We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

4,500
Open access books available

118,000
International authors and editors

130M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

ADME Profiling in Drug Discovery and a New Path Paved on Silica

Arne Krüger, Vinicius Gonçalves Maltarollo, Carsten Wrenger and Thales Kronenberger

Abstract

The drug discovery and development pipeline have more and more relied on *in vitro* testing and *in silico* predictions to reduce investments and optimize lead compounds. A comprehensive set of *in vitro* assays is available to determine key parameters of absorption, distribution, metabolism, and excretion, for example, lipophilicity, solubility, and plasma stability. Such test systems aid the evaluation of the pharmacological properties of a compound and serve as surrogates before entering *in vivo* testing and clinical trials. Nowadays, computer-aided techniques are employed not just in the discovery of new lead compounds but embedded as part of the entire drug development process where the ADME profiling and big data analyses add a new layer of complexity to those systems. Herein, we give a short overview of the history of the drug development pipeline presenting state-of-the-art ADME *in vitro* assays as established in academia and industry. We will further introduce the underlying good practices and give an example of the compound development pipeline. In the next step, recent advances at *in silico* techniques will be highlighted with special emphasis on how pharmacogenomics and *in silico* PK profiling can enhance drug monitoring and individualization of drug therapy.

Keywords: ADME, drug discovery, *in silico* prediction, pharmacokinetics prediction, QSAR

1. Introduction

Drug discovery and development grew into a wide interdisciplinary field during the last decades and many factors played and play an important role in the successful evolution from a bioactive compound, or so-called new molecular entity (NME), into a potential drug [1]. Herein, we discuss the drug discovery and development (DDD) process where the pharmacokinetic profiling in terms of ADME assessment is concerned. Therefore, we provide a short overview of the *in vitro, ex vivo*, and *in vivo* state-of-the-art techniques used in academy and industry with special emphasis on how recent advances in computer science paved the path for *in silico* prediction in the DDD process for small molecules. However, the discussion of the whole topic is out of the scope of this review, which only aims to give insights into the principle process of (computer-aided) drug discovery and development.

The current state of pharmaceutical DDD estimates that only up to ten compounds out of thousand screened hits would result in optimized leads and enter preclinical testing, with a chance of 9.6% to pass the clinical testing phase [1, 2]. Additionally, the drug approval process is estimated to last in average 15 years, with major expenses in phases II and III of clinical trials, which highlights the drawback a failure in (pre-) clinical testing causes [3–6], where the overall DDD cost for each drug can reach as high as ~$2.56 billion preapproval rising to $2.87 billion including postapproval investments [6–8]. From the initial small molecule screened as hit to the optimized lead, a variety of *in vitro* tests are performed to guarantee efficacy and safety, but also to find structure-activity relationships (SAR), which can then be connected to specific physicochemical properties of the compound and further aid in the lead optimization phase [8–10].

The drug development phase starts with preclinical testing followed by the clinical stage comprising phase I–III human trials. Each of the phases aims to answer a specific question. Initially, preclinical trials are conducted in animals and can provide information about whether a drug is toxic or not. Compounds that show no toxicity in animals then advance to phase I trials, which will study whether the drug is also safe in healthy humans and provide an initial idea for appropriate dosage. In phase II, the efficacy of the drug is examined in parallel to potential side effects to answer the question if it principally meets the expected performance. Phase II presents the biggest hurdle with a transition success rate as low as 30%. Ultimately, drug candidates enter clinical phase III in which the preliminary results found so far need to be proofed and any adverse reactions monitored to make sure that the drug really helps treating the disease [2, 11].

Starting from the generation of a lead compound assessment, and optimization of pharmacokinetic properties and correlation to pharmacodynamic effects increases in importance as one of the three major attrition causes among toxicity and efficacy [8, 12]. In this sense, it is not surprising that the period between lead and the clinical candidate is sometimes referred to as "valley of death" due to the often occurring failures and dead ends during this time of the DDD process, which results in high costs and missing deadlines [13].

2. Role of computer-aided techniques in drug discovery

In a long ongoing effort, more and more *in silico* techniques are being integrated into several points of DDD with different purposes. *In silico* techniques can ease the process of SAR assessment as well as the generation of compound series by guiding combinatorial chemistry since they allow fast and easy evaluation of compounds prior to synthesis from big libraries. For instance, combinatorial chemistry offered an option to readily produce a broad range of potentially pharmaceutical active small molecules in a short time, while SAR data in combination with complex mathematical algorithms, such as regression analyses based or machine-learning–based approaches, allow to determine the potential effects of the analogues and derivative's structures *a priori* [8].

Latter approach can save time and resources by eliminating in early stages molecules that have predicted low efficacy against the target or to suggest the next round of chemical modifications [14, 15]. Still, lead generation and/or optimization will eventually also include *in vivo* testing after no toxic side effect was shown *in vitro*. *In vivo* efficacy testing will be carried out as proof of concept followed by PK assessment and ultimately animal models of human disease to find correlations between preliminary data and potential performance later on in humans [12].

In silico ADME prediction aims to generate tools and models based on experimental data to calculate *in vivo* behavior of compounds by finding quantitative structure-property relationships (QSPRs), which connect structural information to physical and chemical characteristics or even biological behavior (quantitative structure-activity relationship; QSAR). Gained empirical data are then related to descriptors/properties thereby supporting the process of hit-to-lead optimization [10, 16].

When using *in silico* methods for prediction, it is important to keep in mind that algorithms and tools applied are only models thus being only as good as the data and idea they are based on. That implies a continuous experimental validation and improvement as a basic principle that is supported by an interdisciplinary team. In this sense, frequently used models include QSPR predictors, matched molecular pair (MMP), and data trend analysis since they allow comparably easy application and are based on a high amount of (end) point data. For instance, some experiments offer highly convenient data but do not contribute much to model design, whereas others show high variability but lead to impactful models. Considering the nature of data, it is important to know which type can be used as input from different sources (low variability biological and activity data or homogeneously calculated chemical descriptors) in contrast to data that should only be used from one source (Caco-2, MDCK). A sophisticated approach to generate reliable data or to determine differences between individual experiments is to use assays with control compounds [10]. The target property must be obtained under the same experimental condition and, in the best scenario, obtained from the same laboratory aiming to avoid interlaboratory and interpersonal data noise [17].

Furthermore, the choice of the number and type of molecular descriptors has a high impact, since it influences the accuracy and interpretability of the model. One would expect that using the maximum number of descriptors would be beneficial, but in reality, the risk of overfitting the data or losing the interpretability is a trade-off. This leads to the point that it is fundamental for a "good" model to find the perfect compromise between quality and quantity. Nevertheless, it is crucial to test and train a model and to evaluate its predictability by different means, such as statistical measures and internal and external validation as recommended by organizations as OECD [18], and also includes outlier analysis to reduce the noise in the model. An extensive review of different adequate validation methods is discussed in [19].

As a result of newly achieved advances in computational capability, more complex models and algorithms can now be applied. Despite this, it is still a challenge to create a model for the pharmacokinetic and pharmacodynamic phenomena and interactions within an organism as complex as a mammal, let alone humans [10]. Finally, notwithstanding the apparent linearity, the development of a new chemical entity into a drug is an iterative process, even more, where modeling is concerned, with data from failed attempts being integrated into the new predictions [13].

3. How specific parameters shape the pharmacology studies

Pharmacology is a major part of the DDD process and describes the interaction of an organism and the drug. It can be divided into two main branches: while pharmacodynamics (PD) describes what the drug does to the body, pharmacokinetics (PK) is interested in what the body does to the drug [20]. The main processes of PK are absorption, distribution, metabolism, and excretion (ADME), finally complemented by toxicity (ADMET). While ADME tries to maximize the pharmacological performance of a small molecule, toxicology aims to ensure that it causes no harm in any kind of side effect [21].

The big hurdle to overcome is to combine appropriate physicochemical proper-ties of the drug, which would drive its interaction with the organism and show biological activity [22]. Or according to Hodgson: "A chemical cannot be a drug, no matter how active nor how specific its action, unless it is also taken appropriately into the body (absorption), distributed to the right parts of the body, metabolized in a way that does not instantly remove its activity, and eliminated in a suitable manner—a drug must get in, move about, hang around, and then get out" [21].

As already reviewed [23] and suggested by the FDA, PK/PD assessment is one of the main focuses for optimization in the drug development process. This is apparently an idea that was shared among many: whereas ADME evaluation was previously addressed in the late stages of preclinical development, currently it became a major concern throughout the whole DDD process, starting from the very beginnings in drug discovery approaches until the very last steps in lead optimization [21, 24].

For each step in the drug's path through the body, several parameters determine the destination of the drug. In respect to this, each of those parameters would be addressed directly and individually. Unfortunately, to address experimentally each potential parameter is timely unviable, due to the complexity of the human body where all those parameters influence each other. This is not only restricted to mecha-nisms within the body between different compartments but also extends to interper-sonal variations introduced through gender, age, genetic state, disease, etc. To find an approximation, most of the important variables are indirectly evaluated by either models or surrogates (**Table 1**). In an approach to characterize the properties of compounds, facilitate calculations, and allow standardization between experiments, descriptors are introduced as numerical representations encoding aspects of the chemical information of a molecule. Examples of descriptors and properties include molecular weight and H-bond donors/acceptors and they can be directly obtained from experimental or generated by computational techniques [25].

	In vitro	*Ex vivo/cells*	*In vivo*	*In silico*
Absorption	Physicochemical properties Dissolution and solubility Cell monolayers Artificial membranes	MDCK Caco-2 Transfected cells	Mouse model Knock-out/ down mice Humanized mice	QSPR/QSAR pKa log$P$, log$D$ Binding and expression of transporters Inhibition of efflux pumps
Distribution	PAMPA IAM HAS-coupled (RP-) HPLC	Plasma/tissue binding	Vd Plasma protein binding	Vd Plasma protein binding Activity and expression of transporters
Metabolism	S9 fraction Liver microsomes Recombinant enzymes	Hepatocytes (HepG2) Isolated tissue Isolated organ	Humanized animals	Half-life Activity and expression of transporters
Excretion		Isolated tissue Isolated organ Transfected cells	Urine analysis	Half-life prediction

Table 1.
Tools for ADME evaluation.

For instance, although the perfect approach of PK profiling would also reflect the kinetics of drug administration and concentration at the site of action, most *in vivo* systems rely on plasma sampling as a medium of drug equilibrium since it is easily accessible. As a consequence, results are highly influenced by intrinsic and extrinsic factors such as interpersonal variances as already stated above [20].

Each compound possesses individual physicochemical properties, such as solubility or lipophilicity, which are influenced by biochemical properties of the body as the different pH of tissues. Although they can be similar, each compound will behave differently, and it is futile to address *in vivo* behavior without any preliminary knowledge of the basic PK parameters *in vitro* [26].

Furthermore, every PK assessment varies depending on the route of administration and requires different models and assays. While some routes depend on absorption mechanisms like oral and transdermal administration, others (i.e., intravenous) directly target the bloodstream and the bioavailability is essentially equal to 100%. Hereafter, we will discuss oral administration parameters of small molecules as the most common form due to many advantages like reliability, safety, price, their experimental approaches, and most common prediction modes [27, 28].

Passive transport across membranes is defined as permeability, which is dependent on lipophilicity, since biological membranes are virtually lipid bilayers, and is by far the most important transport for small molecules, especially in oral absorption [8, 24, 29]. Nonlipophilic compounds normally do not traverse membranes passively, while highly lipophilic molecules run the risk to get stuck within the membranes [30].

Properties utilized for measuring lipophilicity are the logarithm of the partition coefficient (log$P$) and the distribution coefficient (log$D$) with the first not differentiating between ionized and nonionized species. Both are normally applied for *n*-octanol/water representing an organic and aqueous phase, respectively [21, 26].

