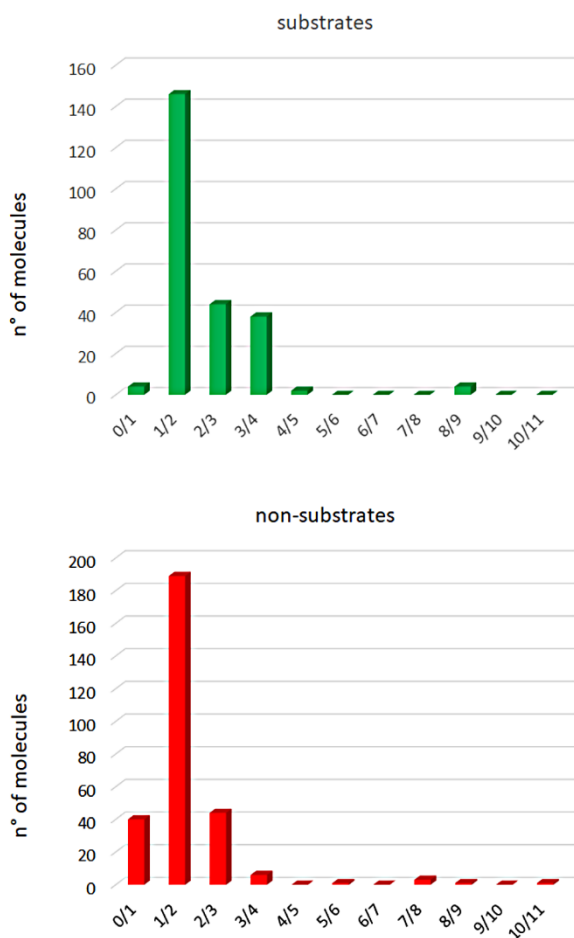
Figure 5. Distribution plot for AOX substrates (top) and non-substrates (bottom), according to Parr's electrophilicity index (ω) values (for molecular structures and experimental details see Supporting Information).

Additionally, we noticed that in some cases AOX liability cannot be rationalized solely by analyzing the initial nucleophile attack by the hydroxyl group bound to molibdenum, and thus the electrophilicity of the ligand sp2 carbon, but also considering the subsequent hydride displacement step. For instance, for a number of phthalazines derivatives an opposite trend between the AOX susceptibility and the electronic effect of the substituting groups was recently observed.[17] Indeed, as reported in Figure 6, 5-nitro-phthalazine (**5**) turned out to be significantly more stable than phthalazine itself (**6**) and than 5-amino-phthalazine (**7**), although the first compound possesses a

strong electron withdrawing group, which is supposed to decrease the electron availability of the ring. DFT analyses performed to investigate the thermodynamics of the hydride displacement[34] step and the nature of the transition state rationalized the observed trend. In particular, the calculations (Table S2-1, Section S2 in Supporting Information) showed that hydride displacement is more thermodynamically favored (ΔG_{solv}) and has a lower activation energy ($\Delta G^{\neq}_{solv}$) for 5-amino-phthalazine than for 5-nitro-phthalazine, mainly due to the possible intramolecular hydrogen bond in 5-amino-phthalazine (Figure S2-1, in Supporting Information). Computational details are provided in section S2 of the Supporting Information.
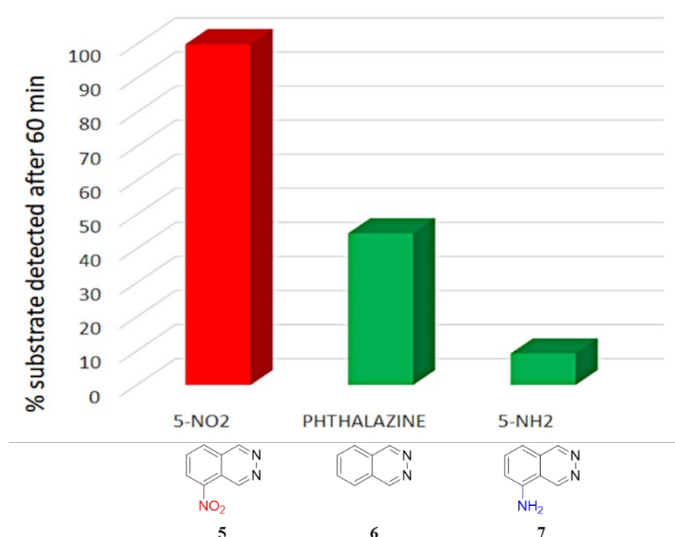

Figure 6. Metabolic stability of 5-nitro-phthalazine (**5**), phthalazine (**6**) and 5-amino-phthalazine (**7**) in HLC.

Obviously, DFT calculations, accounting for the energetics and the nature of intermediates and transition states of the overall process are not feasible (in term of time and computational cost) in a drug design approach, where hundreds of candidates are screened at the same time. However, such results demonstrate that the assumption that electron-withdrawing group (EWG) will increase electrophilicity, and therefore AOX liability, might sometimes fail. At the same time, a general

11

assumption that a nitro- substituent can lead to a more stable compound cannot be made. Indeed, 5-nitroquinoline was recently found to be a substrate of *h*AOX, and we also confirmed this data with the experimental conditions used to generate our library.[35,36] Thus, it is evident that other effects like exposure to the catalytic site must be taken into account for a more accurate prediction.

From an experimental point of view, with the intention of finding the susceptibility of heteroarenes to AOX metabolism, O'Hara et al.[37] recently reported a "Litmus test" for the early identification of AOX drug substrates. This method uses zinc difluoromethanesulfinate (DFMS) as a source of CF_2H radical to simulate AOX metabolic reaction. A binary qualitative answer (reactive or not reactive) was the desired readout of this chemical test (here called AOX-Litmus test). Searching for reliable correlations with AOX liability, we also tested nine compounds, selected among the list of experimentally proved non-substrates to cover different scaffolds (Table S3-1, Supporting Information), by the AOX-Litmus test. Interestingly, we found that for six compounds the test failed (see Figure 7 for compounds containing different heteroarene moieties (imidazo[1,2-a]pyrimidine, isoquinoline, and quinoxaline)). When compounds **8**, **10** and **12** were submitted to the AOX-Litmus test, all compounds resulted "reactive", and thus AOX susceptible substrates. However, when tested in cytosolic *h*AOX, no one of the compounds was affected by AOX and all three compounds were totally stable. Furthermore, AOX-Litmus test on compounds **10** and **12** produced different fluorinated regioisomers (Figure S3-1 in Section 3 of the Supporting Information). This fact demonstrates that chemical reactivity alone and, consequently, all those methods taking into account only chemistry cannot be used to predict the susceptibility of chemical compounds to AOX. Again, chemical reactivity is only a component of a mechanism that is biologically much more complex.
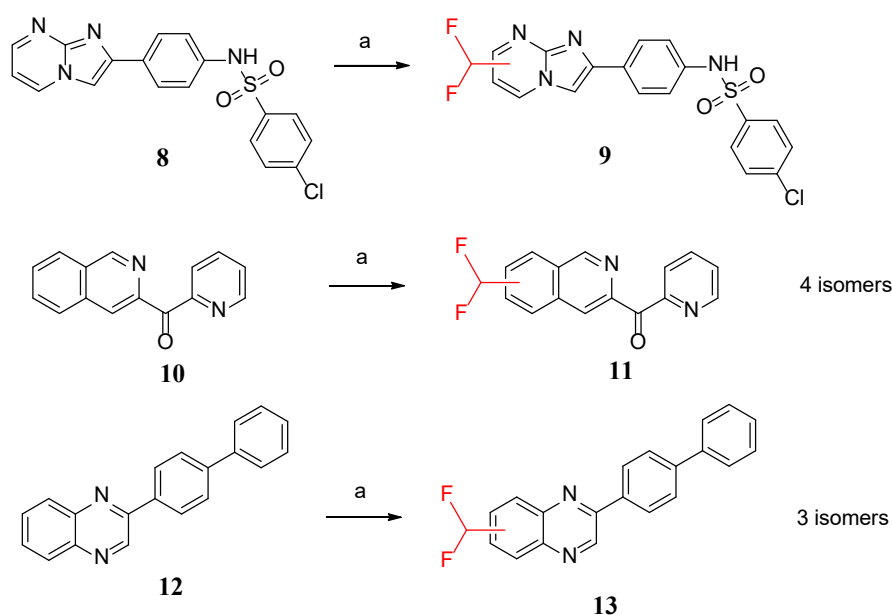
Figure 7. Three *h*AOX non-substrates (**8**, **10**, and **12**) from our database (see Supporting Information) were tested with AOX-Litmus test conditions. Experimental conditions: a) DFMS, DMSO, TFA, TBHP.

