HHS Public Access

Author manuscript

Curr Drug Metab. Author manuscript; available in PMC 2018 April 10.

Published in final edited form as:

Curr Drug Metab. 2017 ; 18(6): 556–565. doi:10.2174/1389200218666170316093301.

Drug Metabolism in Preclinical Drug Development: A Survey of the Discovery Process, Toxicology, and Computational Tools

Naiem T. Issa[1,#], **Henri Wathieu**[1,#], **Abiola Ojo**[3], **Stephen W. Byers**[1,2], and **Sivanesan Dakshanamurthy**[1,2,*]

[1]Georgetown-Lombardi Comprehensive Cancer Center and Department of Oncology, Georgetown University Medical Center, Washington DC, 20057 USA

[2]Department of Biochemistry & Molecular Biology, Georgetown University, Washington DC, 20057, USA

[3]College of Pharmacy, Howard University, Washington, DC 20059, USA

Abstract

Increased R & D spending and high failure rates exist in drug development, due in part to inadequate prediction of drug metabolism and its consequences in the human body. Hence, there is a need for computational methods to supplement and complement current biological assessment strategies. In this review, we provide an overview of drug metabolism in pharmacology, and discuss the current *in vitro* and *in vivo* strategies for assessing drug metabolism in preclinical drug development. We highlight computational tools available to the scientific community for the *in silico* prediction of drug metabolism, and examine how these tools have been implemented to produce drug-target signatures relevant to metabolic routes. Computational workflows that assess drug metabolism and its toxicological and pharmacokinetic effects, such as by applying the adverse outcome pathway framework for risk assessment, may improve the efficiency and speed of preclinical drug development.

Keywords

Preclinical drug development; computational tools; *in silico* prediction; drug-target signatures; toxicology

I. INTRODUCTION

Small-molecule drug development remains an economically risky endeavor. New drug development spending is constantly increasing, currently at over $1.5 billion per drug [1] with a timeline spanning 10–15 years depending upon the complexity of the disease and pipeline [2]. Unfortunately, the failure rate is extremely high, reaching above 90% [3]. Drugs

*Address correspondence to this author at the Georgetown-Lombardi Comprehensive Cancer Center and Department of Oncology, Georgetown University Medical Center, Washington DC, 20057 USA; Tel: 202-687-2347; sd233@georgetown.edu.
#contributed equally

CONFLICT OF INTEREST

The authors confirm that this article content has no conflict of interest.

can fail at various stages of the drug development pipeline spanning the preclinical and clinical phases. Inaccurate and imprecise prediction of drug metabolism can play a major role in such failures.

In the last decade, there has been significant progress in large-scale *in vitro* drug screening technologies such as automated cell proliferation assays and enzyme binding and kinetics [4, 5]. *In vivo* automated drug screening such as the zebrafish assay are also gaining visibility [6]. However, such methods are focused on drug efficacy measured by a change in pathology (*e.g.* halting tumor cell growth) while lacking other relevant physiologic parameters. In the preclinical phase, drug absorption, distribution, metabolism and excretion (ADME) pharmacokinetic properties are as important as efficacy and drug lead optimization. Route of exposure, subsequent bioavailability, and metabolic biotransformation affect the ability of a drug to reach the intended target in the desired bioactive form as well as its toxicity.

Drug metabolism poses a unique optimization problem in preclinical development. Metabolism exhibits heterogeneity among organs as well as between patients. Chemical modifications include hydroxylation, reduction, and hydrolysis, among others. These reactions are mediated by proteins which are differentially expressed in different organs (*e.g.* liver hepatocytes predominantly expressing cytochrome P450 enzymes) and may have a wide range of nucleotide polymorphisms [7]. Furthermore, the drug must make contact with these proteins for biotransformation to occur. Consideration of drug metabolism also encompass drug transport into and out of target cells. Transport is mediated by diffusion dynamics as well as specialized proteins such as P-glycoprotein for drug efflux and organic anion transporters (OATs) for drug uptake [8]. Thus, metabolism of small-molecule drugs is an important aspect of the drug discovery pipeline.

This review will focus on clinical aspects of metabolism in human physiology and discuss current state-of-the-art *in vitro* and *in vivo* tools for studying drug metabolism in the preclinical setting. We also discuss the importance of *in silico* strategies to overcome bottlenecks in the discovery pipeline and survey current computational methods and tools employed in academic and industry settings for modeling drug metabolism.

II. DRUG METABOLISM IN CLINICAL HUMAN PHARMACOLOGY

Drug metabolism is the metabolic process in which the chemical structure (parent compound) of a drug is converted into metabolites to facilitate elimination from the body. The principal sites of drug metabolism are the gut and liver due to high levels of metabolic enzymes in these tissues.

Drug enzyme metabolism involves Phase I reactions (oxidation, reduction and hydrolysis), with subsequent Phase II (conjugation) reactions. The primary goal of this enzymatic activity is to make the drug easier to excrete. Phase I reactions involve the termination of drug activity or the conversion of a prodrug into its active form. Phase I provides a reactive functional group on the compound that inactivates the drug, while a Phase II reaction consists of a conjugation reaction with an exogenous substance (*i.e.* glucuronic acid, sulfate, glycine). Metabolites that stem from Phase II reactions are more readily excreted in the urine

(by the kidneys) and bile (by the liver) than those formed in Phase I. It is important to note that the Phase designation reflects functional, not sequential, classification.

Pharmacokinetics is generally defined as what the human body does to a drug. Pharmacokinetics consists of the study of the time course of drug absorption, distribution, metabolism, and excretion (ADME). The ADME scheme is broken down into the following components: absorption deals with the process of a substance entering blood circulation; distribution deals with the dissemination of a substance to various tissues and organs of the body; metabolism deals with biotransformation; excretion consists of the removal of the substance from the body. Figure 1 shows a general schematic representation of oral drug absorption and metabolism en route to hypothetical lung tumor cells, which also contain their own unique metabolic enzymes and transporters that alter drug efficacy and toxicity.

In the preclinical setting of drug development, clinicians utilize pharmacokinetics to assess drug levels (*i.e.* toxicity) in order to optimize drug therapy. During preclinical development, a common optimization problem deals with heterogeneity in the rate at which humans metabolize drugs. In some cases, humans can metabolize a drug so rapidly that the desired therapeutic blood and tissue concentrations are not met, while in other cases, the metabolism may occur so slowly that the dose leads to toxic effects. Normal changes in human physiology can affect the metabolism of a drug. An individual's drug metabolism rate is influenced by genetic factors, comorbid conditions (*i.e.* chronic liver disorder and heart failure) and drug interactions (*i.e.* induction or inhibition of metabolism).

Cytochrome P450 (CYP450) enzymes are primarily located in the liver and intestine and metabolize the majority of drugs through oxidation. CYP450 enzymes can either be induced or inhibited by various drugs and substances, which results in drug interactions that lead to toxicity or reduction in therapeutic effect. The alteration of ADME characteristics during inflammation can often be attributed to an enzymatic change in drug metabolism. The decrease in the expression of CYP450, which is the primary drug metabolizing enzyme in the liver, following inflammation would be an example of this [9].

Tissue-specific extracts from liver or muscle can be utilized to investigate the role of a transporter in energy related metabolism [9]. Common transporters involved with the alteration of a drug's ADME consists of P-gp, BCRP, OATP1B1/1B3, OAT1/3, OCT1 and OCT2 [10]. The unique expression level and transporting state of a transporter may be modified in relation to a systemic response (*i.e.* immune response) to an exogenous or endogenous stimulant [9]. The alteration of compound ADME during inflammation can also be attributed to transporter changes in drug metabolism. For example, recent *in vitro* and *in vivo* studies showed that inflammatory conditions caused the down regulation of MDRs, MRPs and OATPs. For this reason, the ADME of drugs, which are substrates of these transporters, are affected overall [11].

III. BIOLOGICAL ASSESSMENT OF DRUG METABOLISM IN PRECLINICAL DRUG DISCOVERY

The metabolic biotransformation of drugs is a key determinant of elimination and toxicity in humans. The primary goals of drug metabolism assessment at the preclinical stage are to resolve metabolic stability, identify and quantify primary metabolites, pinpoint metabolic routes by enzyme catalysis as well as inhibition and induction, and measure the potential for drug-drug interactions [12]. Because the major site of drug metabolism is the liver, most *in vivo* and *in vitro* drug metabolism assessments center around hepatic models. Drug metabolism nonetheless occurs in other organs and tissues, such as the lungs, kidneys, and intestine, as illustrated in Fig. (1), with varying consequences on the pharmacokinetic profile of a drug [13]. Therefore, comprehensive biological evaluation of drug metabolism will require both hepatic and extra-hepatic models.

Here, we assess to what extent tools and models can pinpoint the kinetics of a drug, thus helping to derive its biological mechanism and inform both pharmacology and toxicology. In addition, we consider practical factors such as cost and complexity.