Ionizability and lipophilicity provide a strong indication if a compound is likely to be orally absorbed or not [21].

Ultimately, also the molecular size of the compound is involved in successful absorption due to the aforementioned effects on permeability and solubility [31]. Usually, increasing molecular weight by adding new chemical moieties leads to decreased solubility in aqueous solutions [32] and while big lipophilic compounds partition passively along membranes (transcellular), small charged molecules can also cross membranes via tight junctions (paracellular) [26]. For oral absorption in terms of permeability, Lipinski and collaborators already proposed in 1997 [33, 34] that orally active compounds should fit at least three of observed four parameters: molecular weight < 500 g mol^{-1}, log$P$ < 5; number of hydrogen bond acceptors <10; number of hydrogen bond donors <10; the well-known Lipinski's rule of 5 (Ro5). In other words, Ro5 stated a physicochemical space in which molecules outside its domain has a low probability to become orally active. Other rules, as Veber rules [35], Daina and Zoete [36], Egan and collaborators [37], Lovering et al. [38], and Ritchie and colleagues' [39] works for example, also included other properties as the sum of hydrogen bond acceptor and donors, rotatable bonds count, polar surface area, number of aromatic rings, and fraction of sp^3 carbon atoms.

Despite the criticism and overinterpretation of Lipinski and derived rules, the influence of physicochemical parameters on oral bioavailability and related parameters (as log$P$ and aqueous solubility) is notable. Moreover, these rules are still being employed nowadays in virtual screening campaigns aiming to reduce the number of compounds from massively large available libraries (*e.g.*, ZINC, which contains more than 750 millions of compounds) [40–42]. Furthermore, those initial steps instigate the generation of more complex models to predict not just oral bioavailability but other PK-related parameters as Caco-2 permeability, aqueous solubility,

and log*P* as indirectly related properties as well as other direct parameters as intestinal absorption, metabolism, clearance, etc.

3.1 Aqueous solubility and lipophilicity

As already mentioned, ionizability is one of the most important properties in PK, thus making pKa the physicochemical property with the highest impact.

Early attempts to increase the efficiency of pKa evaluation were reported by Morgan and colleagues by scaling down the classical titration and spectrophotometric methods introducing microscale versions [43].

These alterations, however, could not overcome the principle demands of each technique, which are moderate precision and frequent calibration (potentiometric), and the need for a chromophore within the analyte (spectrophotometric) [44]. Starting in 1998, capillary electrophoresis (CE) was effectively used to determine pKa of many compounds and was further upgraded from Pfizer by implementing pressure-assisted capillary electrophoresis (PACE) as a standard method, which is nowadays readily applied in industry settings showing superior features compared to the aforementioned methods [44, 45]. Other variants such as vacuum-assisted multiplexed capillary electrophoresis also exist (VAMCE) [46]. A different approach better suited for HTS is called pH gradient titration offered from Sirius Analytical Instruments but is still limited due to the UV spectroscopy technology [30].

It is well established that solubility in aqueous media is one of the most important physicochemical properties to be evaluated in oral administration. It is not only necessary for absorption in the GI tract but also a requirement for almost all *in vitro* and *in vivo* assays, which depend on a solved compound. Poor solubility can affect the reproducibility of assay results by introducing high variability and further increase development costs of leads with low solubility [26, 47]. Traditionally, solubility measurements were conducted via labor-intensive potentiometric techniques [48] or equilibrium solubility (thermodynamic; *e.g.*, shake flask) [26]. HTS alternatives comprise laser nephelometric scans (kinetic) [47] and LC-MS/HPLC techniques, which can also be performed with DMSO solutions of the compound—the standard for HTS applications [47, 49]. It should be noted, though, that aqueous solubility, as described above, is not an optimal model for GI solubility since it does not consider the composition of the GI fluids [49].

On the other hand, generally speaking, lipophilicity is the ability of a compound to dissolve in lipids and/or organic solvents thus being able to pass biological membranes. Descriptors for lipophilicity are the logarithm of the partition coefficient (log*P*) or distribution coefficient (log*D*). Classically, log*P* was determined using the shake flask method applying *n*-octanol/water phases. Later, UV spectroscopy became the standard, which unfortunately is not applicable for compounds without absorption in the UV range [50]. Today, RP-HPLC methods are frequently in use due to superior properties [25, 51]. As with many methods, comparison of results obtained under different conditions and in different laboratories proves to be difficult with RP-HPLC. A solution offers the implementation of a standardized lipophilicity value, for example, the chromatographic hydrophobicity index (CHI).

In recent years, a great effort has been made to improve the ability of *in silico* models to accurately predict aqueous solubility. One of the most developed model is Yalkowsky and Jain's [52] general solubility equation (GSE), which is based on the melting point (m.p. °C − 25) and log*P* (the octanol-water partition coefficient of the un-ionized molecule) of a chemical substance (Eq. (1)), with a relevant prediction power as represented by the coefficient of determination (R^2) = 0.96 and root-mean-square error (RMSE) = 0.53 in a dataset of 1026 organic compounds [53].

General solubility equation as proposed by Yalkowsky and Jain's [52]:

$$logS = 0.5 - 0.01(\text{m.p.}°C - 25) - logP \qquad (1)$$

Modifications in terms of the GSE have been proposed, for instance with the SCRATCH model, which replaces the melting point by molar aqueous activity coefficient, with comparable accuracy (R^2 = 0.956, RMSE = 0.859 in a dataset of 883 compounds) [54]. Ali and collaborators suggested replacing the melting point descriptor of the GSE with TPSA, aiming to overcome the issues with compounds with high melting points and also to explicitly take into account the effect of polar and polarizable atoms on the aqueous solubility [55].

The argument that real drugs are actually more soluble than drug-like molecules, filtered by Lipinski's rule of five [56], pointed out the studies in the direction of more complex models. Indeed, nowadays, the quantitative structure-property relationship (QSPR) models correlating the aqueous solubility with various molecular descriptors are often employed. As an example, Chevillard et al. reported the use of a random forest protocol to select the most accurate model among several available, both in commercial or free software packages, for each compound [57]. They report that the multimodel approach can enlarge the applicability domain given that more accurate results for solubility prediction were obtained in comparison to using individual models. This approach agrees with other reports that consensus of local QSAR models can generate predictive workflows, especially for datasets with large structural diversity [58, 59]. It is worth noting that Lipinski himself recently revisited his own rules [60], in vision of new potential classes of drugs, such as natural products, peptide-like, and fragments, which, despite the validated effect, would defy the original Ro5 limits.

3.2 Ionization state and pKa prediction

Early pKa measurement proves beneficial in lipophilicity assessment since log*D* values at any pH can be calculated from the pKa and log*P* values [25, 50]. Although octanol/water log*P* is similar to most components in the body, not all biological partition processes (i.e., blood-brain barrier and gastrointestinal absorption) can be easily modeled by it [25].

The prediction of ionization state of compounds, which is indicated by the pKa value, is relevant to derive several other physicochemical and ADME properties of drugs, including solubility, lipophilicity, and pharmacokinetic profile. The use of pKa prediction can be placed in two different stages along the DDD, in the beginning with fast models for larger libraries, intending to generate all possible state populations of particular compounds, and/or later on with more refined semiempirical and, computationally expensive, the density functional theory (DFT), in which more accurate ionization states can be accessed. Examples of fast prediction methods for ionization states, which are available as computer programs, are SPARC [61], MoKa [62], and Epik, which use the Hammett and Taft approaches for the pKa prediction [63]. On the other hand, once smaller subsets of molecules are being addressed, the use of semiempirical or density functional theory (DFT) with more computationally expensive models was reported to accurately incorporate the structural features and diversity into the pKa prediction [64, 65].

3.3 Permeability and the use of cellular and noncellular models

As already seen, lipophilicity (log*P*, log*D*) is highly involved in membrane permeability. Apart from the already described *in vitro* methods for log*P* and log*D* determination, systems for *ex vivo/in situ* but also *in vivo* assessment exist as "direct empirical" determination of permeability. When talking about permeability, the

difference between passive diffusion and transporter-mediated active transport needs to be considered.

Cell culture methods have been applied to study intestinal absorption for several decades already [66]. Finding the correct model or cell line is crucial to assess the desired parameters such as passive or active transport. In general, it cannot be distinguished between the different transport mechanisms when using cell culture approaches, but several models exist to shift the focus on one of the parameters.

Two main cell lines are in use as models for intestinal absorption: Caco-2 and MDCK cells. Caco-2 cells are derived from a human colorectal carcinoma and possess many of the typical properties of the small intestine, therefore representing a well-established and validated assay system for absorption, permeability, and secretion studies [21, 67]. This assay is mainly used for rank ordering of compounds in terms of oral absorption and permeability in early phases of drug design. Unfortunately, results obtained in different batches and laboratories vary heavily due to several reasons, which make control compound usage necessary and represent a drawback of the technique [26]. Additional disadvantages include long preparation times (about 3 weeks) and no specific permeation mechanism evaluation. Caco-2 assays are usually used as a primary assay followed and complemented by other *in vitro* and *ex vivo* methods [68]. Recently, a 3D version of the Caco-2 assay, "Caco-2 3D spheroid permeability assay" was reported, increasing the overall performance and correlation to a human *in vivo* data [69].

As already stated above, transcellular permeation either occurs passively via diffusion of lipophilic molecules or is driven by membrane transporters. Important transporter includes ATP-dependent efflux transporter such as MRP2, BCRP, and P-gp and the organic solute transporter and the multidrug resistance protein 3 (MRP3) on the luminal and basolateral membranes, respectively [26].

Madin-Darby canine kidney (MDCK) cells are an alternative to Caco-2 cell-based assays and the next most common cell line for passive permeability assessment as well as drug-receptor interaction [70]. MDCK cells also are ideal for transfection and overexpression experiments with human transporters and receptors due to the lack of P-glycoprotein [68, 71]. For instance, the MDCK-MDR1 cell line overexpresses the multidrug resistance protein 1 (MDR1, P-glycoprotein) and can be used in concert with other cell-based assays to specifically address the influence of MDR1 in drug efflux [72].

Immobilized artificial membranes (IAMs) were already used very early on for lipophilicity determination and are gaining interest again in recent years for direct permeability measures [25]. IAMs are also intensively used in the measurement of the volume of distribution to mimic *in vivo* binding to phospholipids and phospholipid bilayers (membranes). Therefore, IAMs are discussed in more detail in the following section.

The parallel artificial membrane permeability assay (PAMPA) [73] is a cheap and fast *in vitro* alternative to cellular-based assay systems. A very comprehensive review of recent PAMPA methodologies and applications is available [74]. In principle, PAMPA was developed to overcome cellular-based systems (Caco-2, MDCK, etc.) for passive permeability evaluation, which are error-prone, more difficult, and labor and time intensive and tend to report false negatives. Another advantage of PAMPA over conventional cell-based assays is the ability to selectively measure passive permeability, while in cell-based systems influence of membrane transporters cannot be left out. PAMPA assays can be readily applied in high throughput processes or scales and different variants exist to address ionic and H-bonding with membranes that influence permeability and complement the use of Caco-2 and other cell assays [73]. Bermejo and colleagues also showed a significant correlation between Caco-2, *in situ* rat perfusion, and PAMPA assay data underlining applicability of the method for ADME assessment [75].

Although high-throughput applications of newly developed and standardized techniques allow gathering of an exorbitant amount of data, it is crucial to also (cor)relate the physicochemical and biomimetic properties to structural features of the compound. This will facilitate the development of QSPRs and allows the construction of *in silico* models ultimately guiding the medicinal chemistry efforts [25].