Exposure component extimation from AOX flexibility and dynamics studies

At this stage, it appeared evident that the rationalization of experimental results could be supported by structural analyses taking into account also the ligand exposure in the AOX catalytic site. Indeed, molecular docking approaches are often used to clarify significant interactions between the substrates and the residual of the active site of enzymes that metabolize the pharmaceutical drugs. The three-dimensional model of *h*AOX[38] has been used to model new compounds by using the algorithm present in the MetaSite software, [version 6.0] to which appropriate modifications have been made and are reported below.

The AOX enzyme is composed of two identical subunits of 150 kDa and is prevalently expressed in the liver but also in other human tissues.

The structure of the human AOX enzyme has recently been solved by x-ray crystallography (code 4UHX).[38] This protein structure is missing two important polar amino acids (Asp881, Glu882) situated in the loop between the Leu880 and Ser883 residues that are only at 4.5 Å distance from the substrate phthalazine. It is clear that they carry out a significant role in molecular interaction with potential substrates. Furthermore, also residues Phe655, Phe656, Thr657, Glu658, Ala659 and Glu660, located in the lower part of the binding site, are missing. These residues, mainly hydrophobic, belong to the gate 1 region, while the previous two are part of the gate 2. Both gate 1 and gate 2 are dynamic loops involved in regulating the ligands entrance into the pocket.[38] To avoid false results from these structural gaps, and from other loops not resolved in the original x-ray structure, we built a model of the entire protein using the 4UHX structure as template and adding the missing loops (Figure 8). The resulting AOX catalytic cavity includes 30 amino acids, which, notwithstanding their different physical-chemical natures and different polarity, have generated a large pathway of hydrophobic interaction with fewer areas of H-bonding interaction with H-bond interaction donors being almost absent. Calculations made using water-Flap[39,40] have shown the presence of a water molecule in proximity to the molybdenum cofactor. This water is mobile but localized in a well-defined region, where the enzyme-substrate nucleophile attack can take place, and the oxygen of the water molecule is transferred to the AOX substrate.

It is clear that the entire protein and, in particular, the binding site displays a significant degree of flexibility, likely fundamental for the protein to perform its action. To better investigate the protein intrinsic dynamic we performed plain Molecular Dynamics (MD) studies starting from the protein model previously mentioned, which was minimized, equilibrated and submitted to 200 ns long MD simulations (see Experimental Section). The obtained trajectory was clustered according to the variation of the Molecular Interaction Fields (MIFs) of the pocket,[41] as defined by the first mentioned thirty residues. The analyses confirmed a high flexibility of the 655-663 and of the 881-885 segments, supporting their role in regulating the pocket accessibility (Figure 8). This intrinsic dynamic triggers a significant variation of the pocket MIFs and volume, and of the chemical and

energetic properties of the catalytic site, but also a dramatic change in the pocket dimension and accessibility. These features could reasonably be responsible of the enzyme capability of accepting and metabolizing a number of different and variable chemical entities.
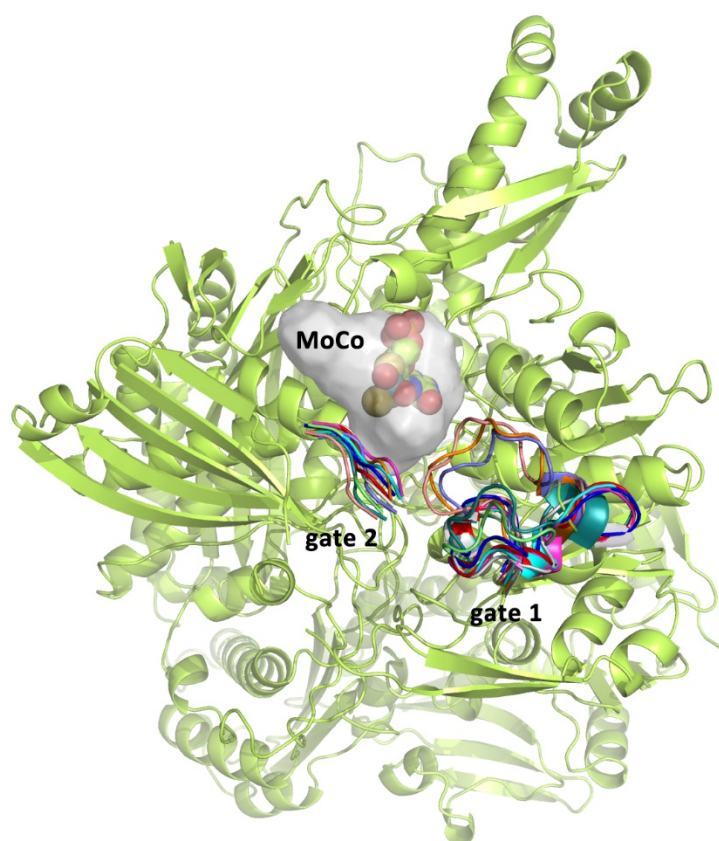

Figure 8. Flexibility of gate 1 and 2. The different conformations of the gates have been taken from the ten medoids obtained by the trajectory clustering. Each color corresponds to a different conformation. The protein binding site cavity was calculated with FLAPsite module in FLAP software (Molecular Discovery, UK)[42-44] and is represented as grey surface. The MoCo cofactor is shown in sphere.

Modelling AOX-catalyzed reactions

We reasoned that, likewise with cytochromes P450 and FMO enzymes, both oxidation and hydrolysis reactions are carried out by means of a combination of substrate reactivity (described here by the electrophilicity) and the substrate interactions with the AOX enzyme.

In previous studies,[3,15,16] we reported that, once the component of accessibility to the enzymatic site and electrophilic reactivity has been calculated, the metabolic site is described by a probability function for the metabolic site (P_{SoM}):

$$P_{SoMi} = (1 + w_e)E_i(1 + w_r)R_i \qquad (1)$$

This function is considered to be an approximation of the free energy of the process including substrate-enzyme interaction,[3,15,16] where exposure (E) and reactivity (R) are opportunely weighted (w_e and w_r, respectively). The two factors that contribute to the P_{SoM} function make contributions that depend on macromolecule utilization (CYP, FMO3, AOX) and on the substrate molecule. In the case of AOX, our analysis of the experimental data collected for more than 600 compounds indicated that the function that describes the substrate exposure plays an important role practically equivalent than that described by the electrophilic reactivity of the substrate. Whatever the case, it is the product of the two components that determines whether a compound is a substrate of the enzyme AOX or not, and in which atom a reaction can occur (if several reactive atoms are present). Figure 9 illustrates an example of reactivity and exposure contributions for experimental results obtained on structurally similar molecules but which have shown totally different experimental behaviors.
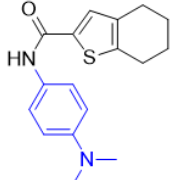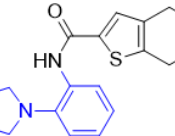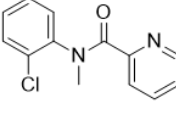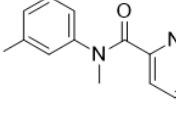
Figure 9. Examples of reactivity and exposure effects. a) Compounds **14** and **15**, having an amide group with similar electrophilicity (due to tertiary amines as anilide substituents) but different shape showed opposite behavior in AOX-mediated hydrolysis reaction; b) compounds **16** and **17**, similar in shape but not in electrophilic nature, were hydrolyzed by AOX with a different extent; c) and d) molecules with the same aza-aromatic electrophilicity (as in **18** and **19**) or molecular shape (**12** and **20**) showed opposite behavior towards oxidation reaction performed by AOX.