In vitro tools have particular advantages in assessing some variables of drug metabolism, even when compared to *in vivo* models. Traditionally used *in vivo* models have low and often fleeting biofluid concentrations of xenobiotic metabolic intermediates that require highly sensitive analytic techniques to capture and characterize [14]. *In vitro* tools, by contrast, allow for precise cellular exposure to a given drug without consideration of how global ADME mechanisms will limit the utility of bioactivity assays and metabolite profiling techniques. One obvious benefit of an *in vivo* animal model is the ability to reproduce the interplay of tissues and organ systems that maintain homeostasis in humans. There are, however, interspecies physiological disparities that cause unanticipated failures in the drug development process, which *in vitro* models may avoid by the use of human cell cultures. Comparative cross-species studies of drug metabolism *in vitro* can in fact highlight human-specific metabolic routes and ultimately improve drug safety estimates [15]. Importantly, *in vitro* tools do not present the ethical concerns that limit the scope and long-term usability of *in vivo* models. While *in vitro* models may never become stand-alone tools for characterizing drug metabolism, they are practical and economical proxies for *in vivo* models, and are a mainstay in informing the latter stages of preclinical and clinical drug development.

Subcellular Fractions

Microsomes are subcellular fractions derived from the endoplasmic reticulum, and present a simple, reproducible, and long-lasting source of membrane-bound metabolic enzymes from which metabolic activity can be assessed on an enzyme-specific level. Implicated enzymes cover multiple stages of metabolism and can include those involved in oxidation, reduction, hydrolysis, and conjugation, but microsomes are primarily useful for the CYP superfamily of enzymes and uridine 5′-diphospho-glucuronosyltransferase (UGT) [16, 17]. Microsomes can easily be derived from any organ- and species-specific tissues, but during preparation, high-lability enzymes can be negatively affected and others require the addition of cofactors

during the preparation process [18]. Nonetheless, models such as human liver microsomes (HLMs) are a reasonably good way to assess inter-individual and demographics-driven differences in enzyme kinetic activity [19]. Such findings can in turn reliably predict the metabolic profile variations that will occur in *in vivo* and clinical studies due to individual or population physiological variations. HLMs are also useful for evaluating intrinsic clearance of drugs, which can predict to what extent metabolism dictates elimination *in vivo*, and are often more reliable than hepatocyte cultures in this regard [20]. To resolve consequential drug-drug interactions, wherein one drug affects the activity of another in either synergistic or antagonistic ways, HLMs can provide a framework for enzyme inhibition assays that indicate how the metabolic routes of other drugs will be suppressed [21]. Enzyme induction similarly modifies the metabolic routes and clearance of chemical compounds, having implications for drug-drug interactions that HMLs can illuminate in preclinical stages [22]. Enzyme induction typically exhibits high inter-species variability [23], underscoring the usefulness of subcellular fractions or other human-derived *in vitro* approaches. S9 fractions are subcellular fractions that contain both cytosolic and membrane-bound metabolic enzymes, thereby benefitting from a greater breadth of Phase I and Phase II mechanisms from which to assess metabolic stability [24]. Subcellular fractions, including HLM, S9, and cytosolic models are often used in conjunction due to variations in enzyme concentrations that in isolation can render some metabolites undetectable [22, 25].

Cell-based Models

At the preclinical stage, two critical characteristics of a proper drug metabolism assessment model are (1) consistency in its bioactivity and reproducibility, and (2) a human-like enzymatic profile as it relates to drug metabolism. In cell-based *in vitro* models, these are often conflicting features. Primary human hepatocytes (PHHs), for example, maintain enzymatic capacity when cultured and are therefore ideal for reflecting bioactivity in the human liver [26]. PHHs are essential tools in the assessment of drug transporters, a feature of drug metabolism that subcellular fraction models do not address [27, 28]. Unfortunately, PHHs lose proliferative ability in isolation and have significant variability in enzyme expression, causing cell source limitations [29]. Nevertheless, there have been a number of promising efforts to induce PPH culture expansion [30, 31]. Liver cell lines such as HepaRG and HepG2, as well as stem cell-derived models such as hepatocyte-like cells (HLCs), exhibit consistent metabolic parameters, are readily available, and have unlimited proliferative potential [32, 33]. Unfortunately, the transformation or immortalization of hepatic cells causes shifts in CYP metabolism and distinct morphological changes [33]. One solution to the difficulty of long-term maintenance of PHHs is the co-culturing of these cells with other cell types, such as non-parenchymal cells of the liver, which has been reported to promote an *in vivo*-like microenvironment, increasing long term applicability and assessment of hepatotoxicity [34]. An important benefit of cell culture models in this light is the ability to reflect both physiological and pathophysiological states, which may serve to illuminate contraindications and the effects of homeostatic imbalance on drug metabolism [35]. This particular dimension of drug metabolism is logistically difficult and uneconomical in *in vivo* approaches. An area where *in vivo* models have the advantage is in multi-organ and inter-organ interactions, but innovative cell-culture systems for liver-based multi-organ models are beginning to have an impact, as reviewed by Bale *et al.* [29].

3D cell culture, especially using multiple cell types, is a major attempt to mitigate the disadvantages of 2D PHH models in evaluating drug metabolism. By increasing longevity and mimicking native cell-cell interactions, cell-matrix interactions, and cell polarity, 3D culture shows promise as a standard for drug metabolism assessments [36]. Costa *et al.* provide a recent description of cutting-edge 3D environment strategies as they relate to drug metabolism [14]. Tissue slice cultures (TSCs) are more representative of *in vivo* conditions and maintain the native cell-cell and cell-matrix interactions, as well as functional structures, that are necessary to preserve the phenotype of a tissue engaged in drug metabolism [37, 38] However, TSCs have presented difficulties in evaluating metabolism in the long term due to limited viability and loss of metabolic properties [39]. TSCs have been proven especially useful to assess multi-organ parameters and pathophysiological states in the context of drug metabolism, especially when associated cell culture models are unavailable or limiting, which is typically the case for extrahepatic purposes [40–42].

In vivo Models

Although *in vitro* preclinical drug metabolism assessments can provide useful results that inform our understanding of how certain cell and tissue types modulate a given drug, *in vivo* biological testing most closely predicts toxicity and many other ADME parameters, and is required for the approval of a drug for clinical use. Developments in *in vivo* assessments, such as an increased understanding of interspecies differences [43], as well as the implementation of chimeric animal models, are improving the predictive power of preclinical drug development. For example, Nishimura *et al.* demonstrated recently that using a chimeric mouse model with a humanized liver improved the accuracy of preclinical drug assessment as it relates to metabolic profiling and drug-drug interactions [44]. *In-vitro-in-vivo* extrapolation (IVIVE), which entails models for the deduction of *in vivo* model response from *in vitro* parameters, are now well-established in pharmacology, especially for hepatic clearance and for predicting drug-drug interactions [45]. Computational models can be employed to substitute or complement traditional *in vitro* and *in vivo* methods for assessing drug metabolism at the preclinical stage [46]. Moreover, strategic biological assessments are crucial for the validation of *in silico*-derived molecular interactions involving drugs and their metabolites, as well as computational platforms that seek to reproduce or expand one or more aspects of the current drug metabolism paradigm.

Metabolomics

Biological assessments of drug metabolism often require high-precision analytical methods to identify and quantify metabolites that arise from the biotransformation of a drug [47]. Metabolomics is most commonly used to characterize the metabolic footprints of pathophysiological states, elucidating disease biomarkers and providing evidence for the discovery of disease perturbation mechanisms [48]. While targeted metabolomics is applied to identify and quantify known and often endogenous metabolites [49], untargeted metabolomics entails a comprehensive analysis of metabolites in the context of a given *in vitro* or *in vivo* model, including unknown compounds. The latter application is most suited to drug metabolism assessments, because novel exogenous metabolites are continuously discovered as new drugs are assessed and increasingly sensitive mass spectrometry technologies are developed [50]. In fact, during modern *in vitro* and *in vivo* studies of

preclinical drug development, achieving the goals of these assessments as stated above depends on such technologies [51, 52] In this context, the difficulty of quantifying highly reactive metabolites can be addressed both by improved assays and by devising novel biological assessment systems, as reviewed by Park *et al.* [53]. Refined and large-scale data arising from metabolomic techniques in turn allows for an improved understanding of metabolic enzyme regulation and other functional insights [49]. Looking forward, continued development of analytical techniques for functional and structural metabolic data will likely contribute to increasingly precise and realistic computational models to bolster the evaluation of drug metabolism in preclinical drug development.

IV. COMPUTATIONAL TOOLS FOR STUDIES OF DRUG METABOLISM

One major aspect of drug metabolism is the interaction of the drug with a protein, often a cytochrome (CYP), to alter its chemical composition. Predicting these interactions is desirable in the preclinical setting. Many computational tools exist that attempt to accurately predict these drug-target interactions. A broad but non-exhaustive survey of these tools follows:

Computer-aided drug design (CADD) tools used to predict drug-target signatures can be functionally categorized as shown in Fig. (2). These include structure-based (or protein target-based), ligand-based, and a combination of the two (proteochemometric). Chemoinformatics approaches, such as quantitative structure-activity relationships (QSAR), also leverage known binding affinity data of drug substrates to various metabolizing enzymes for the prediction of other drug-enzyme pairs. Here, we detail these methods and examine some recent applications to drug metabolism.