When dealing with oral administration, it is important to note that the drug is not only confronted with the hurdles of solubility and permeability in the absorption process but is also facing metabolizing mechanisms (*i.e.*, enzymes) in the gastrointestinal tract, which are referred to as first-pass metabolism [24, 76]. These include but are not limited to P-glycoproteins, uridine diphosphate glucuronosyltransferase, and mainly cytochrome P450s (CYP450) [24]. This will be discussed more deeply in the Metabolism section.

Permeability has a direct influence on the drug absorption rate and, as discussed, despite the several *in vitro* cellular models available (*e.g.*, Caco-2, PAMPA, and MDCK), the high costs justify the use of *in silico* prediction. Further, QSPR study developed using a large compound dataset of Caco-2 permeability data (1272 compounds) presented good apparent permeability prediction accuracy ($R^2 = 0.81$ for the test set) using the polar volume, number hydrogen bond donors, and the surface area as main descriptors [77].

However, we are far from a model that can predict overall permeability and, the current status, rather focuses on individual compartments and tissues, such as the gastrointestinal (GI) tract, skin, buccal membrane, and the blood-brain barrier (BBB). Since the first BBB permeability correlations with log*P* in 1977 [78], models to predict BBB permeability, particularly log*BB* (Eq. (2)), have greatly advanced. Current models using an array of machine-learning methods such as multilinear regression, support vector machine (SVM), and artificial neural network (ANN) against a dataset of 320 unique compounds had good predictive power ($R^2 = 0.89$) [79]. The work of Shen *et al*. developed SVM models using 1593 compounds (1283 BBB[+] and 310 BBB[−]) by using different pattern selection methods and obtained the overall accuracy of 98.2% [80]. Both methods have the limitation of unbalanced datasets (where the number of BBB+ is higher than the BBB− within the training set), which was addressed on the work of Wang et al. by using resampling methods coupled with the machine-learning techniques, to achieve accuracy rates of 0.919 in external test data [81]. Wang and collaborators compiled a dataset of 439 unique molecules, which were employed to generate a diverse set of QSAR models and consensus ($R^2 = 0.504$ for external dataset prediction). They also reported the use of transporter profiles as additional biological descriptors to develop hybrid QSAR BBB models, with an improved correlation coefficient $R^2 = 0.526$ for external validation [82].

Log*BB* can be calculated by the log of the ratio between the concentration within the brain (C_{brain}) by the bloodstream concentration (C_{blood}) of the determined chemical entitiy

$$logBB = \log\left(\frac{C_{brain}}{C_{blood}}\right) \qquad (2)$$

Finally, beyond the usual ADME parameters of interest in DDD, there are several other unusual ones that also can be predicted; as examples, we here point the permeability of the models for skin permeability, which evolved from simple diffusion models based on molecular weight and *n*-octanol/water partition coefficient [83, 84], until more sophisticated models, such as (non)-linear QSPR models and even molecular dynamics simulation (as extensively reviewed by [85, 86]).

4. ADME properties: experimental approaches and *in silico* models

4.1 Absorption

Oral bioavailability is defined as the amount of drug that reaches the site of action after oral administration and is influenced by factors like drug solubility and dissolution, chemical and enzymatic stability in the gastric and intestinal lumen, interacting luminal contents (food), gastrointestinal transit time, enterocyte permeability, and intestinal and hepatic metabolism [24]. Recently, bioavailability has been also described as the rate and speed of the drug to reach systemic bloodstream, considering the initial formulation as the starting point.

Oral administration includes a pharmaceutical phase—prior to PK and PD phases—that comprises disintegration and dissolution of the dosage form. When using oral dosage forms, the shape and chemical composition (*e.g.*, tablets) play an important role since they contribute to the time needed for disintegration and dissolution.

Following the pharmaceutical phase, absorption is the first step in the pharmacokinetic phase and is defined as the movement of the drug from the site of administration to the bloodstream. The main properties determining the rate of oral absorption for small molecules are permeability and solubility [87].

As such, the rate of dissolution and ionization, which are described by the Noyes-Whitney and Henderson-Hasselbalch equation, respectively, is the key factors in lead optimization for oral administration and is complemented by lipophilicity as an additional factor influencing membrane permeation and solubility of the compound [31].

Dissolution can be expressed by a function of the aqueous solubility of a compound, the surface area of the administered tablet (or the particles in other solid formulation), and a specific dissolution rate constant. Altering any of these parameters directly affects the dissolution profile [26]. While solubility is an endpoint value indicating the amount of a compound that is soluble in a solvent, dissolution describes the kinetic process of a compound being solved in a solvent [88].

On the other hand, ionization reflects if a compound is present in the charged or uncharged state and is at least influenced by two major parameters. The physicochemical property responsible for ionization is the pKa and describes the ionization state of that entity at a given pH. It is also referred to as aqueous ionization constant [30]. Thereby, it is directly influenced by the pH of the environment, the second parameter, which drastically changes on the way through the GI tract, from about pH 1 to 8 in the stomach and ileum, respectively.

The determination of the ionization state of a compound in the gastrointestinal system (stomach, jejunum, ileum, and colon) is crucial for absorption since it not only influences the solubility of a compound but also the lipophilicity and permeability [26, 30, 89]. About 60–70% of all drugs (effective 1999) are ionizable, which underlines the role that ionization plays in ADME assessment [30, 90]. While charged molecules easily dissolve in aqueous systems (GI tract), they do not permeate membranes via passive diffusion and are reliant on active transport. The contrary is true for uncharged molecules, which pass biological membranes passively but show low solubility in aqueous solutions. Mechanisms of drug absorption include passive diffusion, active transport, and receptor-mediated endocytosis, which are influenced by different factors and can themselves influence the bioavailability.

Similar to model and prediction, the absorption of a drug is a complex process, which is influenced not only by the physicochemical properties of drugs themselves but also by the physiological state of the tissue in question. As such, there are a large

number of prediction models available, which were generated based on the physico-chemical properties involved in the absorption process, such as membrane perme-ability and drug solubility. These models can help formulation scientists to optimize drugs with poor absorption due to low aqueous solubility.

Initial absorption models can be separated into dispersion and compartmental models [91]. While dispersion models treat the GI as a continuous system, with variable pH and surface area, compartmental models take into account physi-ological factors such as transporters. The compartmental absorption transit (CAT) was one of the first models to regard distinct physiological properties, such as the minimal absorption in the stomach and colon, while assuming some mathemati-cal simplifications, such as the instant dissolution of the drug and linear kinetics [92]. CAT was further modified as advanced CAT (ACAT), by including nonlinear absorption kinetics and the effects of the first-pass metabolism. ACAT also consid-ers the gastrointestinal tract as nine subsections, each with unique physicochemical properties, such as pH, allowed solubility, particle size, and permeability [93]. Novel developments have included other absorption routes other than the GI, which have been recently included in commercially available software, such as oral absorption for the development of sublingual zolpidem tablets [94]. The absorption constant (K_a, expressed in terms of $h^{-1} min^{-1}$), or also called first-order absorp-tion rate constant (to not be confounded with pKa), is employed in most of the aforementioned models and is determined as a result from the changes in mass of absorbable drug over time at the site of administration. K_a can be derived from the decrease in the drug amount of absorbable present at the site of administration over time; however, it is often indirectly determined by the drug amounts measured in the blood and/or urine.

Along physicochemical models, which have a global application, machine-learn-ing techniques were extensively employed to model absorption (as comprehensively reviewed by Kumar et al. [95]) and are inclined to be local models, since they are mostly based on a small, homogeneous dataset that influences their applicability domain.

4.2 Distribution and the role of plasma-binding proteins

After being absorbed and entering the circulatory system, the drug moves reversibly between different compartments within the body, which is described as distribution and influenced by several physicochemical properties of the drug and biological factors of the body. One of the most important properties is lipophilicity, and as such log$P$/log$D$, since it reflects the ability of the compound to pass biological membranes to reach other sites, tissues, and organs within the body [25]. Additional factors include phospholipid and (plasma) protein binding, which reduces the free drug concentration within the body, can prevent the migra-tion to the receptor side/side of action, and causes drug-drug interactions [25, 96]. Interestingly, binding to plasma proteins can also prolong the drug action by releasing the drug over a longer period of time. It is also important to note that the influence of lipophilicity on plasma protein binding is hypothesized to be higher for acidic compounds than for bases, meaning that negative charges contribute highly to plasma protein binding and prevent tissue binding, which leads to diminished volumes of distribution (V_d, Eq. (3)). The V_d is the amount of drug that is freely available in the blood, thus not bound to plasma proteins or other components [25, 97, 98]. V_d is an apparent volume that increases proportionally to the extravascular drug binding and not an anatomically defined volume. Consequently, exten-sive drug binding outside the bloodstream leads to increasing values of volume distribution.

Volume distribution (V_d) is defined by the ratio between the amount of drug in the body (A) and the drug concentration in plasma (C, comprising both free drug and protein-bound drug):

$$Vd = \frac{A}{C} \qquad (3)$$

The parameter describing protein binding is the plasma protein affinity constant K_i. Many efforts to determine distribution led to chromatography-based methods, such as (RP-)HPLC to mimic *n*-octanol/water log*P* or lipophilicity to measure distribution. In general, chromatographic methods are believed to resemble biological partition processes more than octanol/water partition [25]. In the beginning, stationary phases in (RP-)HPLC were either silica-based or polymer-based but both had difficulties to reproduce log*P* and log*D* values despite several additives in the mobile phases [99]. The introduction of biomimetic (stationary) phases coated with human serum albumin (HAS), α_1 acid glycoprotein (AGP), or immobilized artificial membranes (IAM) revolutionized the methodology since they allowed a better approximation of the biological system [25, 100].

A method to address plasma protein binding is the use of HSA and other plasma proteins (*e.g.*, α_1 acid glycoprotein) coupled with RP-HPLC [25, 101]. On the other hand, HPLC combined with IAMs is a popularly accepted technique for phospholipid interaction and partition and several IAM columns are commercially available for DDD projects. Both techniques represent good assay systems to model *in vivo* V_d in high-throughput scale [98]. Problems with HPLC techniques, which are also true for biomimetic phases, include the lack of a gold standard that is needed to calibrate and later standardize results to make a comparison possible [25].

In vitro standard methods for unbound plasma fraction calculation include equilibrium dialysis and ultrafiltration among several others as the two most commonly used methods and are considered the gold standard for binding assessment [26].

To calculate the V_d "a priori"/nonexperimentally, plasma protein binding, experimental log*D* and pKa are necessary. Then again, based on the Vd, the half-life ($t_{1/2}$) of a compound can be calculated [102]. Apart from protein binding, tissue binding is also involved in the distribution of the compound. Generally, "tissue" here comprises several components of the human body such as lipids, DNA, or RNA and is also referred to as nonspecific binding [26].

In silico models to predict the V_d are often based on lipophilicity and solubility descriptors, which correlate with the fractions of the drugs that are either bound to plasma proteins or freely available. The work of del Amo et al. not just accurately predicts V_d and unbound drug fraction but also compares the model's performance against the commercially available software VolSurf+ with comparable accuracy ($R^2 = 0.70$ and 0.71, respectively) [103].

Expanding these studies, the work of Lombardo and Jing generated a set of models to predict the V_d in the steady state (V_{ss}), using a dataset of 1096 diverse compounds [104]. They compared models generated by linear (PLS) with nonlinear (Random Forest) models, recommending the latter, with 33 descriptors, as the optimal method for V_{ss} prediction.