In Figure 9-a, the carbonyl carbon of the compounds **14** and **15** has the same electrophilicity; nevertheless, only compound **15** resulted to be an AOX substrate. This observation suggests that in

this case the portion of exposure to the reactive enzyme site will be the dominant part of the equation (1).

Differently, when compounds **16** and **17** are compared (Figure 9b), the electron withdrawing effect of chlorine affecting the carbonyl electrophilicity seems to be responsible for the different AOX liability. Therefore, even small structural changes can make a significant difference in the metabolic biochemical response. Similar considerations can be made for the cases reported in Figure 9c and 9d related to the oxidation reaction.

The weights of the exposure and reactivity effects in equation 1 were accurately computed taking into consideration our dataset of more than 600 molecules with different scaffolds, sizes, and shapes; different substituents have been designed, synthesized, or acquired, and lastly experimentally tested, to significantly enlarge the chemical diversity of AOX substrates.[17]

Once the models of exposure and reactivity had been introduced in MetaSite, this procedure was used to predict the substrate selectivity and the site of metabolism for a series of xenobiotics, for which experimental data were produced *in house* using the LC-MS/MS technique. Consequently, it is worth recalling that MetaSite is not a dependent training-set method as the two functions used in P_{SoM} are obtained entirely independently from the experimental results and only the weights of the functions were optimized to improve the recalculation of the experimental data.

In summary, the metabolism of azaheterocycle compounds is a complex function of several factors that range from the atomic charge at the most positive C-H, presence and effects of substituents (EDG or EWG) in the azaheterocycle ring, and protonation of the azaheterocycle for compounds with basic centers having a pK_a greater than 7.0. Furthermore, hindered and hydrophobic substituents are often able to dramatically change the substrate exposure, thus enhancing the metabolic stability. This is often caused by interfering with the water network inside or in the proximity of the catalytic center. Regarding amides, the same effects are present but they are more relevant in the amine counterpart, which is more sensitive to *h*AOX susceptibility being established.

Validation

The methodology was validated by comparing the experimental results obtained from more than 600 molecules with the predictions obtained from MetaSite.[3] Approximately 200 molecules experimentally susceptible to oxidation by AOX and approximately 45 molecules experimentally susceptible to hydrolysis by the amide group were used together with approximately 200 molecules that, even if having the structural requirements to be considered potential AOX substrates, were experimentally non-AOX substrates. Two items of information were obtained from the *in silico* methodology. Firstly, the character of the substrate or non-substrate for a compound, and secondly the site of metabolism (or the site*s* of metabolism if experimentally more than one). The prediction was considered to be correct if both the character of the substrate and the site (or sites) of metabolism were correct.

Table 1 summarizes the results obtained from the comparison of the experimental tests with those obtained *in silico*. Among the tested compounds, 195 molecules presented the oxidation reaction in one single site while six presented it in at least two sites of metabolism. In addition, 37 molecules presented the hydrolysis reaction in one single site while seven presented both hydrolysis and oxidation. Furthermore, the molecules studied showed great structural diversity including both rigid and very flexible compounds with more than 10 flexible bonds not to mention a wide range of molecular weights and great variability in lipophilicity.

Table 1. Validation results for the proposed model.

AOX metabolism	Oxidation[1]	Hydrolysis[2]
SUBSTRATE	208	44
NOT SUBSTRATE	303	126
UNCERTAIN	23	3
CORRECT SUBSTRATE[3]		
CORRECT NON-SUBSTRATE[3]		
WRONG SUBSTRATE[3]		

WRONG NON-SUBSTRATE[3]		
% GOOD PREDICTION	88	95
ONE SoM	195	37
TWO SoMs	6	0
TWO SOMs OX/HYDRO	7	7

[1] Database of 534 compounds. [2] Database of 173. [3]According to MetaSite predictions.

Sources of experimental variability

A significant variation of hydroxylase activity analyzed in liver cytosol from human donors[45] in relation to the used substrates has been reported.[46] This is probably due to a different quantity of enzyme present in cytosol in addition to the effect of polymorphism by the human AOX1 gene.[47]

In order to limit the variability and standardize the experimental results obtained as much as possible, the experimental study presented in this paper used a method for quantifying the AOX enzyme present in the cytosol (see Experimental Section) that facilitated the quantification of the enzyme in both the various batches of human acquired cytosol and in human hepatocytes and the normalization of the obtained results. Furthermore, all the compounds found to be potential AOX substrates were also tested with AOX inhibitors at different concentrations, confirming the isoform selectivity. In the opinion of the authors the database obtained is extremely homogeneous and at the same time structurally different, and is now available to the scientific community to *in silico* predict the liability of new compounds to the human AOX enzyme.

Chemistry

All synthesized compounds were obtained according to the procedures reported in Schemes 1-3.

According to Scheme 1, diamine **21** was obtained by nucleophilic substitution of ethylendiamine on 2-chloropyrimidine. Finally, target compound **1** was obtained by Pd$_2$(dba)$_2$-mediated coupling between diamine **21** and triflate **22**.

Scheme 1. Synthesis of compound **1**.[a]
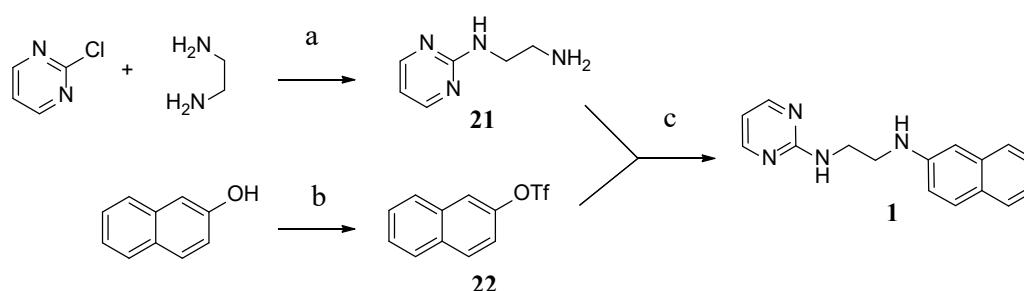
[a]Reagents and conditions: a) reflux, 2 h (62%); b) Triflic anhydride, Pyridine, DCM, 0 °C, 2 h (99%); c) Pd$_2$(dba)$_3$, tBuXPhos, Cs$_2$CO$_3$, dioxane, reflux overnight (7%).

Amine **4** was obtained in quantitative yield by LiAlH$_4$ reduction of amide **3**, in turn prepared by HATU-promoted coupling between 1H-indole-2-carboxylic acid and 2-(pyrrolidin-1-yl)aniline. (Scheme 2).

Scheme 2. Synthesis of amine **3**.[a]
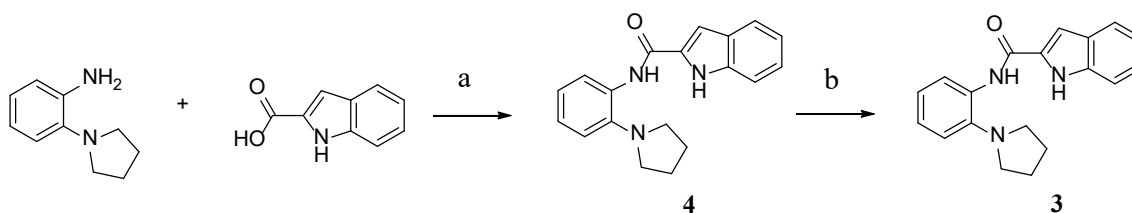
[a]Reagents and conditions: a) HATU, DIPEA, DMF rt (48%); b) LiAlH4, THF, 0 °C (99%).