Structure-Based Computational Methods

When the structure of the metabolizing enzyme is known, structure-based approaches can be utilized to predict drug-enzyme interactions. X-ray crystallography and nuclear magnetic resonance (NMR) are typically used to determine biologically relevant enzyme structures in apo (ligand absent) and holoenzyme (ligand bound) states. These structures are publicly available in online repositories such as the PDB (www.rcsb.org) and EMBL (www.embl.de). If the structure for the enzyme of interest has not been determined, homology modeling can be used to predict reliable three-dimensional models. Homology modeling leverages known structures of proteins that have >30% similarity in amino acid sequence as templates for building the 3D model [54]. Free and commercial software such as SWISS-MODEL [54] and Prime [55], respectively, are available for this purpose. Successful use of homology modeling has been used for human cytochrome P450 enzymes and drug transporters. Unwalla *et al.* utilized ligand-bound CYP2C5 complexes as templates to build a homology model of CYP2D6, an important enzyme involved in first-pass metabolism [56]. The CYP2D6 model was then used for docking of known substrates to identify putative catalytic binding sites.

Docking is the process of computationally “fitting” two molecular entities and predicting their free energy of binding [57]. A variety of docking programs exist and have been widely employed in drug discovery to predict drug-protein interactions, such as GLIDE [58],

Autodock [59], and GOLD [60]. In small-molecule drug discovery, the drug(s) are docked into the specified binding pocket within the 3D structure of the protein target (crystal structure, homology model, *etc.*). Docking has been widely successful in predicting effective drug-protein target interactions important disease therapeutics [61–64].

While largely successful, docking may miss many true drug interactions. This may be due to non-optimal calculations of the free energy of binding due to simplified conditions (*e.g.* protein and drug docking simulation performed in a vacuum devoid of water solvent or membrane effects). de Graaf *et al.* found that crystallographic water molecules in various CYP structure active sites increased docking accuracy [65]. In addition, the protein starting structure is a critical consideration prior to running structure-based methods. At times, the 3D protein structure model may not reflect the true binding capacity of the target. This is particularly important for CYP enzymes as they exhibit great flexibility and multiple binding modes for a broad range of substrates [66]. Thus, defining the appropriate binding site region becomes difficult. One way to overcome this limitation is to use molecular dynamics simulation (MDS). MDS simulates Newtonian motion of the protein's atoms and the interatomic forces within the protein itself as well as forces between protein atoms and solvent atoms, membrane atoms, drug atoms, and other protein-protein atoms [67]. Doing so allows to the prediction of physically plausible structural fluctuations of the protein target in the desired in situ environment (*e.g.* aqueous solvent with physiological salt concentration). If simulations are conducted over a considerable length of time, low-energy conformational states with different binding modes may be uncovered and can serve as the structures for further simulations [68]. In addition, MDS is useful for assessing the stability of proposed drug-target binding signatures. Long-scale simulations help assess the conformational stability of a drug within the binding pocket as well as the diffusion into and sampling of the binding pocket space while maintaining protein flexibility. This method has been successful in assessing complicated structural mechanisms of action such as G protein-coupled receptors [69, 70]. Watanabe *et al.* utilized MDS to study the structural flexibility of wild-type and mutant CYP1A2 [71]. They showed that polymorphisms not only affect local structural integrity but also distant structural dynamics, thus altering ligand binding and recognition. Similarly, Fukuyoshi *et al.* investigated the effects of polymorphisms on CYP2D6 using MDS [72]. CYP2D6 contains the flexible F–G loop structure, which is thought to contribute to the substrate access channel and ligand diversity [73]. Through MDS, the team discovered that some mutations (*e.g.* CYP2D6.10, 14A, and 61) were enzymatically relevant by keeping the F–G loop in a closed position, inhibiting substrate entry into the catalytic site. Since CYP2D6 has over 100 naturally occurring gene mutations [74], MDS is an efficient feasible strategy for selecting mutants to undergo experimental observations. Studies such as these have great utility in computational pharmacogenomics and personalized medicine.

Structure-based methods tend to be used in combination to optimize for both accuracy and computational efficiency. For example, docking is less computationally expensive and can be used to screen very large databases of drugs efficiently [75]. Top-ranked drugs from docking can then undergo MDS. Alternatively, in the absence of crystal structures, homology models can be built then refined using MDS. Though MDS, multiple low-energy conformations, or "snapshots", can be utilized for docking. This "consensus approach"

increases accuracy by investigating drug binding of multiple protein conformations, some of which may be important in catalysis but may be of higher energy and short-lived relative to the conformational state achieved by the global minimum. This approach is highlighted by Mendieta-Wejebe *et al.* who built and refined homology models of CYP1A1 and CYP2B1 to predict potential biotransformation of two acetylcholinesterase inhibitors 4-(4′-hydroxy-phenylamino)-4-oxo propanoic acid and 1H-pyrrolidine-1-(4′-hydroxy-phenyl)-2,5-dione [76]. For each isoform, 15 snapshots were obtained via MDS, all of which were used for the docking of the drugs. Docking results implied that neither molecule would adopt the proper conformation for biotransformation into toxic metabolites in either CYP enzyme, which was confirmed by subsequent experimental studies.

The aforementioned strategies depend on the availability of a 3D structure for the enzyme of interest, which may be obtained via crystallization or homology modeling. At times, it may be difficult to obtain a structure given difficulties in crystallization or the lack of an appropriate template from which to build a reasonable homology model. Thus, ligand-centric methods may be required, which have their own benefits and drawbacks. These methods are discussed next.

Ligand-Centric, Proteochemometric & QSAR Computational Methods

Information extracted from drug chemical structures can be used to predict drug-enzyme interactions. The similarity approach is commonly used where the similarity between a drug known to bind the enzyme of interest and a query drug are calculated. Under the similarity principle, it is assumed that the greater the similarity, the greater the likelihood of the drug interacting with the protein [77]. Similarity can be calculated based on a number of metrics, such as topological parameters (*e.g.* shape [78]), physicochemical descriptors (*e.g.* electrostatics [79, 80]) as well as different combinations (*e.g.* pharmacophore similarity where the type of functional groups and their orientation in 3D space are considered [81]). In addition, these descriptors can be used for building quantitative structure-activity relationship (QSAR) models for a target of interest.

Pharmacophores provide a three-dimensional topological representation of important functional groups implicated in substrate binding. Pharmacophore models are derived from bioactive conformations of known active ligands using software such as LigandScout [82] or SPARTAN [83]. A limitation is inherently imposed by small numbers of known active ligands or potential lack of chemical diversity of those ligands. Nonetheless, pharmacophore models can be used to efficiently screen large libraries of molecules for potential substrates that are able to adopt the appropriate structural conformation and functional group orientation. An early study by de Groot *et al.* exemplifies the development of pharmacophore modeling for CYP2D6-catalyzed N-dealkylation reactions [84]. An example where pharmacophore screening is integrated into a larger *in silico* pipeline for drug metabolism is highlighted by Rakers *et al.* [85]. They developed a pharmacophore model for sulfotransferase 1E1 and used it to successfully screen for drug-like molecules that were known to bind SULT1E1 as well as experimentally validated nine compounds previously not known to bind. A combination approach was used by first utilizing molecular dynamic simulations of SULT1E1 apo and cofactor-bound structures to model the flexibility of the

active site and obtain different conformational states for subsequent ensemble docking of known SULT1E1 active ligands. Docking conformations were then prioritized based on catalytic competency (*e.g.* distance between the ligand and cofactor) and used to establish different pharmacophore models, which were further refined using support vector machines (SVM). Ultimately, a prediction model was created combining a pharmacophore fit score (determined using LigandScout) with other SVM-determined physicochemical descriptors. This model was used to screen the DrugBank library (www.drugbank.ca) resulting in 68 hits, 28% of which were validated in the literature and nine new compounds which were experimentally validated. Pharmacophore models can be especially useful for large active sites capable of binding diverse substrates, whereby different models can highlight different substrate classes.

To date, there are two approaches to QSAR methods that predict the metabolism of substrates and inhibitors by CYP enzymes. The first consists of the application of QSAR to develop ADMET (absorption, distribution, metabolism, elimination and toxicity) models [86–91]. This model is known to be difficult to interpret and less predictable [92]. The second approach is focused on the development of QSAR models for molecules that pertain to specific individual CYPs [92]. For example, CYP1A1 is known to be found in the lungs, lymphocytes and skin and has be inferred in cancers that stem from polycyclic aromatic hydrocarbons (PAHs). PAHs are the primary substrates for metabolism by CYP1A1. Genestes *et al.* developed a QSAR model for the metabolism of 32 different compounds of the PAH series by CYP1A1 [93] utilizing multilinear regression analysis. Shimada *et al.* investigated the metabolic activities of the PAHs [94]. The QSAR analysis utilized energy descriptors developed during the docking process. The three main descriptors consisted of: (1) HOMO energy (energy of the highest occupied molecular orbital), (2) Hydrogen bond acceptor atoms in the ligand, and (3) PMF04 scoring function derived from docking studies correlated with inhibition potency. The predictability of the linear model utilized the neural methods for improvement. Using the same descriptors, the artificial neural networking (ANN) model illustrated a better prediction for CYP1A1 pro-metabolic property compared to the linear model.