The V_d of drugs is greatly influenced by binding to plasma proteins with several machine-learning–based models generated to predict this interaction. Protein-protein interaction (PPI) information derived from molecular docking was employed to derive a PPI-QSAR model for a small dataset of antibiotics (65 unique compounds), which resulted in an accurate model ($R^2 = 0.86$ for the test set) [105]. Additionally, global quantitative models using an array of classification and regression models using physicochemical and molecular descriptors derived from

a dataset of 794 compounds were shown to correctly classify the binding status of the test set compounds and could be used as a prescreening [106]. Another recent QSAR study using an extensively curated training set of 967 diverse pharmaceuticals aimed to predict plasma protein-bound fractions (fb) using models generated by six machine-learning algorithms with 26 molecular descriptors [107]. This study is particularly interesting where the applicability domain is concerned allowing to differentiate whether the classification derives from (un-)favorable regions.

del Amo et al. recently reported one of the first QSPR models to predict intravitreal volume of distribution and clearance of small molecules [108]; the model relies on the Log*D* and hydrogen bond capacity to understand phenomena such as intraocular pressure and guide drug discovery. Complementarily, the prediction of the drug passage through the blood-ocular barrier was described to be an important factor to evaluate volume distribution in this organ [109].

Recently, as a novel approach bridging the animal experiments with human results, it was shown that in PXB mice, a chimeric mice linage with a humanized liver, plasma concentration-time profiles could be used to infer human's compound half-life [110].

Volume of distribution is also closely related to half-life and clearance parameters. As the V_d is a relative measurement of the free concentration of drug in the blood, this same amount could be excreted by kidneys in the glomerular filtration (clearance). Consequently, the rate of clearance (discussed below in Excretion section) directly influences the amount of available drug. Naturally, the concentration of free drug that can bind its molecular target is related to the therapeutic dosage and the half-life of the administered drug (as seen in Eq. (4)).

Half-life definition. Half-life is calculated by a ratio between the Napierian logarithm multiplied by the volume of distribution (V_d) and renal clearance (CL):

$$t1/2 = \frac{Ln2.Vd}{CL} \tag{4}$$

4.3 Metabolism

Drug metabolism normally involves enzymatic modification or degradation of the compound to facilitate excretion via one of the major clearance organs: liver, kidney, spleen, or bile. While phase I enzymatic reactions include modifications such as oxidation, hydrolysis, and reduction to either introduce a functional group to the molecule or make it accessible, phase II reactions are conjugation mechanisms (*e.g.*, methylation, acetylation, glutathione conjugation, amino acid conjugation, and others) that result in polar products that can be actively effluxed [26]. Thus, isozymes of the CYP450 family and efflux transporters such as P-glycoprotein and members of the multidrug resistance transporter MRP family are highly involved in the metabolism of drugs as well as drug-drug interactions, which are a major attrition cause. For instance, CYP3A4, CYP2C9, and CYP2D6 together catalyze the hepatic metabolism of about 50% of drugs, which underlines the importance of the superfamily. Interestingly, when CYP3A4 is expressed, usually P-glycoprotein is as well [8, 10, 14, 24, 111]. An approximation for metabolic behavior analysis is the use of either liver microsomes or S9 fractions although also recombinantly expressed proteins are partially in use [24, 26].

When available, the 3D structure of those proteins could be employed in molecular docking and molecular dynamics simulations aiming to predict the binding affinity of drugs or drug candidates aiming the estimation of a PK profile [112]. The metabolism prediction combines mathematical models to predict whether

the target compound could be a substrate of a specific enzyme in combination with metabolism site predictions. Usually, those initial predictions are followed by molecular docking simulations and quantum mechanics simulations due to the dependency of electronics structure from both substrate and enzyme in catalyzed reaction [113, 114].

Nowadays, several attempts have been made to develop *in silico* models for predicting drug metabolism, specifically site-of-metabolism (SOM), and quite often are also converted into online server prediction tools for general use, for instance, the FAst MEtabolizer (FAME) model, which was generated from a diverse chemical datasets of more than 20,000 molecules and their respective experimentally determined metabolism sites. FAME prediction rates were comparable to other metabolism site predictors focused on specific enzyme families, despite using only seven chemical descriptors [115]. Similarly, SMARTCyp server, which initially relied on the 2D structure of the molecule, without considering electronic properties or generating 3D structures, to predict CYP2D6 [116], was later expanded for other CYP isoforms. A more refined version was later updated to include the atomic solvent accessible surface area, which is independent of 3D coordinates, slightingly improving the overall prediction accuracy for different CYP isoforms [117]. The newest SMARTCyp version (3.0) uses the activation energies calculated by the density functional theory (DFT), meaning the difference between the energy of the transition state and the reactant complex, to predict SOMs of molecular fragments of the query in an unsupervised fashion. SMARTCyp 3.0 also calculates the similarity between the query molecule and the model fragment [118].

IDSite approach aims to overcome the ligand-based bias of SOM prediction by using it as a part of a large framework, more precisely by combining it with molecular docking, where an atom can be considered a significant SOM by a P450 enzyme when accessible to the reactive heme iron center, and/or quantum calculations, where the candidate atom must have some degree of reactivity in the absence of the enzyme [119]. Similarly, the work of Kingsley et al. combined different approaches into a framework to predict CYP2C9 substrates. They validated the predictions from SMARTCyp in an ensemble docking, followed by a QSAR model to account for influences of both the inherent reactivity of each atom and the physical structure of the CYP2C9 binding site [120]. This combined approach resulted in 88% of true SOMs accurately predicted among the top ranked sites.

4.4 Excretion

Excretion is guided by one of the major clearance organs, and the assessment of clearance behavior sometimes involves isolated organs or tissues [24]. Humans rely on the kidney clearance as a major route for xenobiotic excretion, despite other available routes such as feces, bile, sweat, and breath. The excretion pathways directly impact the concentration of available drugs and are often measured in terms of half-life and the initial administered dose.

The renal clearance of a drug is another important parameter, which is usually employed to predict drug excretion. Experimentally, clearance is defined by the drug concentration drug along a defined time of renal excretion by a linear equation (Eq. (5)).

Equation for renal clearance. m is the substance's mass generation rate, K is the clearance and C is the concentration at the time, and V is the volume where the drug is distributed, or for systemic approaches the total body water.

$$V.\frac{dC}{dt} = -K.C + m \qquad (5)$$

Gombar et al. developed SVM- and MLR-based QSAR models to predict both systemic clearance and apparent volume of distribution from intravenous data [121] using as input structural fingerprints and electro-topological states (so-called E-states), respectively. The model performed with high accuracy, despite the highly diverse initial dataset employed for its generation, which points the importance of those models in early steps of the drug-discovery pipeline.

Also, the work of Kusama et al. established a chemoinformatic-based classification model to predict the major clearance pathways of 141 approved drugs based on four physicochemical parameters: charge, molecular weight, lipophilicity, and protein unbound fraction in plasma, resulting in a final model with an accuracy of 88% [122]. This model approach was further refined by using support vector machine and increasing the number of relevant descriptors [123]. In order to better model the biotransformation processes, often the major triggers of excretion, the work of Berellini et al. used ELASTICO (Enhanced Leave Analog-Structural, Therapeutic, Ionization Class Out) to provide an appropriate sampling during the validation process. Their partial least-square models resulted in a highly accurate model derived from 754 compounds [124].

On another topic, ABCB1, also known as P-glycoprotein (P-gp or MDR1), is a membrane protein member of the ATP-binding cassette (ABC) transporters superfamily. Together with the hERG channel and CYP3A4, P-gp is one of the most widely studied antitarget, where its inhibition could bring consequences for several processes, such as the absorption, distribution, and excretion of drugs. Classical studies used chemometric methods to describe bioavailability in terms of P-gp and CYP enzyme activities, generating QSAR models based on 805 unique drug molecules with high accuracy ($R^2 = 0.80$ for the test set) [125]. Alternatively, an approach to predict P-glycoprotein inhibition using molecular interaction fields, derived from a literature collection of more than 1200 structures, generated a pharmacophore model for competitive P-gp inhibition [126].

The most recent reported studies involving prediction of drug clearance, both from human and rat hepatic *in vitro* systems, were based on microsomes, with a recent emphasis on the use of hepatocytes. Wood et al. discuss the inherent limitation of using human hepatocyte predictions, due to underprediction when compared to *in vivo* clearance data, and the comments on the potential causes for those divergences [127].

As the pinnacle of ADME *in silico* approaches, the holistic physiologically based pharmacokinetic (PBPK) modeling was initially conceptualized by Teorell [128], aiming to enable the prediction of drugs' pharmacokinetic behavior in humans using preclinical data. Recent PBPK models benefit from the large amount of available ADME data not only to aid the drug discovery process and dose regiment selection but also to guide the risk assessment for regulatory reviews [129]. PBPK models are compartmentalized representations of the different organs, and each compartment can be described by a specific tissue volume and blood flow rate, which communicates with the blood (venous and arterial). Each organ/tissue has a unique volume, permeability, and eliminating anatomical constants and terms, which are determined independently from the studied drug, while other physiological drug-specific parameters are later incorporated, such as affinity toward plasma proteins, tissue-to-plasma distribution rate, and even on target activity (K_m, V_{max}, or binding kinetics).

One of PBPK models' important features is the perspective for the mechanistic and prospective prediction of a drug's pharmacokinetic profiles. The use of drug-dependent parameters includes, but is not limited to, physicochemical properties, solubility and permeability values, and also the role of individual enzymes and transporters in the metabolism. Those parameters can be determined *in vitro*

Name	Description	Link/reference
ADME prediction		
vNN-ADMET	Public web server for ADMET property prediction based on 15 nearest neighbor models.	https://vnnadmet.bhsai.org/ [132]
Swiss-ADME	Public web server for ADME property prediction. It has a very unique Log*P* calculation (i.Log*P*) based on free energy.	http://www.swissadme.ch [133]
pkCSM	ADME web server based on chemical fragment similarity (the so-called graph-based signatures).	http://biosig.unimelb.edu.au/pkcsm/ [134]
ADMETlab	Web server using similarity-based ADME calculator models and drug-likeness analyses.	http://admet.scbdd.com/home/index/ [135]
Schrodinger—QikProp	Calculates pKa; Log*P*; water solubility—Schrodinger also offers other tools for property calculation.	https://www.schrodinger.com/ QikProp, Schrödinger, LLC, NY, 2019
DDI-Predictor	DDI-Predictor is able to make quantitative predictions of drug exposure even in cases where the interaction has not been studied yet.	https://www.ddi-predictor.org
PBPK models and platforms		
GastroPlus	Comprises 10 different modules including PBPK modeling and *in vitro* vs. *in vivo* correlation, can be parameterized for different disease states and age groups.	www.simulations-plus.com
PKSIM	PBK modeling tool with integrated database of anatomical and physiological parameters for humans, mouse, rat, dog, and monkey. Can model different scenarios depending on the chosen building blocks.	www.system sbiology.com/products/ pk-sim.html
Simcyp	Incorporates databases of genetic, physiological, and epidemiological information to enable simulation of different populations and species, ultimately is able to predict ADME parameters.	www.simcyp.com
ADMEWORKS DDI Simulator	As a differential is able to predict drug-drug interactions using nonlinear models.	http://www.fqs.pl/chemistry_ materials_life_science/products/ ddi_simulator

Table 2.
Tools for ADME prediction and PBPK modeling.

or calculated from the compound structure with other *in silico* approaches, which allows the early use of PBPK in the DDD (the *bottom-up* approach). Concurrently, it is also noteworthy that the model construction and parameter fine-tuning are a source of knowledge for the hit development, where the predictions from the ongoing model can help to understand the model's accuracy itself along the way (called as *middle-out* approach) and then prospectively be applied to simulate unstudied scenarios. Currently, there are several free-to-use and commercially available PBPK and ADME prediction options (**Table 2**), which are also extensively reviewed and discussed by the works of Madden et al. [130].