N-methyl amides **16** and **17** were obtained by methylation of amides **23a-b** (previously synthesized),[17] respectively (Scheme 3).

Scheme 3. Synthesis of amides **16** and **17**.

23a, R = Cl, R' = H
b, R = H, R' = Cl

16, R = Cl, R' = H
17, R = H, R' = Cl
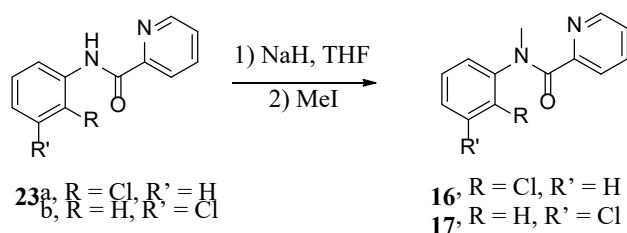
Conclusions

AOX is a cytosolic enzyme able to metabolize pharmaceutical drugs and has currently become one of the main themes of interest in Medicinal Chemistry, Drug Metabolism and Pharmacokinetics. Numerous original research papers focusing on the role of AOX in the metabolism of xenobiotics have been, in fact, published in the last years. This high level of interest has been caused by the numerous unfortunate clinical outcomes, mainly given to unacceptable pharmacokinetic properties and drug safety problems, and often related to an increased liability to AOX. Interestingly, the highest sensitivity to AOX seems to be related to the modifications drug candidates generally undergo to escape CYP metabolism. In this situation, a practical and fast *in silico* strategy to predict human AOX substrate selectivity and to identify the site of metabolism, for both the well-known oxidation reactions and for the less known hydrolysis, is urgently needed. The numerosity and diversity of the used dataset, together with the homogeneity of the produced experimental information should guarantee the reliability of the presented methodology (included in the MetaSite procedure) for all typical drug discovery and new chemical entities optimization projects. Applied in an early phase, the method should facilitate a good understanding of the passivity of AOX-catalyzed metabolism and suggest potential modifications to increase the stability of a xenobiote to AOX-catalyzed reactions.

EXPERIMENTAL SECTION

Chemistry

All commercial products were acquired from Sigma Aldrich and used without further purification (≥ 97% pure). Anhydrous solvents were acquired from Acros Organics. Compounds **23a-b** were obtained as previously reported,[17] from HATU-mediated coupling between picolinic acid and 2- or 3-chloroaniline, respectively. ^{1}H and ^{13}C NMR nuclear magnetic resonance (NMR) spectra were recorded at 400 and 100.6 MHz, respectively, on Bruker Avance III HD 400 MHz spectrometer at room temperature. Chemical shifts (δ) are given in parts per million (ppm) relative to the internal standard tetramethylsilane. Peak multiplicities are reported as s (singlet), d (doublet), dd (double doublet), t (triplet), dt (double triplet), q (quartet), m (multiplet), or br s (broad singlet). Coupling constants (*J*) are given in Hz. HRMS spectra were registered on Agilent Technologies 6540 UHD Accurate Mass Q-TOF LC/MS system. Purity for all final compound was determined by LC-MS/MS and resulted >98%.

General Procedure A for the synthesis of amides.

HATU (1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxid hexafluorophosphate, 1.25 mmol) was added at rt to a stirred solution of suitable carboxylic acid (1.10 mmol) and DIPEA (*N,N*-diisopropylethylamine, 2.25 mmol) in anhydrous *N,N*-dimethylformamide (1.00 mL). After 6 min, the amine was added (1.00 mmol) and stirring was continued overnight. The mixture was dropped in water (10 mL), the resulting solid filtered off, washed with water and dried. The solid was purified by column chromatography to give the desired amide.

***N*^{1}-(naphthalen-2-yl)-*N*^{2}-(pyrimidin-2-yl)ethane-1,2-diamine (1).** Pd$_2$(dba)$_3$ (18.0 mg, 0.02 mmol) was added to a degased solution of amine **21** (0.150 g, 1.1 mmol), trifluoromethanesulfonate **22** (0.276 g, 1.0 mmol), Cs$_2$CO$_3$ (0.487 g, 1.5 mmol), and *t*BuXPhos (13.0 mg, 0.03 mmol) in

anhydrous dioxane (10 mL). The reaction was stirred under reflux for 20 h, cooled to rt, diluted with dichloromethane and filtered over a celite pad. The solution was concentrated and the residue purified by flash chromatography on SiO_2 (eluent: dichloromethane, followed by dichloromethane/methanol, 9:1). The resulting oil was further purified by flash chromatography (eluent petroleum ether/ethyl acetate, from 6:4 to 4:6). The titled product was obtained as a yellow solid (19.1 mg, 7% yield). ^{1}H NMR (400 MHz, CDCl$_3$) δ 8.25 (d, J = 4.8 Hz, 2H), 7.64 – 7.50 (m, 3H), 7.34 – 7.24 (m, 1H), 7.14 (t, J = 7.5 Hz, 1H), 6.82 (dd, J = 2.3, 11.0 Hz, 2H), 6.57 – 6.47 (m, 1H), 5.85 (s, 1H), 3.71 (q, J = 6.0 Hz, 2H), 3.60 (bs, 1H), 3.45 (t, J = 5.9 Hz, 2H); ^{13}C NMR (101 MHz, CDCl$_3$) δ 162.6, 158.0, 145.8, 135.2, 128.8, 127.5, 127.4, 126.2, 125.8, 121.8, 118.0, 110.8, 104.1, 43.9, 40.6, 23.9; HRMS: calcd. for $C_{16}H_{16}N_4$ 265.1148 (M+H[+]), found 265.1148 (M+H[+]).

N-((1H-indol-2-yl)methyl)-2-(pyrrolidin-1-yl)aniline (3). LiAlH$_4$ was added under argon at 0 °C to a stirred solution of amide **4** in anhydrous tetrahydrofuran. After stirring at rt overnight, the reaction was carefully quenched under nitrogen with ethyl acetate, followed by water. The aqueous phase was extracted with ethyl acetate (x3), and reunited organic phases were evaporated to dryness. Flash column chromatography on SiO_2 (eluent dichloromethane /methanol 95:5) allowed to recover the titled product as white solid, which was characterized as follows. ^{1}H NMR (400 MHz, CDCl$_3$) δ 8.36 (s, 1H), 7.63 (d, J = 7.7 Hz, 1H), 7.35 (d, J = 8.1 Hz, 1H), 7.26 – 7.09 (m, 3H), 7.00 (td, J = 1.2, 7.6 Hz, 1H), 6.87 – 6.67 (m, 2H), 6.57 – 6.45 (m, 1H), 5.03 (t, J = 5.3 Hz, 1H), 4.56 (d, J = 5.5 Hz, 2H), 3.17 – 3.03 (m, 4H), 2.13 – 1.89 (m, 4H); ^{13}C NMR (101 MHz, CDCl$_3$) δ 143.5, 137.7, 137.6, 135.9, 128.8, 124.3, 121.5, 120.2, 119.8, 118.8, 118.0, 110.9, 110.8, 99.6, 51.5 (2C), 42.7, 24.2 (2C); HRMS: calcd. for $C_{19}H_{21}N_3$ 292.1808 (M+H[+]), found 292.1802 (M+H[+]).