CYP1A2 is known to be expressed in the liver and plays a role in the activation of aromatic and heterocyclic amines, PAHs and various therapeutic drugs. Procarcinogen activation by CYP1A2 increases a patient's susceptibility to cancer while inhibition of CYP1A2 can contribute to cancer prevention. For example, flavonoids have been known to attribute to preventive cancer effects due to the interaction with CYP1A enzymes [92]. Roy *et al.* utilized a cluster technique with 2D descriptors addressing topological, physicochemical and structural indices combined with 3D descriptors for QSAR analyses [95]. This analysis utilized chemometric tools such as the genetic partial least square (G/PLS).

Proteochemometric methods are also QSAR-related models that further integrate information from proteins as well as drug-centric fingerprints [96]. For example, a unified proteochemometric (PCM) model for the prediction of the inhibition of CYP450 isoforms was developed and made publicly available under the Bioclipse Decision Support open source system [97]. PCM creates a unification of models for multiple proteins that interaction with multiple ligands through the correlation descriptors [98, 99]. The descriptors

are based on the following amino acid properties: hydrophobicity, normalized van der Waals volume, polarity, polarizability, charge, secondary structure and solvent accessibility [97]. The PCM model utilized non-linear data analysis techniques such as Support Vector Machine, Random Forest and k-Nearest Neighbor method to perform binary classification and probability estimates. This proteochemometric CYP model provides the opportunity to draw or import chemical structures and predict quetiapine inhibition of CYP2D6, propranolol inhibition of CYP1A2 and CYP2D6, and fluvoxamine inhibition of CYP1A2, CYP2C9, CYP2C19, CYP2D6 and CYP3A4 [97]. This PCM model advantageous in its extensibility, potentially extending to new CYP isoforms and polymorphic CYP forms.

Adverse Outcome Pathways in the Computational Assessment of Drug Metabolism

As the complexity of metabolism extends beyond immediate drug-enzyme interactions, it is critical to consider higher-order pathway and cellular perturbations that the parent drug may have on the cellular metabolic machinery. These perturbations are consequential in biological assessments of drug metabolism as described earlier in this review, wherein they can potentially explain tissue-specific and inter-organ metabolic routes, drug-drug interactions, and drug toxicity [100]. Equally important in this context are the metabolites arising from biotransformation of a drug, which can contribute to both the therapeutic action of that drug as well as its adverse effects [101]. The adverse outcome pathway (AOP) is a recently developed framework in computational toxicology which seeks to bridge molecular initiating events (MIEs) such as those predicted by the computational tools discussed in this review, and adverse events at the organ or organismal level [102]. The AOP Knowledge Base (AOP-KB) is the primary collaborative effort currently underway to catalog crowdsourced experimentally derived AOPs (www.aopkb.org). Current efforts in computationally predicted AOPs (cpAOPs) were recently reviewed by Oki *et al.* [103].

There are a number of advantages to incorporating the AOP framework to preclinical stages of drug development, particularly in drug metabolism and its relevance to toxicity. Of note is that MIEs are not restricted to chemical-protein interactions; chemical interactions with other macromolecular structures can lead to a wide variety of adverse effects [104]. Another central characteristic of AOPs is that they are defined by MIEs that do not depend on the nature of the implicated chemical compound [105]. For example, AOPs have been experimentally developed for the case of environmental chemicals inhibiting the acetylcholinesterase enzyme at distinct synaptic sites [106]. This MIE has associated adverse effects that range from mild to lethal, and is shared by toxic chemicals such as venoms and insecticides, but also by drugs with broad medicinal relevance [107]. Thus, while ecotoxicity is the primary focus of AOP studies, there is a distinct potential for the application of AOPs to preclinical drug development. Vinken *et al*, for example, developed AOPs describing the stepwise mechanistic association between the drug-implicated MIE of inhibition of the bile salt export pump and the organismal development cholestatic liver injury [108]. When combined with exposure and ADME assessments, the AOP framework has recently been demonstrated by Philips *et al.* to identify drug metabolites and parent compounds that in fact engage in an MIE and therefore avoid unnecessary costs of comprehensive biological assessment [109]. Figure 3 outlines how the AOP concept might contribute to a preclinical drug metabolism assessment workflow, wherein predicted or experimentally established

exogenous metabolites are screened for adverse outcomes according to the current knowledgebase. The application of AOPs to the preemptive assessment of drug metabolism is preclinical stages is just beginning to emerge. The development of the AOP-KB and complementary cpAOP models, if regularly implemented, will likely render preclinical assessment of drug metabolism more efficient.

CONCLUSION

It is clear that drug metabolism is a critical component in understanding drug efficacy, resistance and adverse effects. Prospective early identification of interactions between drugs and the enzymes implicated in human drug metabolism is desirable in drug discovery to avoid excessive costs and failures, particularly in the clinical phases. In addition, it is important to identify drugs that may potentiate or inhibit enzyme metabolic activity. Computational tools such as docking, molecular dynamics and QSAR models facilitate this process.

Pharmacogenomics is becoming increasingly utilized in the clinic as it has become understood that drug metabolism exhibits inter-patient variability. A clinically relevant example worth noting is dihydropyrimidine dehydrogenase (DPD), encoded by the *DPYD* gene, which serves as the rate-limiting step in the metabolism of the cancer chemotherapeutics 5-fluorouracil (5-FU) [110]. DPD deficiency excess 5-FU accumulation and toxicity such as mucositis and neutropenia [111]. A spectrum of DPD deficiency due to single nucleotide polymorphisms has been identified. Most notable is the IVS14+1G>A mutation in intron 14 coupled with exon 14 deletion [110]. Others include 2846A>T in exon 22 [112] and T1679G in exon 13 [113]. Similarly, CYP450 pharmacogenomics has been used to understand inter-patient differences in warfarin metabolism, a widely used anticoagulant that requires close blood level monitoring [114]. This information helps identify patient populations at risk of toxicity, the need for dose adjustments and potentially antidotes in cases where the proposed drug provides life-saving treatment. With the growing body of knowledge in pharmacogenomics, understanding the functional significance of the genetic variations becomes increasingly crucial. Large online databases such as PharmGKB (www.pharmgkb.org) have been built for functional annotations as well as their clinical relevance. However, many genetic variants remain to be characterized through painstaking *in vitro* and *in vivo* methods. Computational tools, especially structure-based approaches such as homology modeling and molecular dynamics simulations, will prove indispensable in characterizing the structures of these variants as well as high-throughput screening of drugs for predicting their putative functional consequences. Prospective computational pharmacogenomics through QSAR and proteochemometrics will also add to our understanding and enhance patient safety throughout the drug development process. Implementation of a potential workflow, one of many possible permutations in the drug discovery process is shown in Fig. (3) for a drug that is already known to bind to a therapeutic target of interest for a disease state. This particular workflow highlights the potentiality of optimizing the starting drug through chemical permutations to form a congener library and subsequent virtual screening of the library against the target of interest, important ADMET enzymes and finally ranking the congener with the optimal combination of target binding, bioavailability and toxicity minimization.

The computational tools listed throughout each have their strengths and weaknesses, especially when methods such as QSAR and pharmacophore models are highly dependent on the training set of ligands. Combinatorial approaches that pool the different methods may be synergistic and increase the accuracy of predictions. The previously noted "consensus approach" utilizing multiple structural conformations, or "snapshots", of a single protein through MDS can increase accuracy when combined with virtual screening workflows such as docking. Furthermore, protein-centric techniques can be integrated with ligand-centric approaches to further optimize predictive accuracy. Multi-component discovery pipelines are becoming more commonplace as computational efficiency is increasing with decreasing costs.

While prediction of drug-protein binding is a critical goal of *in silico* methods, that alone is insufficient for formulating physiologically relevant models of drug biotransformation and the downstream effects on efficacy, toxicity and dosing parameters. Moving forward, enzyme kinetics will be an important consideration for clinical translation as rates of metabolite production as well as the effects on polypharmacy when drugs compete for the same metabolizing catalytic site. In addition, consideration must be given to enzyme inducers/inhibitors and drug-drug interactions. Understanding the regulatory elements behind biotransformation and transport enzyme expression, such as the human nuclear receptor PXR inducing CYP3A [115], will prove indispensable in personalized medicine and pharmacogenomics, allowing us to better understand the interpersonal differences in pharmacological biotransformation that extends beyond structural differences in key enzymes due to genetic polymorphisms.

In silico methods are increasing in accuracy as our understanding of drug metabolism grows from pharmacological studies and structural biology. Their use can greatly facilitate the drug development process by reducing downstream failures due to patient harm from toxic metabolites, lack of efficacy due to quick inactivation of active metabolites, and can also diminish the need of live animal use in preclinical development. Future discovery pipelines, especially in the PK/PD sector, will undoubtedly incorporate computational modeling and screening so as to predict the appropriate bioavailability, efficacy and toxicity of new investigational drugs and ultimately mitigate the downstream risk in animals and humans.

Acknowledgments

The authors wish to acknowledge DOD grant CA140882 (SB, SD), CCSG grant NIH-P30 CA51008 and Georgetown Lombardi Cancer Center.