Early PBPK models, such as the work of Varma et al., described another layer of complexity by including drug-drug interactions (DDI). The dosing time-dependent model considering the interaction between repaglinide with rifampicin was able to predict repaglinide plasma concentrations along a day. The model also predicted the drug interaction with other CYP3A4 and OATP1B1 inhibitors, which could result in further DDIs. Reports of DDI leading to complications in patients with particular genotype stimulated studies such as the one performed by Fermier et al. [131], where the effects of polymorphic cytochromes provided the basis for a more accurate DDI prediction.

5. Biological (large) molecules

During recent years, larger molecules (LM) have gained in significance and popularity, due to achievements and approvals, as new molecular entities. These "*biologics*" are normally biotechnologically synthesized or recombinantly produced compounds of biological origin such as peptides, antibodies, and nucleic acids [136]. From a historical perspective, drug discovery and development of LMs are heavily delayed in comparison to SMs with their first approved entity happening in the 1980s [137]. At about the same time, two major inventions allowed huge progress in pharmacokinetics assessment of small molecules, contributing to smaller drop-out rates in later DDD stages [136]. One of them was the improved understanding of CYP450 mechanism and the other, the invention of (HP)LC-MS technology, fueled the assessment of the ADME parameters. LMs' discovery and development face many challenges, which demand high efforts to overcome but also offer unique opportunities in comparison to those of small molecules [138, 139].

The main differences between small and large molecules, despite the molecular weight, the number of heavy atoms, and torsions, can be found in the physicochemical properties, such as permeability, oral bioavailability, stability, specificity, and immunogenicity [138, 139]. New parameters, unique for large molecules, are also of interest, such as the physical particle size and the hydrodynamic radius, which has a dramatic effect on the absorption. Both parameters are related to the overall shape and correlate well with MW for globular proteins, but not necessarily for unstructured or highly modified entities. As a result, biologics are normally administered parenterally, only targeting extracellular structures; they are also more likely to trigger an immune response; and their production costs are considerably higher [139]. Interestingly, with the exception of the costs, these disadvantages can potentially be circumvented by appropriate delivery systems, for example, nanoparticle-based delivery to facilitate membrane permeation.

Other parameters, such as charges, which were previously modeled by pKa in case of small molecules, are heavily heterogeneous in LMs. The charge can be represented by the use of isoelectric points (pI), which are calculated from the available amino-acid sequence, and surface charge, which can use individual pKa's and structural information to be inferred. Overall protein charge often influences the biologic excretion [140], since negatively charged molecules undergo less renal filtration disregarding size effect [141].

While representing difficulties in the development of new molecular entities, the aforementioned properties also offer special advantages that small molecules cannot cover. As such, LMs normally have longer $t_{1/2}$, slower clearance, and higher selectivity; are multifunctional; and rarely expose drug interactions [139]. Apart from those, it was suggested that only 2–5% of the human genes can be targeted by small molecules, offering a niche for LMs' application against several diseases [138].

The increasing effort and development of new technologies, driven by the belief in higher success rates, enabled the latest advances in the field [138]. For instance, currently, peptide drugs only account for ~2% of the drug market but are in use in a wide range of diseases such as acromegaly and multiple sclerosis, together with different cancer types such as prostate and breast cancer.

Several other biologics are currently in use, namely monoclonal antibodies (mAbs) and bispecific antibodies (bsAbs), as example agents that activate or enhance the immunologic response. Of special interest in cancer therapy is a sub-class of bsAbs, so-called bispecific T-cell engager (BiTEs), which can recruit CD3 cells at the tumor site by binding to both cell types thereby directing the immuno-logical response [142].

Other interesting examples for biologics comprise hormones (e.g., insulin), cytokines (such as erythropoietin, EPO; IL-1; IL-2; IL-6) [143], nucleic acids such as siRNA (ONPATTRO) [144], and aptamers (Pegaptanib) [145]. While such a broad spectrum of molecule classes offers also a wide range of treatments, at the same time, it exacerbates the need for new developments since every molecule type exhibits different properties. In the field of predicting the biologics activity against specific targets, classical modeling tools, such as Monte Carlo sampling, genetic algo-rithms, docking, and molecular dynamics simulation, were adapted or even devel-oped anew to accommodate the specifics (as extensively reviewed by [146, 147]).

On the other hand, the absence of standard techniques to assess ADME proper-ties hampers the PK profiling and thus further development [136]. In fact, the current knowledge of LM pharmacokinetics is even impaired compared to the basic knowledge of ADME principles for small molecules in the 1980s [136]. Although the basic PK principles are similar between SMs and LMs, the specific mechanisms influencing each step of ADME are different. To begin with, the route of adminis-tration between them can differ, which leads to different mechanisms of absorption and first-pass metabolism. Furthermore, LMs are not metabolized by CYPs but can still trigger the release of pro-inflammatory cytokines leading to heavy side effects known as cytokine storm [136, 139]. Also, other modifications play a role in biolog-ics ADME, namely glycosylation, PEGylation, and neonatal Fc receptor (FcRn) interactions [139, 148]. Unfortunately, up until now, most of the evaluation of those factors is only addressed on *in vivo* level systems, which are not suited for HTS, are expensive and labor intensive, and require longer bioethics evaluation.

In this regard, the development of *in vitro* and *in silico* methods to evaluate ADME should be a high-profile goal. One of the main challenges will be to find a way to integrate as many of the biologics into the process in order to facilitate ADME assessment and guide large molecules' DDD as already implemented for their smaller counterparts.

6. Conclusions

The main difficulties in PK profiling lie in the high costs and comparable low throughput of *in vivo* models. The extensive use of animals in DDD also raises ethi-cal issues and is further affected since animal models not always translate readily to the humans, especially in terms of metabolism [149, 150]. Furthermore, the advent of combinatorial chemistry coupled with HTS for efficacy evaluation leads to an explosion in the number to an extent that the classical PK assays could not compen-sate [29, 47]. *In vitro* PK screens are supposed to offer a solution to the problem by complementing *in vivo* assessment to reduce costs while increasing efficiency, but they also suffer from shortcomings. In general, one must distinguish between two main forms of *in vitro* systems: static and dynamic models. Only dynamic models

are suited for PK evaluation because they allow variation of compound concentrations, a key factor in pharmacokinetics. In this sense, diffusion-based dynamic *in vitro* models offer a solution but still are quite limited in terms of high throughput and costs. An alternative was presented by Lockwood and colleagues in the form of a 3D-printed fluidic device utilizing trans-well technique generating dynamic *in vitro* PK profiles also applicable for HTS infrastructure [149].

What distinguishes the DDD "then" and "now" is principally two main changes. First, in the past, pharmaceutical companies as well as academic laboratories were not that concerned with ADMET assessment in the early stages of drug discovery (hit and lead generation) and only addressed PK from preclinical stages on forward. Instead, HTS/HCS, genomics, and computational chemistry were high-profile areas. Today, almost all pharmaceutical big-players have shifted pharmacokinetic profiling to discovery phases. However, only the future will tell whether those changes will yield fruit.

Second, CADD became more and more part of the DDD pipeline in different stages facilitating fast screening of compounds *in silico* and supporting QSAR. Although bioinformatics techniques already substituted many *in vitro* tests, basically all of them require *in vitro* and/or *in vivo* validation and standardization to guarantee trustable predictions. Another important aspect, recently addressed by the work of Ferreria and Andricopulo [151], is the importance of translating those models into well-structured and user-friendly (online) platforms that can be accessed and used by the drug discovery community. Still, the efficacy and reliability of computer simulations increase permanently and drastically, and many see a future of solely virtual drug discovery. Thankfully, these failures resulted in the consequence of addressing safety and efficacy concern earlier in the drug discovery process, for instance, via *in vitro* screens to assess metabolic stability and absorption properties and diminish failure rates later on [13].

Acknowledgements

Notes

ORCID numbers of all authors:

Thales Kronenberger: 0000-0001-6933-7590
Carsten Wrenger: 0000-0001-5987-1749
Vinicius Goncalves Maltarollo: 0000-0001-9675-5907
Arne Krüger: 0000-0002-5531-9508

Conflict of interest

The authors declare no conflict of interest.

Author details

Arne Krüger[1], Vinicius Gonçalves Maltarollo[2], Carsten Wrenger[1]*
and Thales Kronenberger[3]

1 Unit for Drug Discovery, Department of Parasitology, Institute of Biomedical
Sciences, University of São Paulo, São Paulo, Brazil

2 Department of Pharmaceutical Products, Pharmacy Faculty, Federal University of
Minas Gerais, Belo Horizonte, Minas Gerais, Brazil

3 Department of Internal Medicine VIII, University Hospital of Tübingen,
Tübingen, Germany

*Address all correspondence to: cwrenger@icb.usp.br

IntechOpen

References

[1] Strovel J, Sittampalam S, Coussens NP, Hughes M, Inglese J, Kurtz A, et al. Early drug discovery and development guidelines: For academic researchers, collaborators, and start-up companies. In: Sittampalam GS, Coussens NP, Brimacombe K, Grossman A, Arkin M, Auld D, et al., editors. Assay Guidance. Manual, Bethesda (MD): Eli Lilly & Company and the National Center for Advancing Translational Sciences; 2004

[2] BIO Industry Analysis Reports. BIO n.d. https://www.bio.org/bio-industry-analysis-reports [Accessed February 13, 2019]

[3] Mullard A. Parsing clinical success rates. Nature Reviews Drug Discovery. 2016;**15**:447. DOI: 10.1038/nrd.2016.136

[4] Hughes JP, Rees S, Kalindjian SB, Philpott KL. Principles of early drug discovery. British Journal of Pharmacology. 2011;**162**:1239-1249. DOI: 10.1111/j.1476-5381.2010.01127.x

[5] Van Norman GA. Drugs, devices, and the FDA: Part 1: An overview of approval processes for drugs. JACC: Basic to Translational Science. 2016;**1**:170-179. DOI: 10.1016/j.jacbts.2016.03.002

[6] Paul SM, Mytelka DS, Dunwiddie CT, Persinger CC, Munos BH, Lindborg SR, et al. How to improve R&D productivity: The pharmaceutical industry's grand challenge. Nature Reviews. Drug Discovery. 2010;**9**:203-214. DOI: 10.1038/nrd3078

[7] DiMasi JA, Grabowski HG, Hansen RW. Innovation in the pharmaceutical industry: New estimates of R&D costs. Journal of Health Economics. 2016;**47**:20-33. DOI: 10.1016/j.jhealeco.2016.01.012

[8] Shearer TW, Smith KS, Diaz D, Asher C, Ramirez J. The role of in vitro ADME assays in antimalarial drug discovery and development. Combinatorial Chemistry & High Throughput Screening. 2005;**8**:89-98

[9] Fraietta I, Gasparri F. The development of high-content screening (HCS) technology and its importance to drug discovery. Expert Opinion on Drug Discovery. 2016;**11**:501-514. DOI: 10.1517/17460441.2016.1165203

[10] Lombardo F, Desai PV, Arimoto R, Desino KE, Fischer H, Keefer CE, et al. In silico absorption, distribution, metabolism, excretion, and pharmacokinetics (ADME-PK): Utility and best practices. An industry perspective from the international consortium for innovation through quality in pharmaceutical development. Journal of Medicinal Chemistry. 2017;**60**:9097-9113. DOI: 10.1021/acs.jmedchem.7b00487

[11] FDA. The Drug Development Process. n.d. https://www.fda.gov/ForPatients/Approvals/Drugs/ [Accessed February 13, 2019].