N-(2-(Pyrrolidin-1-yl)phenyl)-1H-indole-2-carboxamide (4). The titled compound was obtained according to the General Procedure A from 1H-indole-2-carboxylic acid and 2-(pyrrolidin-1-yl)aniline. After column chromatography (eluent dichloromethane), the product was obtained as a turquoise solid (48% yield). ^{1}H NMR (400 MHz, CDCl$_3$) δ 9.37 (s, 1H), 9.11 (s, 1H), 8.38 (d, J = 7.5 Hz, 1H), 7.74 (d, J = 8.0 Hz, 1H), 7.50 (d, J = 8.3 Hz, 1H), 7.35 (t, J = 7.6 Hz, 1H), 7.26 – 7.08

(m, 4H), 6.94 (s, 1H), 3.13 (s, 4H), 2.07 (s, 4H); ^{13}C NMR (101 MHz, CDCl$_3$) δ 162.6, 142.9, 139.8, 138.5, 131.3, 125.6, 125.2, 121.7, 120.8, 120.7, 119.8, 118.4, 114.9, 111.1, 110.7, 54.4 (2C), 25.5 (2C); HRMS: calcd. for C$_{19}$H$_{19}$N$_3$O 306.1601 (M+H$^+$), found 306.1608 (M+H$^+$).

N-(4-(dimethylamino)phenyl)-4,5,6,7-tetrahydrobenzo[*b*]thiophene-2-carboxamide (14). The titled compound was obtained according to the General Procedure A from 4,5,6,7-tetrahydrobenzo[*b*]thiophene-2-carboxylic acid and N^1,N^1-dimethylbenzene-1,4-diamine. After column chromatography (eluent dichloromethane /ethyl acetate, 9:1), the product was obtained as a yellow solid (27% yield). ^{1}H NMR (400 MHz, Chloroform-*d*) δ 7.60 – 7.40 (m, 3H), 7.29 (d, *J* = 3.8 Hz, 1H), 6.75 (d, *J* = 8.9 Hz, 2H), 2.95 (s, 6H), 2.82 (t, *J* = 5.8 Hz, 2H), 2.64 (t, *J* = 5.8 Hz, 2H), 2.05 – 1.68 (m, 4H); ^{13}C NMR (101 MHz, Chloroform-*d*) δ 160.09, 148.02, 141.51, 136.31, 134.86, 129.24, 127.68, 121.95 (2C), 113.12 (2C), 40.96 (2C), 25.47, 25.34, 23.28, 22.62; HRMS: calcd. for C$_{17}$H$_{20}$N$_2$Os 301.1369 (M+H$^+$), found 301.1379 (M+H$^+$).

N-(2-(Pyrrolidin-1-yl)phenyl)-4,5,6,7-tetrahydrobenzo[*b*]thiophene-2-carboxamide (15). The titled compound was obtained according to the General Procedure A from 4,5,6,7-tetrahydrobenzo[*b*]thiophene-2-carboxylic acid and 2-(pyrrolidin-1-yl)aniline. After column chromatography (eluent dichloromethane), the product was obtained as a yellow solid (48% yield). ^{1}H NMR (400 MHz, Chloroform-*d*) δ 8.84 (s, 1H), 8.35 (d, *J* = 7.8 Hz, 1H), 7.32 (s, 1H), 7.18 (d, *J* = 7.7 Hz, 1H), 7.24 – 6.87 (m, 2H), 3.08 (t, *J* = 5.9 Hz, 4H), 2.83 (t, *J* = 5.8 Hz, 2H), 2.67 (t, *J* = 5.8 Hz, 2H), 2.02 (p, *J* = 3.1 Hz, 4H), 1.98 – 1.70 (m, 4H); ^{13}C NMR (101 MHz, Chloroform-*d*) δ 159.84, 141.69, 140.10, 136.56, 135.16, 133.14, 129.68, 124.24, 123.91, 120.29, 119.35, 52.48 (2C), 25.53, 25.40, 24.45 (2C), 23.28, 22.62; HRMS: calcd. for C$_{19}$H$_{22}$N$_2$OS 327.1526 (M+H$^+$), found 327.1522 (M+H$^+$).

N-(2-chlorophenyl)-N-methylpicolinamide (16). To a solution of **23a** in DMF (1 mL) was added sodium hydride (26 mg, 60% in mineral oil, 0.64 mmol), and the mixture was stirred at room temperature for 30 min. Methyl iodide (55 μL, 0.64 mmol) was then added at 0 °C, and stirring was continued at rt for 24 h. The reaction was quenched with water, and extracted 3 times with ethyl

acetate. The reunited organic phases were washed with water and brine, and dried over Na_2SO_4. Filtration, concentration of the filtrate under vacuum, and purification of the residue by silica gel flash column chromatography (ethyl acetate/petroleum ether, 4:6) gave the titled product as a yellow oil (56.0 mg, 54% yield), which was characterized as follows. Major atropoisomer (89%, in ^{13}C NMR spectra indicated with carbons C). ^{1}H NMR (400 MHz, CDCl$_3$) δ 8.25 (d, J = 4.7 Hz, 1H), 7.68 – 7.54 (m, 2H), 7.32 – 7.24 (m, 1H), 7.24 – 7.18 (m, 1H), 7.16 – 7.10 (m, 2H), 7.10 – 7.07 (m, 1H), 3.52 – 3.32 (m, 3H). Minor atropoisomer (11 %, in ^{13}C NMR spectra indicated with carbons C'). ^{1}H NMR (400 MHz, CDCl$_3$) δ 8.68 (d, J = 3.9 Hz, 1H), 7.84 (t, J = 7.3 Hz, 1H), 7.78 (d, J = 7.5 Hz, 1H), 7.52 (d, J = 7.8 Hz, 1H), 7.47 – 7.32 (m, 4H), 3.41 (s, 3H). ^{13}C NMR (101 MHz, CDCl$_3$) δ 168.7 (C), 153.6 (C), 148.3 (C'), 148.0 (C), 141.8 (C), 137.1 (C'), 136.1 (C), 132.2 (C), 130.5 (C'), 130.4 (C), 130.0 (C), 129.3 (C'), 129.1 (C'), 128.7 (C), 128.1 (C'), 127.4 (C), 124.8 (C'), 124.1 (C), 124.1 (C'), 123.2 (C), 39.8 (C'), 36.9 (C). Three carbons of the minor atropoisomer are not detected. HRMS: calcd. for $C_{13}H_{11}{}^{35}ClN_2O$ 247. 0638 (M+H$^+$), found 247.0639 (M+H$^+$).

N-(3-chlorophenyl)-N-methylpicolinamide (17). The titled product was obtained according to the procedure described for **16**, but starting from **23b**. ^{1}H NMR (400 MHz, CDCl$_3$) δ 8.38 (s, 1H), 7.69 (t, J = 7.4 Hz, 1H), 7.64 – 7.49 (m, 1H), 7.25 – 7.18 (m, 1H), 7.14 (s, 3H), 6.96 (s, 1H), 3.53 (s, 3H); ^{13}C NMR (101 MHz, CDCl$_3$) δ 168.6, 153.7, 148.4, 145.6, 136.5, 134.4, 129.8, 126.7, 126.7, 124.9, 124.3, 123.9, 38.2. HRMS: calcd. for $C_{13}H_{11}{}^{35}ClN_2O$ 247. 0638 (M+H$^+$), found 247.0637 (M+H$^+$).

N^1-(pyrimidin-2-yl)ethane-1,2-diamine (21). A mixture of 2-chloropyrimidine (1.02 g, 8.9 mmol) and ethylendiamine (10 mL, 667 mmol) was refluxed for 2 h. The mixture was diluted with brine and extracted with dichloromethane. The reunited organic phases were evaporated to dryness to afford the titled product as a yellow oil (0.75 g, yield 62%), which was used without further purification in the following step. ^{1}H NMR (400 MHz, CDCl$_3$) δ 8.28 (d, J = 4.8 Hz, 2H), 6.53 (t, J = 4.8 Hz, 1H), 5.74 – 4.93 (m, 2H), 3.50 (q, J = 6.0 Hz, 2H), 3.06 – 2.78 (m, 2H).