References

1. Dickson M, Gagnon JP. Key factors in the rising cost of new drug discovery and development. Nat Rev Drug Discov. 2004; 3(5):417–429. [PubMed: 15136789]
2. Paul SM, Mytelka DS, Dunwiddie CT, Persinger CC, Munos BH, Lindborg SR, Schacht AL. How to improve R&D productivity: The pharmaceutical industry's grand challenge. Nat Rev Drug Discov. 2010; 9(3):203–214. [PubMed: 20168317]
3. Calcoen D, Elias L, Yu X. What does it take to produce a break-through drug? Nat Rev Drug Discov. 2015; 14(3):161–162. [PubMed: 25722234]

4. Mosmann T. Rapid colorimetric assay for cellular growth and survival: Application to proliferation and cytotoxicity assays. J Immunol Methods. 1983; 65(1–2):55–63. [PubMed: 6606682]
5. Michelini E, Cevenini L, Mezzanotte L, Coppa A, Roda A. Cell-based assays: Fuelling drug discovery. Anal Bioanal Chem. 2010; 398(1):227–238. [PubMed: 20623273]
6. Laggner C, Kokel D, Setola V, Tolia A, Lin H, Irwin JJ, Keiser MJ, Cheung CY, Minor DL Jr, Roth BL, Peterson RT, Shoichet BK. Chemical informatics and target identification in a zebrafish phenotypic screen. Nat Chem Biol. 2012; 8(2):144–146.
7. Sim SC, Kacevska M, Ingelman-Sundberg M. Pharmacogenomics of drug-metabolizing enzymes: A recent update on clinical implications and endogenous effects. Pharmacogenomics J. 2013; 13(1):1–11. [PubMed: 23089672]
8. DeGorter MK, Xia CQ, Yang JJ, Kim RB. Drug transporters in drug efficacy and toxicity. Annu Rev Pharmacol Toxicol. 2012; 52:249–273. [PubMed: 21942630]
9. Liang Y, Li S, Chen L. The physiological role of drug transporters. Protein Cell. 2015; 6(5):334–350. [PubMed: 25797421]
10. Tweedie D, Polli JW, Berglund EG, Huang SM, Zhang L, Poirier A, Chu X, Feng B. Transporter studies in drug development: Experience to date and follow-up on decision trees from the international transporter consortium. Clin Pharmacol Ther. 2013; 94(1):113–125. [PubMed: 23588318]
11. Petrovic V, Teng S, Piquette-Miller M. Regulation of drug transporters: During infection and inflammation. Mol Interv. 2007; 7(2):99. [PubMed: 17468390]
12. Pelkonen O, Turpeinen M, Uusitalo J, Rautio A, Raunio H. Prediction of drug metabolism and interactions on the basis of *in vitro* investigations. Basic Clin Pharmacol Toxicol. 2005; 96(3):167–175. [PubMed: 15733211]
13. Krishna DR, Klotz U. Extrahepatic metabolism of drugs in humans. Clin Pharmacokinet. 1994; 26(2):144–160. [PubMed: 8162658]
14. Costa A, Sarmento B, Seabra V. An evaluation of the latest *in vitro* tools for drug metabolism studies. Expert Opin Drug Metab Toxicol. 2014; 10(1):103–119. [PubMed: 24205859]
15. Wrighton SA, Ring BJ, Vandenbranden M. The use of *in vitro* metabolism techniques in the planning and interpretation of drug safety studies. Toxicol Pathol. 1995; 23(2):199–208. [PubMed: 7569675]
16. Penner N, Xu L, Prakash C. Radiolabeled absorption, distribution, metabolism, and excretion studies in drug development: Why, when, and how? Chem Res Toxicol. 2012; 25(3):513–531. [PubMed: 22309195]
17. Bertrand M, Jackson P, Walther B. Rapid assessment of drug metabolism in the drug discovery process. Eur J Pharm Sci. 2000; 11:S61–S72. [PubMed: 11033428]
18. Ekins S, Ring BJ, Grace J, McRobie-Belle DJ, Wrighton SA. Present and future *in vitro* approaches for drug metabolism. J Pharmacol Toxicol Methods. 2000; 44(1):313–324. [PubMed: 11274898]
19. Zhang H, Gao N, Tian X, Liu T, Fang Y, Zhou J, Wen Q, Xu B, Qi B, Gao J, Li H, Jia L, Qiao H. Content and activity of human liver microsomal protein and prediction of individual hepatic clearance *in vivo*. Sci Rep. 2015; 5:17671. [PubMed: 26635233]
20. Lu C, Li P, Gallegos R, Uttamsingh V, Xia CQ, Miwa GT, Balani SK, Gan LS. Comparison of intrinsic clearance in liver microsomes and hepatocytes from rats and humans: Evaluation of free fraction and uptake in hepatocytes. Drug Metab Dispos. 2006; 34(9):1600–1605. [PubMed: 16790553]
21. Obach RS, Walsky RL, Venkatakrishnan K, Gaman EA, Houston JB, Tremaine LM. The utility of *in vitro* cytochrome P450 inhibition data in the prediction of drug-drug interactions. J Pharmacol Exp Ther. 2006; 316(1):336–348. [PubMed: 16192315]
22. Hariparsad N, Sane RS, Strom SC, Desai PB. *In vitro* methods in human drug biotransformation research: Implications for cancer chemotherapy. Toxicol In vitro. 2006; 20(2):135–153. [PubMed: 16359840]
23. Martignoni M, Groothuis GM, de Kanter R. Species differences between mouse, rat, dog, monkey and human CYP-mediated drug metabolism, inhibition and induction. Expert Opin Drug Metab Toxicol. 2006; 2(6):875–894. [PubMed: 17125407]

24. Fasinu P, Bouic JP, Rosenkranz B. Liver-based *in vitro* technologies for drug biotransformation studies-a review. Curr Drug Metab. 2012; 13(2):215–224. [PubMed: 22300020]
25. Van den Eede N, Maho W, Erratico C, Neels H, Covaci A. First insights in the metabolism of phosphate flame retardants and plasticizers using human liver fractions. Toxicol Lett. 2013; 223(1):9–15. [PubMed: 23994729]
26. Sahi J, Grepper S, Smith C. Hepatocytes as a tool in drug metabolism, transport and safety evaluations in drug discovery. Curr Drug Discov Technol. 2010; 7(3):188–198. [PubMed: 20843293]
27. Jigorel E, Le Vee M, Boursier-Neyret C, Parmentier Y, Fardel O. Differential regulation of sinusoidal and canalicular hepatic drug transporter expression by xenobiotics activating drug-sensing receptors in primary human hepatocytes. Drug Metab Dispos. 2006; 34(10):1756–1763. [PubMed: 16837569]
28. McGinnity DF, Soars MG, Urbanowicz RA, Riley RJ. Evaluation of fresh and cryopreserved hepatocytes as *in vitro* drug metabolism tools for the prediction of metabolic clearance. Drug Metab Dispos. 2004; 32(11):1247–1253. [PubMed: 15286053]
29. Bale SS, Moore L, Yarmush M, Jindal R. Emerging *in vitro* liver technologies for drug metabolism and inter-organ interactions. Tissue Eng Part B Rev. 2016; 22(5):383–394. [PubMed: 27049038]
30. Shan J, Schwartz RE, Ross NT, Logan DJ, Thomas D, Duncan SA, North TE, Goessling W, Carpenter AE, Bhatia SN. Identification of small molecules for human hepatocyte expansion and iPS differentiation. Nat Chem Biol. 2013; 9(8):514–520. [PubMed: 23728495]
31. Levy G, Bomze D, Heinz S, Ramachandran SD, Noerenberg A, Cohen M, Shibolet O, Sklan E, Braspenning J, Nahmias Y. Long-term culture and expansion of primary human hepatocytes. Nat Biotechnol. 2015; 33(12):1264–1271. [PubMed: 26501953]
32. Lübberstedt M, Müller-Vieira U, Mayer M, Biemel KM, Knöspel F, Knobeloch D, Nüssler AK, Gerlach JC, Zeilinger K. HepaRG human hepatic cell line utility as a surrogate for primary human hepatocytes in drug metabolism assessment *in vitro*. J Pharmacol Toxicol Methods. 2011; 63(1): 59–68. [PubMed: 20460162]
33. Wilkening S, Stahl F, Bader A. Comparison of primary human hepatocytes and hepatoma cell line Hepg2 with regard to their biotransformation properties. Drug Metab Dispos. 2003; 31(8):1035–1042. [PubMed: 12867492]
34. Kostadinova R, Boess F, Applegate D, Suter L, Weiser T, Singer T, Naughton B, Roth A. A long-term three dimensional liver co-culture system for improved prediction of clinically relevant drug-induced hepatotoxicity. Toxicol Appl Pharmacol. 2013; 268(1):1–16. [PubMed: 23352505]
35. Bhogal RH, Hodson J, Bartlett DC, Weston CJ, Curbishley SM, Haughton E, Williams KT, Reynolds GM, Newsome PN, Adams DH, Afford SC. Isolation of primary human hepatocytes from normal and diseased liver tissue: A one hundred liver experience. PloS One. 2011; 6(3):e18222. [PubMed: 21479238]
36. Lu HF, Chua KN, Zhang PC, Lim WS, Ramakrishna S, Leong KW, Mao HQ. Three-dimensional co-culture of rat hepatocyte spheroids and NIH/3T3 fibroblasts enhances hepatocyte functional maintenance. Acta Biomater. 2005; 1(4):399–410. [PubMed: 16701821]
37. Gebhardt R, Gaunitz F. Cell-cell interactions in the regulation of the expression of hepatic enzymes. Cell Biol Toxicol. 1997; 13(4–5):263–273. [PubMed: 9298247]
38. Edwards RJ, Price RJ, Watts PS, Renwick AB, Tredger JM, Boobis AR, Lake BG. Induction of cytochrome P450 enzymes in cultured precision-cut human liver slices. Drug Metab Dispos. 2003; 31(3):282–288. [PubMed: 12584154]
39. Graaf IAD, Groothuis GM, Olinga P. Precision-cut tissue slices as a tool to predict metabolism of novel drugs. Expert Opin Drug Metab Toxicol. 2007; 3(6):879–898. [PubMed: 18028031]
40. De Kanter R, Monshouwer M, Draaisma AL, De Jager MH, De Graaf IAM, Proost JH, Meijer DK, Groothuis GMM. Prediction of whole-body metabolic clearance of drugs through the combined use of slices from rat liver, lung, kidney, small intestine and colon. Xenobiotica. 2004; 34(3):229–241. [PubMed: 15204696]
41. Maund SL, Nolley R, Peehl DM. Optimization and comprehensive characterization of a faithful tissue culture model of the benign and malignant human prostate. Lab Invest. 2014; 94(2):208–221. [PubMed: 24296879]