[12] Sinha S, Vohora D. Chapter 2—Drug discovery and development: An overview. In: Vohora D, Singh G, editors. Pharmaceutical Medicine and Translational Clinical Research. Boston: Academic Press; 2018. pp. 19-32. DOI: 10.1016/B978-0-12-802103-3.00002-X

[13] Hefti FF. Requirements for a lead compound to become a clinical candidate. BMC Neuroscience. 2008;**9**:S7. DOI: 10.1186/1471-2202-9-S3-S7

[14] Andrews KA, Wesche D, McCarthy J, Möhrle JJ, Tarning J, Phillips L, et al. Model-informed drug development for malaria therapeutics. Annual Review of Pharmacology and Toxicology. 2018;**58**:567-582. DOI: 10.1146/annurev-pharmtox-010715-103429

[15] Lombardino JG, Lowe JA. The role of the medicinal chemist in drug discovery—Then and now. Nature Reviews. Drug Discovery. 2004;**3**:853-862. DOI: 10.1038/nrd1523

[16] Maltarollo VG, Kronenberger T, Espinoza GZ, Oliveira PR, Honorio KM. Advances with support vector machines for novel drug discovery. Expert Opinion on Drug Discovery. 2019;**14**:23-33. DOI: 10.1080/17460441.2019.1549033

[17] Huang J, Fan X. Why QSAR fails: An empirical evaluation using conventional computational approach. Molecular Pharmaceutics. 2011;**8**:600-608. DOI: 10.1021/mp100423u

[18] Validation of (Q)SAR Models—OECD. n.d. http://www.oecd.org/chemicalsafety/risk-assessment/validationofqsarmodels.htm [Accessed February 13, 2019]

[19] Gramatica P, Sangion A. A historical excursus on the statistical validation parameters for QSAR models: A clarification concerning metrics and terminology. Journal of Chemical Information and Modeling. 2016;**56**:1127-1131. DOI: 10.1021/acs.jcim.6b00088

[20] Abdel-Rahman SM, Kauffman RE. The integration of pharmacokinetics and pharmacodynamics: Understanding dose-response. Annual Review of Pharmacology and Toxicology. 2004;**44**:111-136. DOI: 10.1146/annurev.pharmtox.44.101802.121347

[21] Hodgson J. ADMET—Turning chemicals into drugs. Nature Biotechnology. 2001;**19**:722-726. DOI: 10.1038/90761

[22] Tietgen H, Walden M. Physicochemical properties. In: Vogel HG, Maas J, Hock FJ, Mayer D, editors. Drug Discovery and Evaluation: Safety and Pharmacokinetic Assays. Berlin, Heidelberg: Springer Berlin Heidelberg; 2013. pp. 1125-1138. DOI: 10.1007/978-3-642-25240-2_48

[23] Vaddady PK, Lee RE, Meibohm B. In vitro pharmacokinetic/pharmacodynamic models in anti-infective drug development: Focus on TB. Future Medicinal Chemistry. 2010;**2**:1355-1369. DOI: 10.4155/fmc.10.224

[24] Dowty ME, Messing DM, Lai Y, Kirkovsky L. Adme. ADMET for Medicinal Chemists. John Wiley & Sons, Ltd; 2010. pp. 145-200. DOI: 10.1002/9780470915110.ch4

[25] Valkó KL. Lipophilicity and biomimetic properties measured by HPLC to support drug discovery. Journal of Pharmaceutical and Biomedical Analysis. 2016;**130**:35-54. DOI: 10.1016/j.jpba.2016.04.009

[26] Fan J, de Lannoy IAM. Pharmacokinetics. Biochemical Pharmacology. 2014;**87**:93-120. DOI: 10.1016/j.bcp.2013.09.007

[27] Amidon GL, Lennernäs H, Shah VP, Crison JR. A theoretical basis for a biopharmaceutic drug classification: The correlation of in vitro drug product dissolution and in vivo bioavailability. Pharmaceutical Research. 1995;**12**:413-420

[28] Lennernäs H. Human intestinal permeability. Journal of Pharmaceutical Sciences. 1998;**87**:403-410. DOI: 10.1021/js970332a

[29] Avdeef A. Physicochemical profiling (solubility, permeability and charge state). Current Topics in Medicinal Chemistry. 2001;**1**:277-351

[30] Comer J, Box K. High-throughput measurement of drug pK a values for ADME screening. Journal of the Association for Laboratory Automation. 2003;**1**:55-59. DOI: 10.1016/S1535-5535(04)00243-6

[31] Valko K, Butler J, Eddershaw P. Predictive approaches to increase absorption of compounds during lead optimisation. Expert Opinion on Drug Discovery. 2013;**8**:1225-1238. DOI: 10.1517/17460441.2013.815613

[32] Tolls J, van Dijk J, Verbruggen EJM, Hermens JLM, Loeprecht B, Schüürmann G. Aqueous solubility–molecular size relationships: A mechanistic case study using C10- to C19-alkanes. The Journal of Physical Chemistry. A. 2002;**106**:2760-2765. DOI: 10.1021/jp011755a

[33] Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Advanced Drug Delivery Reviews. 2001;**46**:3-26

[34] Hou T, Wang J, Zhang W, Xu X. ADME evaluation in drug discovery. 6. Can oral bioavailability in humans be effectively predicted by simple molecular property-based rules? Journal of Chemical Information and Modeling. 2007;**47**:460-463. DOI: 10.1021/ci6003515

[35] Veber DF, Johnson SR, Cheng H-Y, Smith BR, Ward KW, Kopple KD. Molecular properties that influence the oral bioavailability of drug candidates. Journal of Medicinal Chemistry. 2002;**45**:2615-2623

[36] Daina A, Zoete V. A BOILED-egg to predict gastrointestinal absorption and brain penetration of small molecules. ChemMedChem. 2016;**11**:1117-1121. DOI: 10.1002/cmdc.201600182

[37] Egan WJ, Merz Kenneth M, Baldwin JJ. Prediction of drug absorption using multivariate statistics. Journal of Medicinal Chemistry. 2000;**43**:3867-3877. DOI: 10.1021/jm000292e

[38] Lovering F, Bikker J, Humblet C. Escape from flatland: Increasing saturation as an approach to improving clinical success. Journal of Medicinal Chemistry. 2009;**52**:6752-6756. DOI: 10.1021/jm901241e

[39] Ritchie TJ, Macdonald SJF, Young RJ, Pickett SD. The impact of aromatic ring count on compound developability: Further insights by examining carbo- and hetero-aromatic and -aliphatic ring types. Drug Discovery Today. 2011;**16**:164-171. DOI: 10.1016/j.drudis.2010.11.014

[40] Meissner KA, Kronenberger T, Maltarollo VG, Trossini GHG, Wrenger C. Targeting the plasmodium flaciparum plasmepsin V by ligand-based virtual screening. Chemical Biology & Drug Design. Mar 2019;**93**(3):300-312. DOI: 10.1111/cbdd.13416

[41] Yao T-T, Xie J-F, Liu X-G, Cheng J-L, Zhu C-Y, Zhao J-H, et al. Integration of pharmacophore mapping and molecular docking in sequential virtual screening: Towards the discovery of novel JAK2 inhibitors. RSC Advances. 2017;**7**:10353-10360. DOI: 10.1039/C6RA24959K

[42] Naz S, Farooq U, Ali S, Sarwar R, Khan S, Abagyan R. Identification of new benzamide inhibitor against α-subunit of tryptophan synthase from *Mycobacterium tuberculosis* through structure-based virtual screening, anti-tuberculosis activity and molecular dynamics simulations. Journal of Biomolecular Structure and Dynamics. 2018;**0**:1-11. DOI: 10.1080/07391102.2018.1448303

[43] Morgan ME, Lui K, Anderson BD. Microscale titrimetric and spectrophotometric methods for determination of ionization constants and partition coefficients of new drug candidates. Journal of Pharmaceutical Sciences. 1998;**87**:238-245. DOI: 10.1021/js970057s

[44] Kibbey CE, Poole SK, Robinson B, Jackson JD, Durham D. An

integrated process for measuring the physicochemical properties of drug candidates in a preclinical discovery environment. Journal of Pharmaceutical Sciences. 2001;**90**:1164-1175

[45] Miller JM, Blackburn AC, Shi Y, Melzak AJ, Ando HY. Semi-empirical relationships between effective mobility, charge, and molecular weight of pharmaceuticals by pressure-assisted capillary electrophoresis: Applications in drug discovery. Electrophoresis. 2002;**23**:2833-2841. DOI: 10.1002/1522-2683(200209)23:17<2833::AID-ELPS2833>3.0.CO;2-7

[46] Zhou C, Jin Y, Kenseth JR, Stella M, Wehmeyer KR, Heineman WR. Rapid pKa estimation using vacuum-assisted multiplexed capillary electrophoresis (VAMCE) with ultraviolet detection. Journal of Pharmaceutical Sciences. 2005;**94**:576-589. DOI: 10.1002/jps.20275

[47] Bevan CD, Lloyd RS. A high-throughput screening method for the determination of aqueous drug solubility using laser nephelometry in microtiter plates. Analytical Chemistry. 2000;**72**:1781-1787

[48] Avdeef A. pH-metric solubility. 1. Solubility-pH profiles from bjerrum plots. Gibbs buffer and pKa in the solid state. Pharmacy and Pharmacology Communications. 1998;**4**:165-178. DOI: 10.1111/j.2042-7158.1998.tb00328.x

[49] Wan H, Holmén AG. High throughput screening of physicochemical properties and in vitro ADME profiling in drug discovery. Combinatorial Chemistry & High Throughput Screening. 2009;**12**:315-329

[50] Valkó K. Chapter 12. Measurements of physical properties for drug design in industry. In: Valkó K, editor. Handbook of Analytical Separations. Vol. 1.

Amsterdam: Netherlands; Elsevier Science B.V.; 2000. pp. 535-583. DOI: 10.1016/S1567-7192(00)80015-7

[51] Hitzel L, Watt AP, Locker KL. An increased throughput method for the determination of partition coefficients. Pharmaceutical Research. 2000;**17**:1389-1395

[52] Jain N, Yalkowsky SH. Estimation of the aqueous solubility I: Application to organic nonelectrolytes. Journal of Pharmaceutical Sciences. 2001;**90**:234-252

[53] Ran Y, He Y, Yang G, Johnson JLH, Yalkowsky SH. Estimation of aqueous solubility of organic compounds by using the general solubility equation. Chemosphere. 2002;**48**:487-509

[54] Jain P, Yalkowsky SH. Prediction of aqueous solubility from SCRATCH. International Journal of Pharmaceutics. 2010;**385**:1-5. DOI: 10.1016/j.ijpharm.2009.10.003

[55] Ali J, Camilleri P, Brown MB, Hutt AJ, Kirton SB. Revisiting the general solubility equation: In Silico prediction of aqueous solubility incorporating the effect of topographical polar surface area. Journal of Chemical Information and Modeling. 2012;**52**:420-428. DOI: 10.1021/ci200387c

[56] Wang J, Hou T. Recent advances on aqueous solubility prediction. Combinatorial Chemistry & High Throughput Screening. 2011;**14**:328-338

[57] Chevillard F, Lagorce D, Reynès C, Villoutreix BO, Vayer P, Miteva MA. In silico prediction of aqueous solubility: A multimodel protocol based on chemical similarity. Molecular Pharmaceutics. 2012;**9**:3127-3135. DOI: 10.1021/mp300234q

[58] Bergström CAS, Wassvik CM, Norinder U, Luthman K, Artursson P. Global and local computational models for aqueous solubility prediction

of drug-like molecules. Journal of Chemical Information and Computer Sciences. 2004;**44**:1477-1488. DOI: 10.1021/ci049909h

[59] Raevsky OA, Polianczyk DE, Grigorev VY, Raevskaja OE, Dearden JC. In silico prediction of aqueous solubility: A comparative study of local and global predictive models. Molecular Informatics. 2015;**34**:417-430. DOI: 10.1002/minf.201400144