Naphthalen-2-yl trifluoromethanesulfonate (22). Trifluoromethanesulfonic anhydride (4.00 mL, 24.0 mmol) was added drop-wise at 0 °C to a stirred solution of naphthalen-2-ol (2.88 g, 20.0 mmol) and pyridine (3.23 mL, 40.0 mmol) in anhydrous dichloromethane (40 mL). The mixture was warmed at rt and left under stirring for additional 2 h. HCl 10% aq. was added and the mixture extracted by dichloromethane. The reunited organic phases were washed with $NaHCO_3$ sat., dried over Na_2SO_4, and concentrated under vacuum. After flash chromatography on SiO_2 (eluent, petroleum ether) a yellow solid was obtained (5.48g, 99% yield). ^1H NMR (400 MHz, $CDCl_3$) δ 8.11 – 7.86 (m, 3H), 7.78 (d, $J = 2.5$ Hz, 1H), 7.70 – 7.55 (m, 2H), 7.41 (dd, $J = 2.5, 9.0$ Hz, 1H).

Metabolism Assay on HLC.

Solvents, reagents, inhibitors and pooled mixed-gender cryopreserved HLC (1 mg/mL) were obtained from Sigma-Aldrich. HLM UltraPool™ HLM 150 were purchased from BD Gentest (Woburm, MA, USA). Tested compounds for in vitro assays were acquired from SPECS, ENAMINE, VITAS-M and MEDCHEMEXPRESS. All tested compounds were > 98% pure as detected by UHPL-LC-MS/MS analysis through 6540 UHD accurate-mass quadrupole time-of-flight (QTOF) system. For AOX quantification, TRIS/HCl, urea, dithiothreitol (DTT), trypsin (sequencing grade), iodoacetamide (IAA), HLC (Sigma Aldrich), H2O (MS grade), formic acid (MS grade) were acquired from Sigma Aldrich and used without further purification. Standard peptide YIQDIVASTLK was obtained provided by the Institute of Biostructure and Bioimaging (IBB) of the Italian National Research Council, (Naples, Italy).

General Procedure for metabolism assay

Metabolism was evaluated upon incubation with HLC according to a modified Dalvie et al.[30] procedure, recently described.[17] Briefly, each compound (10 μM) was incubated at pH 7.4 and 37 °C in the presence of magnesium chloride hexahydrate (MgCl2·6H2O, 1 mM) and HLC. The reaction was stopped after 0 and 60 min by addition of 250 μL of ice-cold acetonitrile (containing

0.6 μM labetalol as internal standard). Corresponding blank was prepared following an identical procedure, in the absence of tested compound. Similarly, in the inhibition study, test compound (10 μM) was incubated in the presence (50 μM) or absence of selective inhibitor DCPIP. Proteins were precipitated by centrifugation at 5,000 rpm for 10 min (Eppendorf, Italy; centrifuge 5810 R; rotor F-45-30-11) at room temperature, and an aliquot of supernatant (1 μL) was analyzed by LC-MS/MS on an Agilent 1200 series HPLC coupled to an Agilent 6540 UHD accurate-mass QTOF with a dual Jet Stream electron spray ionization source. Depending on the nature of the substrates three different analytical methods (1, 2, 3) were used in LC analysis. The mobile phase was a mixture of water (solvent A) and acetonitrile (solvent B), both containing formic acid at 0.1%. Method 1: Aeris Widepore 3.6-μm (C4, 100 × 4.6-mm) column at 30 °C using a flow rate of 0.850 mL/min in a 10-min gradient elution. Gradient elution: 100:0 (A/B) to 70:30 (A/B) over 9 min, 5:95 (A/B) for 1 min, and then reversion back to 100:0 (A/B) over 0.1 min. Method 2: Acquity UPLC BEH C18 1.7-μm (C18,150 ×2.1-mm) column at 40°C using a flow rate of 0.650mL/minina10-min gradient elution. Gradient elution: 99.5:0.5 (A/B) to 5:95 (A/B) over 8 min, 5:95 (A/B) for 2 min, and then reversion back to 99.5:0.5 (A/B) over 0.1 min. Method 3: Luna® Omega 1.6 um Polar (C18, 100 x 2.1 mm) column at 40 ºC using a flow rate of 0.650 mL/min in a 10 minutes gradient elution. Gradient elution: 99.5:0.5 (A/B) to 50.0:50.0 (A/B) over 8 minutes, 50.0:50.0 (A/B) for 2 minutes, and then reversion back to 99.5:0.5 (A/B) over 0.1 min. The MS/MS data were processed using Mass-MetaSite (version 3.3.2) (MolecularDiscovery, Ltd., UK).[6] LC-MS chromatograms in figure S3-1 were extracted by MassHunter (Mass Hunter Workstation Software B.06.00 Qualitative Analysis; Agilent).

AOX quantification

HLC was digested following Barr et al. method [48] with modifications. Briefly, 20 mg/ml of HLC (25 μL) were mixed with 8 M urea, 50 mM DTT and incubated for 1 h at 56 °C. 200 mM iodoacetamide was added to the mixture and incubated in the dark at 37 °C for 30 min.

Subsequently the mixture was diluted with 0.1 M TRIS/HCl buffer containing sequence grade trypsin (0.25 µg/µL) and incubated overnight at 37 °C. Digestion was terminated by adding a solution of 50% v/v H_2O/HCO_2H (final pH <2). Aliquots of supernatant (5 µL) were analysed by LC-MS/MS. Quantification of digested proteins was performed by MassHunter Quantitative Analysis B.06.00 (Agilent Technologies, CA) using the standard peptide YIQDIVASTLK provided by IBB (Italian National Research Council, Naples, Italy).

Modelling and Molecular Dynamics

The missing loops in the original x-ray structure were built with the homology modelling program MODELLER (http://salilab.org/modeller/). The x-ray human structure of AOX (pdb code 4UHX) was used as template upon removal of the co-crystallized ligands and of the malonate. The standard protonation state at physiological pH was assigned to ionizable residues.

The parameters for the FeS group and for the MoCo cofactor were kindly provided by Cerqueira et al.[49] The FAD cofactor was parameterized within the software BiKi[50] used to setup the MD simulation.

The protein was parametrized by Amber99SB-ILDN force Field[51] and Gromacs 5.1.4 was used to run MD Simulations.[52] The water model employed was TIP3P. The solvated system was preliminary minimized by 5000 steps of steepest descent. The Verlet cutoff scheme, the Bussi−Parrinello thermostat LINCS for the constraints (all bonds), and the particle mesh Ewald for electrostatics, with a short-range cutoff of 11 Å, were applied. The system was equilibrated in seven subsequent steps: 500 ps in NVT ensemble at 50 K, 500 ps in NVT ensemble at 100 K, 500 ps in NVT ensemble at 150 K, 500 ps in NVT ensemble at 200 K, 500 ps in NVT ensemble at 250 K, 500 ps in NVT ensemble at 300 K and, finally, a 1 ns long NPT simulation to reach the pressure equilibrium condition. In the first two equilibration steps harmonic positional restraints were set on the backbone of the protein with a spring constant of 1000 kJ/(mol· Å2). The integration step was set equal to 1 for the equilibration process and inceased to 2 fs during the following production. The

production run was carried out in the NVT ensemble at 300 K without any restraint for 200 ns. The trajectory was clustered according to the variation of the pocket MIFs.[41] Ten medoids were identified and analyzed. The pockets were calculated with the FLAPsite tool implemented FLAP[53] and the MIFs calculated and compared with the new "medoids" tool also implemented in FLAP.

AUTHOR INFORMATION

Corresponding Author

* To whom correspondence should be addressed. phone, + 39 075 585 5629; e-mail, gabriele.cruciani@unipg.it.

Author Contributions

The manuscript was written by G.C., L.G. and F.S.

NOTES

Disclosure: MetaSite 6.0 is freely available for nonprofit research institutions at www.moldiscovery.com.