42. van de Kerkhof EG, de Graaf IA, Groothuis GM. *In vitro* methods to study intestinal drug metabolism. Curr Drug Metab. 2007; 8(7):658–675. [PubMed: 17979654]
43. Sjögren E, Abrahamsson B, Augustijns P, Becker D, Bolger MB, Brewster M, Brouwers J, Flanagan T, Harwood M, Heinen C, Holm R, Juretschke HP, Kubbinga M, Lindahl A, Lukacova V, Münster U, Neuhoff S, Nguyen MA, Peer AV, Reppas C, Hodjegan AR, Tannergren C, Weitschies W, Wilson C, Zane P, Lennernäs H, Langguth P. *In vivo* methods for drug absorption–comparative physiologies, model selection, correlations with *in vitro* methods (IVIVC), and applications for formulation/API/excipient characterization including food effects. Eur J Pharm Sci. 2014; 57:99–151. [PubMed: 24637348]
44. Nishimura T, Hu Y, Wu M, Pham E, Suemizu H, Elazar M, Liu M, Idilman R, Yurdaydin C, Angus P, Stedman C, Murphy B, Glenn J, Nakamura M, Nomura T, Chen Y, Zheng M, Fitch WL, Peltz G. Using chimeric mice with humanized livers to predict human drug metabolism and a drug-drug interaction. J Pharmacol Exp Ther. 2013; 344(2):388–396. [PubMed: 23143674]
45. Rostami-Hodjegan A. Physiologically based pharmacokinetics joined with *in vitro–in vivo* extrapolation of ADME: A marriage under the arch of systems pharmacology. Clin Pharmacol Ther. 2012; 92(1):50–61. [PubMed: 22644330]
46. Kirchmair J, Göller AH, Lang D, Kunze J, Testa B, Wilson ID, Glen RC, Schneider G. Predicting drug metabolism: Experiment and/or computation? Nat Rev Drug Discov. 2015; 14(6):387–404. [PubMed: 25907346]
47. Kumar B, Prakash A, Ruhela RK, Medhi B. Potential of metabolomics in preclinical and clinical drug development. Pol J Pharmacol. 2014; 66(6):956–963.
48. Xia J, Broadhurst DI, Wilson M, Wishart DS. Translational biomarker discovery in clinical metabolomics: An introductory tutorial. Metabolomics. 2013; 9(2):280–299. [PubMed: 23543913]
49. Roberts LD, Souza AL, Gerszten RE, Clish CB. Targeted metabolomics. Curr Protoc Mol Biol. 2012; 30(2):1–24.
50. Patti GJ, Yanes O, Siuzdak G. Innovation: Metabolomics: The apogee of the omics trilogy. Nat Rev Mol Cell Biol. 2012; 13(4):263–269. [PubMed: 22436749]
51. Robertson DG, Frevert U. Metabolomics in drug discovery and development. Clin Pharmacol Ther. 2013; 94(5):559–561. [PubMed: 24145714]
52. Spaggiari D, Geiser L, Rudaz S. Coupling ultra-high-pressure liquid chromatography with mass spectrometry for *in-vitro* drug-metabolism studies. Trends Analyt Chem. 2014; 63:129–139.
53. Park BK, Boobis A, Clarke S, Goldring CE, Jones D, Kenna JG, Lambert C, Laverty HG, Naisbitt DJ, Nelson S, Nicoll-Griffith DA, Obach RS, Routledge P, Smith DA, Tweedie DJ, Vermeulen N, Williams DP, Wilson ID, Baillie TA. Managing the challenge of chemically reactive metabolites in drug development. Nat Rev Drug Discov. 2011; 10(4):292–306. [PubMed: 21455238]
54. Bordoli L, Kiefer F, Arnold K, Benkert P, Battey J, Schwede T. Protein structure homology modeling using SWISS-MODEL workspace. Nat Protoc. 2009; 4(1):1–13. [PubMed: 19131951]
55. Schrödinger Release 2016-3: Prime. Schrödinger, LLC; New York, NY: 2016.
56. Unwalla RJ, Cross JB, Salaniwal S, Shilling AD, Leung L, Kao J, Humblet C. Using a homology model of cytochrome P450 2D6 to predict substrate site of metabolism. J Comput Aided Mol Des. 2010; 24(3):237–256. [PubMed: 20361239]
57. Lengauer T, Rarey M. Computational methods for biomolecular docking. Curr Opin Struct Biol. 1996; 6(3):402–406. [PubMed: 8804827]
58. Friesner RA, Banks JL, Murphy RB, Halgren TA, Klicic JJ, Mainz DT, Repasky MP, Knoll EH, Shelley M, Perry JK, Shaw DE, Francis P, Shenkin PS. Glide: A new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. J Med Chem. 2004; 47(7): 1739–1749. [PubMed: 15027865]
59. Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, Olson AJ. AutoDock4 and AutoDock-Tools4: Automated docking with selective receptor flexibility. J Computat Chem. 2009; 30(16):2785–2791.
60. Verdonk ML, Cole JC, Hartshorn MJ, Murray CW, Taylor RD. Improved protein–ligand docking using GOLD. Proteins: Structure, Function, and Bioinformatics. 2003; 52(4):609–623.
61. DesJarlais RL, Seibel GL, Kuntz ID, Furth PS, Alvarez JC, Ortiz de Montellano PR, DeCamp DL, Babé LM, Craik CS. Structure-based design of nonpeptide inhibitors specific for the human

immunodeficiency virus 1 protease. Proc Nat Acad Sci. 1990; 87(17):6644–6648. [PubMed: 2204060]