[60] Lipinski CA. Rule of five in 2015 and beyond: Target and ligand structural limitations, ligand chemistry structure and drug discovery project decisions. Advanced Drug Delivery Reviews. 2016;**101**:34-41. DOI: 10.1016/j.addr.2016.04.029

[61] Lee PH, Ayyampalayam SN, Carreira LA, Shalaeva M, Bhattachar S, Coselmon R, et al. In silico prediction of ionization constants of drugs. Molecular Pharmaceutics. 2007;**4**:498-512. DOI: 10.1021/mp070019+

[62] Cruciani G, Milletti F, Storchi L, Sforna G, Goracci L. In silico pKa prediction and ADME profiling. Chemistry & Biodiversity. 2009;**6**:1812-1821. DOI: 10.1002/cbdv.200900153

[63] Shelley JC, Cholleti A, Frye LL, Greenwood JR, Timlin MR, Uchimaya M. Epik: A software program for pKa prediction and protonation state generation for drug-like molecules. Journal of Computer-Aided Molecular Design. 2007;**21**:681-691. DOI: 10.1007/s10822-007-9133-z

[64] Ho J. Predicting pKa in implicit solvents: Current status and future directions. Australian Journal of Chemistry. 2014;**67**:1441. DOI: 10.1071/CH14040

[65] Jensen JH, Swain CJ, Olsen L. Prediction of pKa values for druglike molecules using semiempirical quantum chemical methods. The Journal of Physical Chemistry. A. 2017;**121**:699-707. DOI: 10.1021/acs.jpca.6b10990

[66] Artursson P. Cell cultures as models for drug absorption across the intestinal mucosa. Critical Reviews in Therapeutic Drug Carrier Systems. 1991;**8**:305-330

[67] Hidalgo IJ, Borchardt RT. Transport of bile acids in a human intestinal epithelial cell line, Caco-2. Biochimica et Biophysica Acta (BBA)—General Subjects. 1990;**1035**:97-103. DOI: 10.1016/0304-4165(90)90179-Z

[68] Press B, Di Grandi D. Permeability for intestinal absorption: Caco-2 assay and related issues. Current Drug Metabolism. 2008;**9**:893-900

[69] Lee JB, Son SH, Park MC, Kim TH, Kim MG, Yoo SD, et al. A novel in vitro permeability assay using three-dimensional cell culture system. Journal of Biotechnology. 2015;**205**:93-100. DOI: 10.1016/j.jbiotec.2014.12.019

[70] Volpe DA. Drug-permeability and transporter assays in Caco-2 and MDCK cell lines. Future Medicinal Chemistry. 2011;**3**:2063-2077. DOI: 10.4155/fmc.11.149

[71] Tang F, Horie K, Borchardt RT. Are MDCK cells transfected with the human MRP2 gene a good model of the human intestinal mucosa? Pharmaceutical Research. 2002;**19**:773-779

[72] Jin X, Luong T-L, Reese N, Gaona H, Collazo-Velez V, Vuong C, et al. Comparison of MDCK-MDR1 and Caco-2 cell based permeability assays for anti-malarial drug screening and drug investigations. Journal of Pharmacological and Toxicological Methods. 2014;**70**:188-194. DOI: 10.1016/j.vascn.2014.08.002

[73] Kansy M, Avdeef A, Fischer H. Advances in screening for membrane permeability: High-resolution PAMPA for medicinal chemists. Drug Discovery

Today: Technologies. 2004;**1**:349-355. DOI: 10.1016/j.ddtec.2004.11.013

[74] Avdeef A. The rise of PAMPA. Expert Opinion on Drug Metabolism & Toxicology. 2005;**1**: 325-342. DOI: 10.1517/17425255.1.2.325

[75] Bermejo M, Avdeef A, Ruiz A, Nalda R, Ruell JA, Tsinman O, et al. PAMPA—A drug absorption in vitro model: 7. Comparing rat in situ, Caco-2, and PAMPA permeability of fluoroquinolones. European Journal of Pharmaceutical Sciences. 2004;**21**: 429-441. DOI: 10.1016/j.ejps.2003.10.009

[76] Cho H-J, Kim J-E, Kim D-D, Yoon I-S. In vitro-in vivo extrapolation (IVIVE) for predicting human intestinal absorption and first-pass elimination of drugs: Principles and applications. Drug Development and Industrial Pharmacy. 2014;**40**:989-998. DOI: 10.3109/03639045.2013.831439

[77] Wang N-N, Dong J, Deng Y-H, Zhu M-F, Wen M, Yao Z-J, et al. ADME properties evaluation in drug discovery: Prediction of Caco-2 cell permeability using a combination of NSGA-II and boosting. Journal of Chemical Information and Modeling. 2016;**56**:763-773. DOI: 10.1021/acs.jcim.5b00642

[78] Timmermans PB, Brands A, van Zwieten PA. Lipophilicity and brain disposition of clonidine and structurally related imidazolidines. Naunyn-Schmiedeberg's Archives of Pharmacology. 1977;**300**:217-226

[79] Yan A, Liang H, Chong Y, Nie X, Yu C. In-silico prediction of blood-brain barrier permeability. SAR and QSAR in Environmental Research. 2013;**24**:61-74. DOI: 10.1080/1062936X.2012.729224

[80] Shen J, Cheng F, Xu Y, Li W, Tang Y. Estimation of ADME properties with substructure pattern recognition. Journal of Chemical Information and Modeling. 2010;**50**:1034-1041. DOI: 10.1021/ci100104j

[81] Wang Z, Yang H, Wu Z, Wang T, Li W, Tang Y, et al. In silico prediction of blood–brain barrier permeability of compounds by machine learning and resampling methods. ChemMedChem. 2018;**13**:2189-2201. DOI: 10.1002/cmdc.201800533

[82] Wang W, Kim MT, Sedykh A, Zhu H. Developing enhanced blood-brain barrier permeability models: Integrating external bio-assay data in QSAR modeling. Pharmaceutical Research. 2015;**32**:3055-3065. DOI: 10.1007/s11095-015-1687-1

[83] Pugh WJ, Roberts MS, Hadgraft J. Epidermal permeability—Penetrant structure relationships: 3. The effect of hydrogen bonding interactions and molecular size on diffusion across the stratum corneum. International Journal of Pharmaceutics. 1996;**138**:149-165. DOI: 10.1016/0378-5173(96)04533-4

[84] Potts RO, Guy RH. Predicting skin permeability. Pharmaceutical Research. 1992;**9**:663-669. DOI: 10.1023/A:1015810312465

[85] Degim IT. New tools and approaches for predicting skin permeability. Drug Discovery Today. 2006;**11**:517-523. DOI: 10.1016/j.drudis.2006.04.006

[86] Pecoraro B, Tutone M, Hoffman E, Hutter V, Almerico AM, Traynor M. Predicting skin permeability by means of computational approaches: Reliability and caveats in pharmaceutical studies. Journal of Chemical Information and Modeling. 2019. DOI: 10.1021/acs.jcim.8b00934. [Epub ahead of print]

[87] Ermondi G, Vallaro M, Caron G. Learning how to use IAM chromatography for predicting permeability. European Journal

of Pharmaceutical Sciences. 2018;**114**:385-390. DOI: 10.1016/j. ejps.2018.01.001

[88] Smith BT. Remington Education: Physical Pharmacy. London, UK: Pharmaceutical Press; 2015. ISBN 978 0 85711 106 7

[89] Dressman JB, Amidon GL, Reppas C, Shah VP. Dissolution testing as a prognostic tool for oral drug absorption: Immediate release dosage forms. Pharmaceutical Research. 1998;**15**:11-22

[90] Comer J, Tam K. Lipophilicity Profiles: Theory and Measurement. Pharmacokinetic Optimization in Drug Research. Postfach, Switzerland; John Wiley & Sons, Ltd; Verlag Helvetica Chimica Acta; 2001. pp. 275-304. ISBN 978 0 85711 106 7. DOI: 10.1002/9783906390437.ch17

[91] Huang W, Lee SL, Yu LX. Mechanistic approaches to predicting oral drug absorption. The AAPS Journal. 2009;**11**:217-224. DOI: 10.1208/ s12248-009-9098-z

[92] Yu LX, Lipka E, Crison JR, Amidon GL. Transport approaches to the biopharmaceutical design of oral drug delivery systems: Prediction of intestinal absorption. Advanced Drug Delivery Reviews. 1996;**19**:359-376

[93] Agoram B, Woltosz WS, Bolger MB. Predicting the impact of physiological and biochemical processes on oral drug bioavailability. Advanced Drug Delivery Reviews. 2001;**50**(Suppl 1):S41-S67

[94] Xia B, Yang Z, Zhou H, Lukacova V, Zhu W, Milewski M, et al. Development of a novel oral cavity compartmental absorption and transit model for sublingual administration: Illustration with Zolpidem. The AAPS Journal. 2015;**17**:631-642. DOI: 10.1208/ s12248-015-9727-7

[95] Kumar R, Sharma A, Siddiqui MH, Tiwari RK. Promises of machine learning approaches in prediction of absorption of compounds. Mini Reviews in Medicinal Chemistry. 2018;**18**:196-207. DOI: 10.217 4/1389557517666170315150116

[96] Trainor GL. The importance of plasma protein binding in drug discovery. Expert Opinion on Drug Discovery. 2007;**2**:51-64. DOI: 10.1517/17460441.2.1.51

[97] Kratochwil NA, Huber W, Müller F, Kansy M, Gerber PR. Predicting plasma protein binding of drugs: A new approach. Biochemical Pharmacology. 2002;**64**:1355-1374

[98] Hollósy F, Valkó K, Hersey A, Nunhuck S, Kéri G, Bevan C. Estimation of volume of distribution in humans from high throughput HPLC-based measurements of human serum albumin binding and immobilized artificial membrane partitioning. Journal of Medicinal Chemistry. 2006;**49**:6958-6971. DOI: 10.1021/jm050957i

[99] Liang C, Lian H. Recent advances in lipophilicity measurement by reversed-phase high-performance liquid chromatography. TrAC Trends in Analytical Chemistry. 2015;**68**:28-36. DOI: 10.1016/j.trac.2015.02.009

[100] Valkó K. Application of high-performance liquid chromatography based measurements of lipophilicity to model biological distribution. Journal of Chromatography. A. 2004;**1037**:299-310

[101] Huang Z, Ung T. Effect of alpha-1-acid glycoprotein binding on pharmacokinetics and pharmacodynamics. Current Drug Metabolism. 2013;**14**:226-238

[102] Lombardo F, Obach RS, Shalaeva MY, Gao F. Prediction of volume of distribution values in humans for neutral and basic drugs using

physicochemical measurements and plasma protein binding data. Journal of Medicinal Chemistry. 2002;**45**:2867-2876

[103] Amo EM d, Ghemtio L, Xhaard H, Yliperttula M, Urtti A, Kidron H. Applying linear and non-linear methods for parallel prediction of volume of distribution and fraction of unbound drug. PLoS One. 2013;**8**:e74758. DOI: 10.1371/journal.pone.0074758

[104] Lombardo F, Jing Y. In silico prediction of volume of distribution in humans. Extensive data set and the exploration of linear and nonlinear methods coupled with molecular interaction fields descriptors. Journal of Chemical Information and Modeling. 2016;**56**:2042-2052. DOI: 10.1021/acs. jcim.6b00044

[105] Li H, Chen Z, Xu X, Sui X, Guo T, Liu W, et al. Predicting human plasma protein binding of drugs using plasma protein interaction QSAR analysis (PPI-QSAR). Biopharmaceutics & Drug Disposition. 2011;**32**:333-342. DOI: 10.1002/bdd.762

[106] Ghafourian T, Amin ZQSAR. Models for the prediction of plasma protein binding. BioImpacts: BI. 2013;**3**:21-27. DOI: 10.5681/bi.2013.011.