ABBREVIATIONS USED

AOX, aldehyde oxidase; CYP, cytochrome P450; DFT, density functional theory; FMO3, flavine monoxygenase 3; *h*AOX, human aldehyde oxidase; HLC, human liver cytosol; HLM, human liver microsomes; MD, molecular dynamics; MetID, metabolism identification; MIFs, molecular interaction fields; P_{SoM}, probability for the metabolic site; SoM, site of metabolism.

ASSOCIATED CONTENT

The Supporting Information is available free of charge via the Internet at http://pubs.acs.org. Parr's electrophilicity index calculated for a subset of compounds, DFT calculation details, AOX-Litmus test results (PDF). Molecular formula strings and the associated AOX effect towards oxidation and hydrolysis reactions (XLSX). Two files containing our databases for oxidation and hydrolysis reactions are also available (TXT). Model for AOX (PDB).

REFERENCES

(1)   Kola, I.; Landis, J. Can the pharmaceutical industry reduce attrition rates? *Nat. Rev. Drug Discov.* **2004**, *3*, 711-715.

(2)   Broccatelli, F.; Brown, N. *Molecular Interaction Fields for Predicting the Sites and Products of Metabolism, in Drug Metabolism Prediction* Wiley-VCH Verlag GmbH & Co. KGaA: Weinheim, Germany, 2014.

(3)   Cruciani, G.; Carosati, E.; De Boeck, B.; Ethirajulu, K.; Mackie, C.; Howe, T.; Vianello, R. MetaSite: understanding metabolism in human cytochromes from the perspective of the chemist. *J. Med. Chem.* **2005**, *48*, 6970-6979.

(4)   de Groot, M. J.; Ackland, M. J.; Horne, V. A.; Alex, A. A.; Jones, B. C. A novel approach to predicting P450 mediated drug metabolism. CYP2D6 catalyzed N-dealkylation reactions and qualitative metabolite predictions using a combined protein and pharmacophore model for CYP2D6. *J. Med. Chem.* **1999**, *42*, 4062-4070.

(5)   Li, J.; Schneebeli, S. T.; Bylund, J.; Farid, R.; Friesner, R. A. IDSite: An accurate approach to predict P450-mediated drug metabolism. *J. Chem. Theory Comput.* **2011**, *7*, 3829-3845.

(6)   Zamora, I.; Fontaine, F.; Serra, B.; Plasencia, G. High-throughput, computer assisted, specific MetID. A revolution for drug discovery. *Drug Discov .Today Technol.* **2013**, *10*, e199-e205.

(7)   Di, L. The role of drug metabolizing enzymes in clearance. *Expert Opin. Drug Metab. Toxicol.* **2014**, *10*, 379-393.

(8)   scifinder. **accessed 22/09/2017**.

(9)   Hutzler, J. M.; Obach, R. S.; Dalvie, D.; Zientek, M. A. Strategies for a comprehensive understanding of metabolism by aldehyde oxidase. *Expert Opin. Drug Metab. Toxicol.* **2013**, *9*, 153-168.

(10)   Akabane, T.; Tanaka, K.; Irie, M.; Terashita, S.; Teramura, T. Case report of extensive metabolism by aldehyde oxidase in humans: pharmacokinetics and metabolite profile of FK3453 in rats, dogs, and humans. *Xenobiotica* **2011**, *41*, 372-384.

(11)   Dittrich, C.; Greim, G.; Borner, M.; Weigang-Kohler, K.; Huisman, H.; Amelsberg, A.; Ehret, A.; Wanders, J.; Hanauske, A.; Fumoleau, P. Phase I and pharmacokinetic study of BIBX 1382 BS, an epidermal growth factor receptor (EGFR) inhibitor, given in a continuous daily oral administration. *Eur. J. Cancer* **2002**, *38*, 1072-1080.

(12)   Kaye, B.; Offerman, J. L.; Reid, J. L.; Elliott, H. L.; Hillis, W. S. A species difference in the presystemic metabolism of carbazeran in dog and man. *Xenobiotica* **1984**, *14*, 935-945.

(13)   Zhang, X.; Liu, H. H.; Weller, P.; Zheng, M.; Tao, W.; Wang, J.; Liao, G.; Monshouwer, M.; Peltz, G. In silico and in vitro pharmacogenetics: aldehyde oxidase rapidly metabolizes a p38 kinase inhibitor. *Pharmacogenomics J.* **2011**, *11*, 15-24.

(14)   Diamond, S.; Boer, J.; Maduskuie, T. P., Jr.; Falahatpisheh, N.; Li, Y.; Yeleswaram, S. Species-specific metabolism of SGX523 by aldehyde oxidase and the toxicological implications. *Drug Metab. Dispos.* **2010**, *38*, 1277-1285.

(15)   Cruciani, G.; Baroni, M.; Benedetti, P.; Goracci, L.; Fortuna, C. G. Exposition and reactivity optimization to predict sites of metabolism in chemicals. *Drug Discov. Today Technol.* **2013**, *10*, e155-e165.

(16)   Cruciani, G.; Valeri, A.; Goracci, L.; Pellegrino, R. M.; Buonerba, F.; Baroni, M. Flavin monooxygenase metabolism: why medicinal chemists should matter. *J. Med. Chem.* **2014**, *57*, 6183-6196.

(17)   Lepri, S.; Ceccarelli, M.; Milani, N.; Tortorella, S.; Cucco, A.; Valeri, A.; Goracci, L.; Brink, A.; Cruciani, G. Structure-metabolism relationships in human-AOX: Chemical insights from a large database of aza-aromatic and amide compounds. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114*, E3178-E3187.

(18)   Sodhi, J. K.; Wong, S.; Kirkpatrick, D. S.; Liu, L.; Khojasteh, S. C.; Hop, C. E.; Barr, J. T.; Jones, J. P.; Halladay, J. S. A novel reaction mediated by human aldehyde oxidase: amide hydrolysis of GDC-0834. *Drug Metab. Dispos.* **2015**, *43*, 908-915.

(19)   Lewis, D. F.; Dickins, M. Substrate SARs in human P450s. *Drug Discov. Today* **2002**, *7*, 918-925.

(20)   Lewis, D. F.; Jacobs, M. N.; Dickins, M. Compound lipophilicity for substrate binding to human P450s in drug metabolism. *Drug Discov. Today* **2004**, *9*, 530-537.

(21)   Smith, D. A. Discovery and ADMET: Where are we now. *Curr. Top. Med. Chem.* **2011**, *11*, 467-481.

(22)   van de Waterbeemd, H.; Smith, D. A.; Jones, B. C. Lipophilicity in PK design: methyl, ethyl, futile. *J. Comput. Aided Mol. Des.* **2001**, *15*, 273-286.

(23)   Linton, A.; Kang, P.; Ornelas, M.; Kephart, S.; Hu, Q.; Pairish, M.; Jiang, Y.; Guo, C. Systematic structure modifications of imidazo[1,2-a]pyrimidine to reduce metabolism mediated by aldehyde oxidase (AO). *J Med Chem* **2011**, *54*, 7705-12.

(24)   Pryde, D. C.; Dalvie, D.; Hu, Q.; Jones, P.; Obach, R. S.; Tran, T. D. Aldehyde oxidase: an enzyme of emerging importance in drug discovery. *J. Med. Chem.* **2010**, *53*, 8441-8460.

(25)   Pryde, D. C.; Tran, T. D.; Jones, P.; Duckworth, J.; Howard, M.; Gardner, I.; Hyland, R.; Webster, R.; Wenham, T.; Bagal, S.; Omoto, K.; Schneider, R. P.; Lin, J. Medicinal chemistry approaches to avoid aldehyde oxidase metabolism. *Bioorg. Med. Chem. Lett.* **2012**, *22*, 2856-2860.