62. Li YY, An J, Jones SJ. A computational approach to finding novel targets for existing drugs. PLoS Comput Biol. 2011; 7(9):e1002139. [PubMed: 21909252]
63. Méndez-Lucio O, Tran J, Medina-Franco JL, Meurice N, Muller M. Toward drug repurposing in epigenetics: Olsalazine as a hypomethylating compound active in a cellular context. ChemMedChem. 2014; 9(3):560–565. [PubMed: 24482360]
64. Chen SJ. A potential target of Tanshinone IIA for acute promyelocytic leukemia revealed by inverse docking and drug repurposing. Asian Pac J Cancer Prev. 2013; 15(10):4301–4305.
65. de Graaf C, Pospisil P, Pos W, Folkers G, Vermeulen NP. Binding mode prediction of cytochrome p450 and thymidine kinase protein-ligand complexes by consideration of water and rescoring in automated docking. J Med Chem. 2005; 48(7):2308–2318. [PubMed: 15801824]
66. Oostenbrink C, de Ruiter A, Hritz J, Vermeulen N. Malleability and versatility of cytochrome P450 active sites studied by molecular simulations. Curr Drug Metab. 2012; 13(2):190–196. [PubMed: 22208533]
67. Durrant JD, McCammon JA. Molecular dynamics simulations and drug discovery. BMC Biology. 2011; 9(1):1. [PubMed: 21214944]
68. Ivetac A, McCammon JA. A molecular dynamics ensemble-based approach for the mapping of druggable binding sites. Computat Drug Discov Des. 2012:3–12.
69. Rosenbaum DM, Zhang C, Lyons JA, Holl R, Aragao D, Arlow DH, Rasmussen SG, Choi HJ, Devree BT, Sunahara RK, Chae PS, Gellman SH, Dror RO, Shaw DE, Weis WI, Caffrey M, Gmeiner P, Kobilka BK. Structure and function of an irreversible agonist-[bgr] 2 adrenoceptor complex. Nature. 2011; 469(7329):236–240. [PubMed: 21228876]
70. Huang W, Manglik A, Venkatakrishnan AJ, Laeremans T, Feinberg EN, Sanborn AL, Kato HE, Livingston KE, Thorsen TS, Kling RC, Granier S, Gmeiner P, Husbands SM, Traynor JR, Weis WI, Steyaert J, Dror RO, Kobilka BK. Structural insights into [micro]-opioid receptor activation. Nature. 2015; 524(7565):315–321. [PubMed: 26245379]
71. Watanabe Y, Fukuyoshi S, Hiratsuka M, Yamaotsu N, Hirono S, Takahashi O, Oda A. Prediction of three-dimensional structures and structural flexibilities of wild-type and mutant cytochrome P450 1A2 using molecular dynamics simulations. J Mol Graph Model. 2016; 68:48–56. [PubMed: 27371932]
72. Fukuyoshi S, Kometani M, Watanabe Y, Hiratsuka M, Yamaotsu N, Hirono S, Manabe N, Takahashi O, Oda A. Molecular Dynamics Simulations to Investigate the Influences of Amino Acid Mutations on Protein Three-Dimensional Structures of Cytochrome P450 2D6. 1, 2, 10, 14A, 51, and 62. PloS One. 2016; 11(4):e0152946. [PubMed: 27046024]
73. Cojocaru V, Winn PJ, Wade RC. The ins and outs of cytochrome P450s. Biochimica et Biophysica Acta. 2007; 1770(3):390–401. [PubMed: 16920266]
74. Teh LK, Bertilsson L. Pharmacogenomics of CYP2D6: Molecular genetics, interethnic differences and clinical importance. Drug Metab Pharmacokinet. 2012; 27(1):55–67. [PubMed: 22185816]
75. Ellingson SR, Dakshanamurthy S, Brown M, Smith JC, Baudry J. Accelerating virtual high-throughput ligand docking: Current technology and case study on a petascale supercomputer. Concurrency Comput Pract Ex. 2014; 26(6):1268–1277.
76. Mendieta-Wejebe JE, Correa-Basurto J, Trujillo-Ferrara JG, Rosales-Hernández MC. Homology modeling and molecular dynamics of CYP1A1 and CYP2B1 to explore aryl derivatives metabolism by docking, followed by experimental assays. FASEB J. 2010; 24(Suppl 1):967–8.
77. Nigsch F, Mitchell JB. How to winnow actives from inactives: Introducing molecular orthogonal sparse bigrams (MOSBs) and multiclass Winnow. J Chem Inf Model. 2008; 48(2):306–318. [PubMed: 18220378]
78. Haigh JA, Pickup BT, Grant JA, Nicholls A. Small molecule shape-fingerprints. J Chem Inf Model. 2005; 45(3):673–684. [PubMed: 15921457]
79. Muchmore SW, Souers AJ, Akritopoulou-Zanze I. The use of three-dimensional shape and electrostatic similarity searching in the identification of a melanin-concentrating hormone receptor 1 antagonist. Chem Biol Drug Des. 2006; 67(2):174–176. [PubMed: 16492165]
80. EON, version 210. OpenEye Scientific Software, Inc; Santa Fe, NM: 2011. www.eyesopen.com

81. Kirchmair J, Williamson MJ, Tyzack JD, Tan L, Bond PJ, Bender A, Glen RC. Computational prediction of metabolism: Sites, products, SAR, P450 enzyme dynamics, and mechanisms. J Chem Inf Model. 2012; 52(3):617–648. [PubMed: 22339582]
82. Wolber G, Langer T. LigandScout: 3-D pharmacophores derived from protein-bound ligands and their use as virtual screening filters. J Chem Inf Model. 2005; 45(1):160–169. [PubMed: 15667141]
83. Spartan. SGI version 4.1.1 OpenGL. Wave function Inc; Irvine, CA 92715: 1996.
84. de Groot MJ, Ackland MJ, Horne VA, Alex AA, Jones BC. 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(20):4062–4070. [PubMed: 10514276]
85. Rakers C, Schumacher F, Meinl W, Glatt H, Kleuser B, Wolber G. *In silico* prediction of human sulfotransferase 1E1 activity guided by pharmacophores from molecular dynamics simulations. J Biol Chem. 2016; 291(1):58–71. [PubMed: 26542807]
86. Lewis DF, Modi S, Dickins M. Structure–activity relationship for human cytochrome P450 substrates and inhibitors. Drug Metab Rev. 2002; 34(1–2):69–82. [PubMed: 11996013]
87. Li H, Sun J, Fan X, Sui X, Zhang L, Wang Y, He Z. Considerations and recent advances in QSAR models for cytochrome P450-mediated drug metabolism prediction. J Comput Aided Mol Des. 2008; 22(11):843–855. [PubMed: 18574695]
88. Gleeson MP, Davis AM, Chohan KK, Paine SW, Boyer S, Gavaghan CL, Arnby CH, Kankkonen C, Albertson N. Generation of *in-silico* cytochrome P450 1A2, 2C9, 2C19, 2D6, and 3A4 inhibition QSAR models. J Comput Aided Mol Des. 2007; 21(10–11):559–573. [PubMed: 18034311]
89. Yamashita F, Hashida M. *In silico* approaches for predicting ADME properties of drugs. Drug Metab Pharmacokinet. 2004; 19(5):327–338. [PubMed: 15548844]
90. Lewis DF, Dickins M. Substrate SARs in human P450s. Drug Discov Today. 2002; 7(17):918–925. [PubMed: 12546935]
91. Lewis DFV, Jacobs MN, Dickins M, Lake BG. Quantitative structure–activity relationships for inducers of cytochromes P450 and nuclear receptor ligands involved in P450 regulation within the CYP1, CYP2, CYP3 and CYP4 families. Toxicology. 2002; 176(1):51–57. [PubMed: 12062929]
92. Sridhar J, Liu J, Foroozesh M, Stevens CLK. Insights on cytochrome P450 enzymes and inhibitors obtained through QSAR studies. Molecules. 2012; 17(8):9283–9305. [PubMed: 22864238]
93. Gonzalez J, Marchand-Geneste N, Giraudel JL, Shimada T. Docking and QSAR comparative studies of polycyclic aromatic hydrocarbons and other procarcinogen interactions with cytochromes P450 1A1 and 1B1. SAR QSAR Environ Res. 2012; 23(1–2):87–109. [PubMed: 22150106]
94. Shimada T, Fujii-Kuriyama Y. Metabolic activation of polycyclic aromatic hydrocarbons to carcinogens by cytochromes P450 1A1 and1B1. Cancer Sci. 2004; 95(1):1–6. [PubMed: 14720319]
95. Roy K, Roy PP. Comparative QSAR studies of CYP1A2 inhibitor flavonoids using 2D and 3D descriptors. Chem Biol Drug Des. 2008; 72(5):370–382. [PubMed: 19012573]
96. Lapinsh M, Prusis P, Mutule I, Mutulis F, Wikberg JE. QSAR and proteo-chemometric analysis of the interaction of a series of organic compounds with melanocortin receptor subtypes. J Med Chem. 2003; 46(13):2572–2579. [PubMed: 12801221]
97. Lapins M, Worachartcheewan A, Spjuth O, Georgiev V, Prachayasittikul V, Nantasenamat C, Wikberg JE. A unified proteochemometric model for prediction of inhibition of cytochrome P450 isoforms. PloS One. 2013; 8(6):e66566. [PubMed: 23799117]
98. Wikberg, J., Eklund, M., Willighagen, E., Spjuth, O., Lapins, M., Engkvist, O., Alvarsson, J. Introduction to pharmaceutical bioinformatics. 2. Oakleaf Academic; Stockholm: 2010.
99. Wikberg, J., Spjuth, O., Eklund, M., Lapins, M. Chemoinformatics taking biology into account: Proteochemometrics. 1. John Wiley & Sons; Hoboken: 2012.
100. Apic G, Ignjatovic T, Boyer S, Russell RB. Illuminating drug discovery with biological pathways. FEBS Lett. 2005; 579(8):1872–1877. [PubMed: 15763566]