[107] Sun L, Yang H, Li J, Wang T, Li W, Liu G, et al. In silico prediction of compounds binding to human plasma proteins by QSAR models. ChemMedChem. 2018;**13**:572-581. DOI: 10.1002/cmdc.201700582

[108] del Amo EM, Vellonen K-S, Kidron H, Urtti A. Intravitreal clearance and volume of distribution of compounds in rabbits: In silico prediction and pharmacokinetic simulations for drug development. European Journal of Pharmaceutics and Biopharmaceutics. 2015;**95**:215-226. DOI: 10.1016/j. ejpb.2015.01.003

[109] Vellonen K-S, Soini E-M, del Amo EM, Urtti A. Prediction of ocular drug distribution from systemic blood circulation. Molecular Pharmaceutics. 2016;**13**:2906-2911. DOI: 10.1021/acs. molpharmaceut.5b00729

[110] Miyamoto M, Iwasaki S, Chisaki I, Nakagawa S, Amano N, Kosugi Y, et al. Prediction of human pharmacokinetics of long half-life compounds using chimeric mice with humanised liver. Xenobiotica. 2019:1-31. DOI: 10.1080/00498254.2019.1579394. [Epub ahead of print]

[111] Eyal S, Hsiao P, Unadkat JD. Drug interactions at the blood-brain barrier: Fact or fantasy? Pharmacology & Therapeutics. 2009;**123**:80-104. DOI: 10.1016/j.pharmthera.2009.03.017

[112] Moroy G, Martiny VY, Vayer P, Villoutreix BO, Miteva MA. Toward in silico structure-based ADMET prediction in drug discovery. Drug Discovery Today. 2012;**17**:44-55. DOI: 10.1016/j.drudis.2011.10.023

[113] Braga RC, Alves VM, Fraga CAM, Barreiro EJ, de Oliveira V, Andrade CH. Combination of docking, molecular dynamics and quantum mechanical calculations for metabolism prediction of 3,4-methylenedioxybenzoyl-2-thienylhydrazone. Journal of Molecular Modeling. 2012;**18**:2065-2078. DOI: 10.1007/s00894-011-1219-9

[114] Sadowski P, Fooshee D, Subrahmanya N, Baldi P. Synergies between quantum mechanics and machine learning in reaction prediction. Journal of Chemical Information and Modeling. 2016;**56**:2125-2128. DOI: 10.1021/acs.jcim.6b00351

[115] Kirchmair J, Williamson MJ, Afzal AM, Tyzack JD, Choy APK, Howlett A, et al. FAst MEtabolizer (FAME): A rapid and accurate predictor of sites of metabolism in multiple species

by endogenous enzymes. Journal of Chemical Information and Modeling. 2013;**53**:2896-2907. DOI: 10.1021/ci400503s

[116] Rydberg P, Olsen L. Ligand-based site of metabolism prediction for cytochrome P450 2D6. ACS Medicinal Chemistry Letters. 2012;**3**:69-73. DOI: 10.1021/ml200246f

[117] Rydberg P, Rostkowski M, Gloriam DE, Olsen L. The contribution of atom accessibility to site of metabolism models for cytochromes P450. Molecular Pharmaceutics. 2013;**10**:1216-1223. DOI: 10.1021/mp3005116

[118] Olsen L, Montefiori M, Tran KP, Jørgensen FS. SMARTCyp 3.0: Enhanced cytochrome P450 site-of-metabolism prediction server. Bioinformatics. 2019. DOI: 10.1093/bioinformatics/btz037. [Epub ahead of print]

[119] Li J, Schneebeli ST, Bylund J, Farid R, Friesner RA. IDSite: An accurate approach to predict P450-mediated drug metabolism. Journal of Chemical Theory and Computation. 2011;**7**:3829-3845. DOI: 10.1021/ct200462q

[120] Kingsley LJ, Wilson GL, Essex ME, Lill MA. Combining structure- and ligand-based approaches to improve site of metabolism prediction in CYP2C9 substrates. Pharmaceutical Research. 2015;**32**:986-1001. DOI: 10.1007/s11095-014-1511-3

[121] Gombar VK, Hall SD. Quantitative structure-activity relationship models of clinical pharmacokinetics: Clearance and volume of distribution. Journal of Chemical Information and Modeling. 2013;**53**:948-957. DOI: 10.1021/ci400001u

[122] Kusama M, Toshimoto K, Maeda K, Hirai Y, Imai S, Chiba K, et al. In silico classification of major clearance pathways of drugs with

their physiochemical parameters. Drug Metabolism and Disposition. 2010;**38**:1362-1370. DOI: 10.1124/dmd.110.032789

[123] Toshimoto K, Wakayama N, Kusama M, Maeda K, Sugiyama Y, Akiyama Y. In silico prediction of major drug clearance pathways by support vector machines with feature-selected descriptors. Drug Metabolism and Disposition. 2014;**42**:1811-1819. DOI: 10.1124/dmd.114.057893

[124] Berellini G, Waters NJ, Lombardo F. In silico prediction of total human plasma clearance. Journal of Chemical Information and Modeling. 2012;**52**:2069-2078. DOI: 10.1021/ci300155y

[125] Xu X, Zhang W, Huang C, Li Y, Yu H, Wang Y, et al. A novel chemometric method for the prediction of human oral bioavailability. International Journal of Molecular Sciences. 2012;**13**:6964-6982. DOI: 10.3390/ijms13066964

[126] Broccatelli F, Carosati E, Neri A, Frosini M, Goracci L, Oprea TI, et al. A novel approach for predicting P-glycoprotein (ABCB1) inhibition using molecular interaction fields. Journal of Medicinal Chemistry. 2011;**54**:1740-1751. DOI: 10.1021/jm101421d

[127] Wood FL, Houston JB, Hallifax D. Clearance prediction methodology needs fundamental improvement: Trends common to rat and human hepatocytes/microsomes and implications for experimental methodology. Drug Metabolism and Disposition. 2017;**45**:1178-1188. DOI: 10.1124/dmd.117.077040

[128] Teorell T. Kinetics of distribution of substances administered to the body, I : The extravascular modes of administration. Archives Internationales de Pharmacodynamie et de Therapie. 1937;**57**:205-225

[129] Shebley M, Sandhu P, Emami Riedmaier A, Jamei M, Narayanan R, Patel A, et al. Physiologically based pharmacokinetic model qualification and reporting procedures for regulatory submissions: A consortium perspective. Clinical Pharmacology and Therapeutics. 2018;**104**:88-110. DOI: 10.1002/cpt.1013

[130] Madden JC, Pawar G, Cronin MTD, Webb S, Tan Y-M, Paini A. In silico resources to assist in the development and evaluation of physiologically-based kinetic models. Computational Toxicology. 2019;**11**:33-49. DOI: 10.1016/j.comtox.2019.03.001

[131] Fermier N, Bourguignon L, Goutelle S, Bleyzac N, Tod M. Identification of cytochrome P450-mediated drug–drug interactions at risk in cases of gene polymorphisms by using a quantitative prediction model. Clinical Pharmacokinetics. 2018;**57**:1581-1591. DOI: 10.1007/s40262-018-0651-8

[132] Schyman P, Liu R, Desai V, Wallqvist A. vNN web server for ADMET predictions. Frontiers in Pharmacology. 2017;**8**:889. DOI: 10.3389/fphar.2017.00889

[133] Daina A, Michielin O, Zoete V. SwissADME: A free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. Scientific Reports. 2017;**7**:42717. DOI: 10.1038/srep42717

[134] Pires DEV, Blundell TL, Ascher DB. pkCSM: Predicting small-molecule pharmacokinetic and toxicity properties using graph-based signatures. Journal of Medicinal Chemistry. 2015;**58**:4066-4072. DOI: 10.1021/acs.jmedchem.5b00104

[135] Dong J, Wang N-N, Yao Z-J, Zhang L, Cheng Y, Ouyang D, et al. ADMETlab: A platform for systematic ADMET evaluation based on a comprehensively collected ADMET database. Journal of Cheminformatics. 2018;**10**:1:11. DOI: 10.1186/s13321-018-0283-x

[136] Prueksaritanont T, Tang C. ADME of biologics-what have we learned from small molecules? The AAPS Journal. 2012;**14**:410-419. DOI: 10.1208/s12248-012-9353-6

[137] Ortho Multicenter Transplant Study Group. A randomized clinical trial of OKT3 monoclonal antibody for acute rejection of cadaveric renal transplants. The New England Journal of Medicine. 1985;**313**:337-342. DOI: 10.1056/NEJM198508083130601

[138] Di L. Strategic approaches to optimizing peptide ADME properties. The AAPS Journal. 2015;**17**:134-143. DOI: 10.1208/s12248-014-9687-3

[139] Shi S. Biologics: An update and challenge of their pharmacokinetics. Current Drug Metabolism. 2014;**15**:271-290

[140] Khawli LA, Goswami S, Hutchinson R, Kwong ZW, Yang J, Wang X, et al. Charge variants in IgG1: Isolation, characterization, in vitro binding properties and pharmacokinetics in rats. MAbs. 2010;**2**:613-624. DOI: 10.4161/mabs.2.6.13333

[141] Haraldsson B, Nyström J, Deen WM. Properties of the glomerular barrier and mechanisms of proteinuria. Physiological Reviews. 2008;**88**:451-487. DOI: 10.1152/physrev.00055.2006

[142] Suryadevara CM, Gedeon PC, Sanchez-Perez L, Verla T, Alvarez-Breckenridge C, Choi BD, et al. Are BiTEs the "missing link" in cancer therapy? Oncoimmunology. 2015;**4**:e1008339. DOI: 10.1080/2162402X.2015.1008339

[143] Schooltink H, Rose-John S. Cytokines as therapeutic drugs.

Journal of Interferon & Cytokine Research. 2002;**22**:505-516. DOI: 10.1089/10799900252981981

[144] Hoy SM. Patisiran: First global approval. Drugs. 2018;**78**:1625-1631. DOI: 10.1007/s40265-018-0983-6

[145] Stein CA, Castanotto D. FDA-approved oligonucleotide therapies in 2017. Molecular Therapy. 2017;**25**: 1069-1075. DOI: 10.1016/j. ymthe.2017.03.023

[146] Roy A, Nair S, Sen N, Soni N, Madhusudhan MS. In silico methods for design of biological therapeutics. Methods. 2017;**131**:33-65. DOI: 10.1016/j. ymeth.2017.09.008

[147] Sormanni P, Aprile FA, Vendruscolo M. Third generation antibody discovery methods: In silico rational design. Chemical Society Reviews. 2018;**47**:9137-9157. DOI: 10.1039/c8cs00523k

[148] Tibbitts J, Canter D, Graff R, Smith A, Khawli LA. Key factors influencing ADME properties of therapeutic proteins: A need for ADME characterization in drug discovery and development. MAbs. 2016;**8**:229-245. DOI: 10.1080/19420862.2015.1115937

[149] Lockwood SY, Meisel JE, Monsma FJ, Spence DM. A diffusion-based and dynamic 3D-printed device that enables parallel in vitro pharmacokinetic profiling of molecules. Analytical Chemistry. 2016;**88**:1864-1870. DOI: 10.1021/acs.analchem.5b04270

[150] Gloede J, Scheerans C, Derendorf H, Kloft C. In vitro pharmacodynamic models to determine the effect of antibacterial drugs. The Journal of Antimicrobial Chemotherapy. 2010;**65**:186-201. DOI: 10.1093/jac/ dkp434

[151] Ferreira LLG, Andricopulo AD. ADMET modeling approaches in drug discovery. Drug Discovery Today. 2019. DOI: 10.1016/j.drudis.2019.03.015. [Epub ahead of print]