(26)   Cruciani, G.; Crivori, P.; Carrupt, P.; Testa, B. Molecular fields in quantitative structure-permeation relationships: the VolSurf approach. *Theochem.* **2000**, *503*, 17-30.

(27)   Cruciani, G.; Pastor, M.; Guba, W. VolSurf: a new tool for the pharmacokinetic optimization of lead compounds. *Eur. J. Pharm. Sci.* **2000**, *11 Suppl 2*, S29-S39.

(28)   Chan, K.; Poon, R.; O'Brien, P. J. Application of structure-activity relationships to investigate the molecular mechanisms of hepatocyte toxicity and electrophilic reactivity of alpha,beta-unsaturated aldehydes. *J. Appl. Toxicol.* **2008**, *28*, 1027-1039.

(29)   Ghafourian, T.; Rashidi, M. R. Quantitative study of the structural requirements of phthalazine/quinazoline derivatives for interaction with human liver aldehyde oxidase. *Chem. Pharm. Bull. (Tokyo)* **2001**, *49*, 1066-1071.

(30)   Dalvie, D.; Sun, H.; Xiang, C.; Hu, Q.; Jiang, Y.; Kang, P. Effect of structural variation on aldehyde oxidase-catalyzed oxidation of zoniporide. *Drug Metab. Dispos.* **2012**, *40*, 1575-1587.

(31)   Torres, R. A.; Korzekwa, K. R.; McMasters, D. R.; Fandozzi, C. M.; Jones, J. P. Use of density functional calculations to predict the regioselectivity of drugs and molecules metabolized by aldehyde oxidase. *J. Med. Chem.* **2007**, *50*, 4642-4647.

(32)   Kiyooka, S.; Kaneno, D.; Fujiyama, R. Parr's index to describe both electrophilicity and nucleophilicity. *Tetrahedron Lett.* **2013**, *54*, 339-342.

(33)   Parr, R.; Szentpaly, L.; Liu, S. Electrophilicity index. *J. Am. Chem. Soc.* **1999**, *121*, 1922-1924.

(34)   Alfaro, J. F.; Jones, J. P. Studies on the mechanism of aldehyde oxidase and xanthine oxidase. *J. Org. Chem.* **2008**, *73*, 9469-9472.

(35)   Konishi, K.; Fukami, T.; Gotoh, S.; Nakajima, M. Identification of enzymes responsible for nitrazepam metabolism and toxicity in human. *Biochem. Pharmacol.* **2017**, *140*, 150-160.

(36)   Paragas, E.; Humphreys, S.; Min, J.; Joswig-Jones, C.; Jones, J. The two faces of aldehyde oxidase: Oxidative and reductive transformations of 5-nitroquinoline. *Biochem. Pharmacol.* **2017**, *145*, 210-217.

(37)   O'Hara, F.; Burns, A. C.; Collins, M. R.; Dalvie, D.; Ornelas, M. A.; Vaz, A. D.; Fujiwara, Y.; Baran, P. S. A simple litmus test for aldehyde oxidase metabolism of heteroarenes. *J. Med. Chem.* **2014**, *57*, 1616-1620.

(38)   Coelho, C.; Foti, A.; Hartmann, T.; Santos-Silva, T.; Leimkuhler, S.; Romao, M. J. Structural insights into xenobiotic and inhibitor binding to human aldehyde oxidase. *Nat. Chem. Biol.* **2015**, *11*, 779-783.

(39)   *FLAP/WaterFLAP*, Molecular Discovery, Ltd, UK, 2013.

(40)   Mason, J.; Bortolato, A.; Weiss, D.; Deflorian, F.; Tehan, B.; Marshall, F. High end GPCR design: crafted ligand design and druggability analysis using protein structure, lipophilic hotspots and explicit water networks. *In Silico Pharmacol.* **2013**, *1*, 1-23.

(41)   Spyrakis, F.; Benedetti, P.; Decherchi, S.; Rocchia, W.; Cavalli, A.; Alcaro, S.; Ortuso, F.; Baroni, M.; Cruciani, G. A pipeline to enhance ligand virtual screening: integrating molecular dynamics and fingerprints for ligand and proteins. *J. Chem. Inf. Model.* **2015**, *55*, 2256-2274.

(42)   Henrich, S.; Salo-Ahen, O. M.; Huang, B.; Rippmann, F. F.; Cruciani, G.; Wade, R. C. Computational approaches to identifying and characterizing protein binding sites for ligand design. *J. Mol. Recognit.* **2010**, *23*, 209-219.

(43)   Mangiatordi, G. F.; Alberga, D.; Siragusa, L.; Goracci, L.; Lattanzi, G.; Nicolotti, O. Challenging AQP4 druggability for NMO-IgG antibody binding using molecular dynamics and molecular interaction fields. *Biochim Biophys. Acta* **2015**, *1848*, 1462-1471.

(44)   Spyrakis, F.; Cellini, B.; Bruno, S.; Benedetti, P.; Carosati, E.; Cruciani, G.; Micheli, F.; Felici, A.; Cozzini, P.; Kellogg, G. E.; Voltattorni, C. B.; Mozzarelli, A. Targeting cystalysin, a virulence factor of treponema denticola-supported periodontitis. *ChemMedChem* **2014**, *9*, 1501-1511.

(45)   Schofield, P. C.; Robertson, I. G.; Paxton, J. W. Inter-species variation in the metabolism and inhibition of N-[(2'-dimethylamino)ethyl]acridine-4-carboxamide (DACA) by aldehyde oxidase. *Biochem. Pharmacol.* **2000**, *59*, 161-165.

(46)   Hutzler, J. M.; Yang, Y. S.; Albaugh, D.; Fullenwider, C. L.; Schmenk, J.; Fisher, M. B. Characterization of aldehyde oxidase enzyme activity in cryopreserved human hepatocytes. *Drug Metab. Dispos.* **2012**, *40*, 267-275.

(47)   Garattini, E.; Terao, M. The role of aldehyde oxidase in drug metabolism. *Expert Opin. Drug Metab. Toxicol.* **2012**, *8*, 487-503.

(48)   Barr, J. T.; Jones, J. P.; Joswig-Jones, C. A.; Rock, D. A. Absolute quantification of aldehyde oxidase protein in human liver using liquid chromatography-tandem mass spectrometry. *Mol. Pharm.* **2013**, *10*, 3842-3849.

(49)   Cerqueira, N. M.; Coelho, C.; Bras, N. F.; Fernandes, P. A.; Garattini, E.; Terao, M.; Romao, M. J.; Ramos, M. J. Insights into the structural determinants of substrate specificity and activity in mouse aldehyde oxidases. *J. Biol. Inorg. Chem.* **2015**, *20*, 209-217.

(50)   Decherchi, S.; Berteotti, A.; Bottegoni, G.; Rocchia, W.; Cavalli, A. The ligand binding mechanism to purine nucleoside phosphorylase elucidated via molecular dynamics and machine learning. *Nat .Commun.* **2015**, *6*, 6155.

(51)   Case, D. A.; Darden, T. A.; Cheatham, T. E.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Walker, R. C.; Zhang, W.; Merz, K. M.; Roberts, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Swails, J.; Götz, A. W.; Kolossváry, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Wolf, R. M.; Liu, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Cai, Q.; Ye, X.; Wang, J.; Hsieh, M.-J.; Cui, G.; Roe, D. R.; Mathews, D. H.; Seetin, M. G.; Salomon-Ferrer, R.; Sagui, C.; Babin, V.; Luchko, T.; Gusarov, S.; Kovalenko, A.; Kollman, P. A. *Amber 12*, University of California: San Francisco, CA, 2012.

(52)   Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.* **2008**, *4*, 435-447.

(53)   Siragusa, L.; Cross, S.; Baroni, M.; Goracci, L.; Cruciani, G. BioGPS: navigating biological space to predict polypharmacology, off-targeting, and selectivity. *Proteins* **2015**, *83*, 517-532.

Table of Contents graphic