101. Drayer DE. Pharmacologically active drug metabolites: Therapeutic and toxic activities, plasma and urine data in man, accumulation in renal failure. Clin Pharmacokinet. 1976; 1(6):426–443. [PubMed: 799549]
102. Ankley GT, Bennett RS, Erickson RJ, Hoff DJ, Hornung MW, Johnson RD, Mount DR, Nichols JW, Russom CL, Schmieder PK, Serrrano JA, Tietge JE, Villeneuve DL. Adverse outcome pathways: A conceptual framework to support eco-toxicology research and risk assessment. Environ Toxicol Chem. 2010; 29(3):730–741. [PubMed: 20821501]
103. Oki NO, Nelms MD, Bell SM, Mortensen HM, Edwards SW. Accelerating adverse outcome pathway development using publicly available data sources. Curr Environ Health Rep. 2016; 3(1):53–63. [PubMed: 26809562]
104. Vinken M. The adverse outcome pathway concept: A pragmatic tool in toxicology. Toxicology. 2013; 312:158–165. [PubMed: 23978457]
105. Allen TE, Goodman JM, Gutsell S, Russell PJ. Defining molecular initiating events in the adverse outcome pathway framework for risk assessment. Chem Res Toxicol. 2014; 27(12):2100–2112. [PubMed: 25354311]
106. Russom CL, LaLone CA, Villeneuve DL, Ankley GT. Development of an adverse outcome pathway for acetylcholinesterase inhibition leading to acute mortality. Environ Toxicol Chem. 2014; 33(10):2157–2169. [PubMed: 24922588]
107. Colovic MB, Krstic DZ, Lazarevic-Pasti TD, Bondzic AM, Vasic VM. Acetylcholinesterase inhibitors: Pharmacology and toxicology. Curr Neuropharmacol. 2013; 11(3):315–335. [PubMed: 24179466]
108. Vinken M, Landesmann B, Goumenou M, Vinken S, Shah I, Jaeschke H, Willett C, Whelan M, Rogiers V. Development of an adverse outcome pathway from drug-mediated bile salt export pump inhibition to cholestatic liver injury. Toxicol Sci. 2013; 136(1):97–106. [PubMed: 23945500]
109. Phillips MB, Leonard JA, Grulke CM, Chang DT, Edwards SW, Brooks R, Goldsmith MR, El-Masri H, Tan YM. A workflow to investigate exposure and pharmacokinetic influences on high-throughput *in vitro* chemical screening based on adverse outcome pathways. Enrivon Health Perspect. 2016; 124(1):53.
110. Lee A, Ezzeldin H, Fourie J, Diasio R. Dihydropyrimidine dehydrogenase deficiency: Impact of pharmacogenetics on 5-fluorouracil therapy. Clin Adv Hematol Oncol. 2004; 2(8):527–532. [PubMed: 16163233]
111. Gross E, Busse B, Riemenschneider M, Neubauer S, Seck K, Klein HG, Kiechle M, Lordick F, Meindl A. Strong association of a common dihydropyrimidine dehydrogenase gene polymorphism with fluoropyrimidine-related toxicity in cancer patients. PloS One. 2008; 3(12):e4003. [PubMed: 19104657]
112. van Kuilenburg AB, Haasjes J, Richel DJ, Zoetekouw L, Van Lenthe H, De Abreu RA, Maring JG, Vreken P, van Gennip AH. Clinical implications of dihydropyrimidine dehydrogenase (DPD) deficiency in patients with severe 5-fluorouracil-associated toxicity: Identification of new mutations in the DPD gene. Clin Cancer Res. 2000; 6(12):4705–4712. [PubMed: 11156223]
113. Johnson MR, Wang K, Diasio RB. Profound dihydropyrimidine dehydrogenase deficiency resulting from a novel compound heterozygote genotype. Clin Cancer Res. 2002; 8(3):768–774. [PubMed: 11895907]
114. Maier CL, Duncan A, Hill CE. Pharmacogenetics in oral anti-thrombotic therapy. Clin Lab Med. 2016; 36(3):461–472. [PubMed: 27514462]
115. Kliewer SA. Nuclear receptor PXR: Discovery of a pharmaceutical anti-target. J Clin Invest. 2015; 125(4):1388–1389. [PubMed: 25831443]
116. Dressman, JB., Reppas, C. Oral Drug Absorption: Prediction and assessment. 2. Taylor & Francis Group; Boca Raton: 2010.
117. Hukkanen J, Pelkonen O, Hakkola J, Raunio H. Expression and regulation of xenobiotic-metabolizing cytochrome P450 (CYP) enzymes in human lung. Crit Rev Toxicol. 2002; 32(5): 391–411. [PubMed: 12389869]

118. Oyama T, Sugio K, Uramoto H, Kawamoto T, Kagawa N, Nadaf S, Carbone D, Yasumoto K. Cytochrome P450 expression (CYP) in non-small cell lung cancer. Front Biosci. 2007; 12:2299–2308. [PubMed: 17127240]

119. Knights KM, Rowland A, Miners JO. Renal drug metabolism in humans: The potential for drug–endobiotic interactions involving cytochrome P450 (CYP) and UDP-glucuronosyltransferase (UGT). Br J Clin Pharmacol. 2013; 76(4):587–602. [PubMed: 23362865]

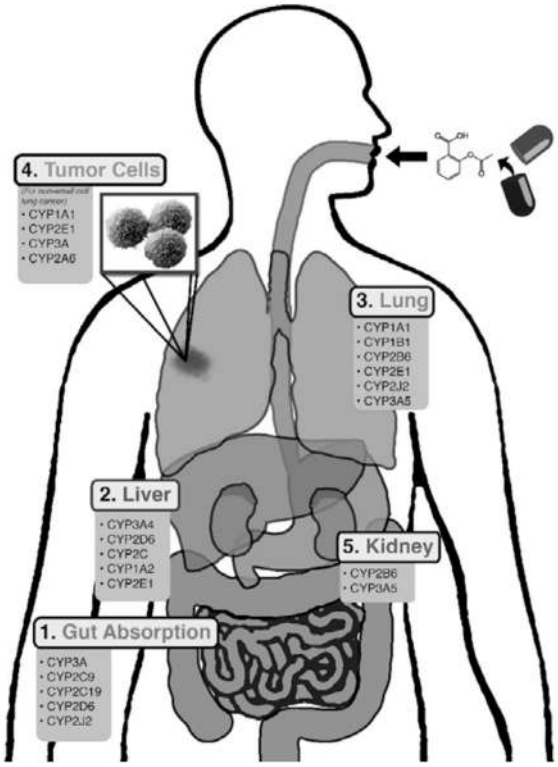

Fig. 1. Major sites of drug metabolism and associated xenobiotic-metabolizing Cytochrome P450 (CYP) isoforms when considering oral drug administration

CYPs expressed in the gut wall [116], liver [16], lung [117], tumor cell [118], and kidney [119] guide the phase 1 preclinical assessment of tissue-specific drug metabolism. Tissue-specific CYPs are not exhaustive but are presented to demonstrate inter-tissue differences and similarities in predominant CYP expression. Other xenobiotic metabolizing enzymes originating from both membraneous and cytosolic media are not included.

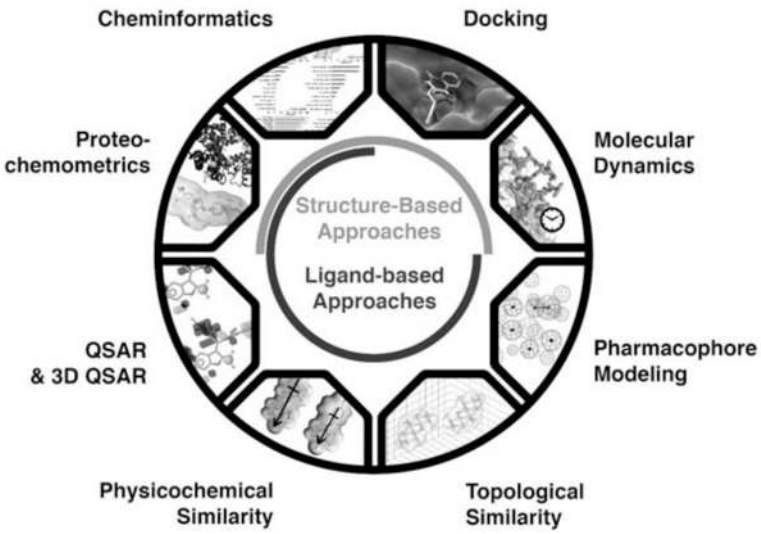

Fig. 2.
Computational approaches utilized in PKPD *in silico* studies during the pre-discovery phase.

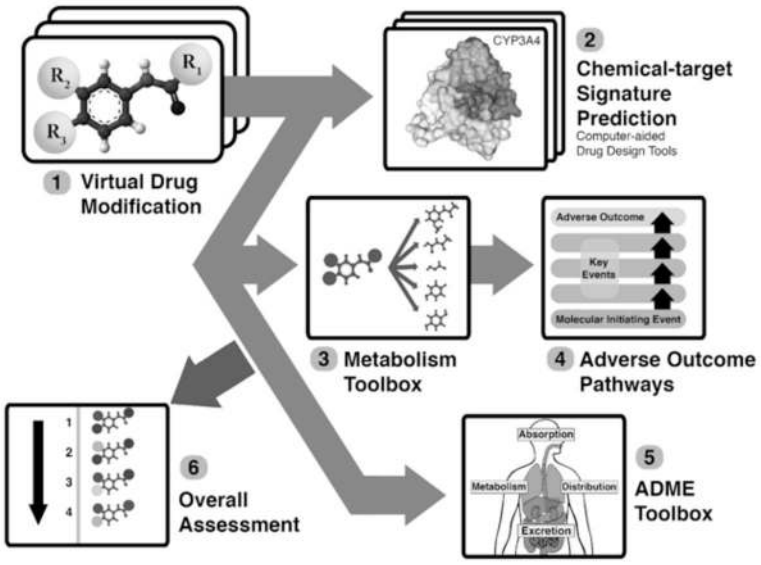

Fig. 3. Computational optimization workflow for a drug known to bind a protein target of interest in a disease state that can serve as a clinical therapeutic candidate

This workflow is intended for drug discovery instances where investigators have a drug with potential efficacy but would like to optimize the drug for ADMET. (1) Virtual modification of R groups to create a virtual library of chemical congeners for the drug. (2) Computational experiments of congener library against the protein target of interest to determine potential binding and potency. (3) Computational experiments of congeners against curated virtual library of all known metabolism-associated enzymes to predict all potential metabolites. (4) Prediction of downstream pathway perturbations of all potential metabolites for each congener that may be related to toxicity. (5) Computational experiments to determine interaction of drug with important proteins implicated in absorption, distribution and excretion such that optimal bioavailability may be determined. (6) Algorithm to rank congeners based on greatest binding potential to given therapeutic protein target, least potential for metabolism into toxic metabolites and greatest potential bioavailability depending on interested route of administration.