Review

In Silico ADME Methods Used in the Evaluation of Natural Products

Robert Ancuceanu , Beatrice Elena Lascu * , Doina Drăgănescu * , Mihaela Dinu

Posted Date: 31 January 2025

doi: 10.20944/preprints202501.2308.v1

Keywords: ADME; in silico methods; natural compounds; quantum mechanics; molecular docking; pharmacophore modeling; QSAR; molecular dynamics; PBPK

Review

In Silico ADME Methods Used in the Evaluation of Natural Products

Robert Ancuceanu [1], Beatrice Elena Lascu [1,*], Doina Drăgănescu [2,*] and Mihaela Dinu [1]

[1] Carol Davila University of Medicine and Pharmacy, Faculty of Pharmacy, Department of Pharmaceutical Botany and Cell Biology; robert.ancuceanu@umfcd.ro (R.A.); mihaela.dinu@umfcd.ro (M.D.)
[2] Carol Davila University of Medicine and Pharmacy, Faculty of Pharmacy, Department of Pharmaceutical Physics and Informatics
* Correspondence: beatrice-elena.lascu@drd.umfcd.ro (B.E.L.); doina.draganescu@umfcd.ro (D.D.)

Abstract: The pharmaceutical industry faces significant challenges when promising drug candidates fail during development due to suboptimal ADME (absorption, distribution, metabolism, excretion) properties or toxicity concerns. Natural compounds are subject to the same pharmacokinetic considerations. In silico approaches offer a compelling advantage: they eliminate the need for phyasical samples and laboratory facilities, while providing rapid and cost-effective alternatives to expensive and time-consuming experimental testing. Computational methods can often effectively address common challenges associated with natural compounds, such as chemical instability and poor solubility. Through a review of the relevant scientific literature, we present a comprehensive analysis of in silico methods and tools used for ADME prediction, specifically examining their application to natural compounds. Whereas we focus on identifying the predominant computational approaches applicable to natural compounds, these tools were developed for conventional drug discovery and are of general use. We examine an array of computational approaches for evaluating natural compounds, including fundamental methods like quantum mechanics calculations, molecular docking, and pharmacophore modeling, as well as more complex techniques such as QSAR analysis, molecular dynamics simulations, and PBPK modeling.

Keywords: ADME; in silico methods; natural compounds; quantum mechanics; molecular docking; pharmacophore modeling; QSAR; molecular dynamics; PBPK

1. Introduction

Many promising drug candidates fail to reach the market due to poor ADME characteristics (how the drug is absorbed, transported throughout the body, metabolized, and excreted) and safety concerns about toxicity. When these problems are uncovered late in drug development, pharmaceutical companies risk large financial losses since they have already invested heavily in clinical trials and other costly research processes. To address this dilemma, the pharmaceutical industry has significantly changed its strategy in the last decades. Companies are increasingly performing extensive ADMET (absorption, distribution, metabolism, excretion, and toxicity) screening considerably earlier in the drug discovery process. This early screening technique identifies and eliminates problematic compounds before they enter costly development phases, thereby saving money and boosting medication development efficiency [1,2]. The presence of pan-assay interference compounds (PAINS) is an additional significant concern. Deceptive findings in research tests can be generated by these challenging compounds, which appear to be active against a diverse array of targets when they are not. It is a significant waste of time and money to investigate these "frequent hitters" in the context of research and development [2,3].

Compared to synthetic molecules, natural compounds possess unique properties that influence drug discovery. They are more structurally diverse and complex, tend to be larger, contain a greater amount of oxygen (and less nitrogen, sulfur and halogens), more chiral centers, less aromatic rings, and are more water-soluble. This provides them with a distinctive potential as drugs, even when they do not adhere to the conventional principles for drug-like properties, such as Lipinski's rule of five [4,5]. Nevertheless, the utilization of natural compounds in the drug discovery process is hindered by a number of obstacles. It can be challenging to test natural extracts, identify the active constituents

 doi:10.20944/preprints202501.2308.v1

in complex mixtures, obtain sufficient material from nature, and protect discoveries with patents. The pharmaceutical industry has experienced a decline in drug discovery that concentrates on natural products as a result of these obstacles [4,6]. Nevertheless, the discovery and development of natural product-based drugs are being substantially improved by recent developments in biosynthetic engineering technologies, chemical synthesis, as well as "smart screening" techniques or use of robotic equipment in chemical separations and progresses in structural analysis. Complex natural compound scaffolds that were previously considered inaccessible can now be optimized with these technologies. This advancement enables the enrichment of screening libraries with natural products, hybrids derived from natural products, natural product analogs and natural products inspired molecules, as well as superior structure functionalization approaches, including late-stage functionalization, for the optimization of natural product leads [4,6–8]. In addition to their potential applications in drug discovery, numerous bioactive natural compounds are essential components of our daily diets. Gaining an understanding of their properties is essential for the advancement of human health, as it facilitates the creation of enhanced foods, such as functional foods and dietary supplements [9].

Many of the challenges applicable in the understanding of the pharmacological or biological properties of natural compounds are also relevant when exploring their ADME properties. For instance, often the available quantities of natural products are limited, and while numerous plant-derived natural products have been isolated and characterized, the amounts available are frequently insufficient for comprehensive ADME testing [10]. Using in silico methods from this point of view has a great advantage as they require no physical sample (not even picograms are necessary once the structural formula is available) or laboratory infrastructure. Besides (and this is an aspect relevant for all products of pharmaceutical interest, irrespective of their origin), the experimental assessment of ADME properties of a substance is costly and time consuming [11], whereas the use of in silico tools is usually very cheap.

Natural compounds are often highly sensitive to environmental factors such as high temperature, moisture, light exposure, oxygen, or pH variations. They may also be volatile or react with other substances and stability issues result in limited shelf-life, making it difficult to develop stable commercial products that maintain their essential properties over time [12]. Besides, many natural compounds may be degraded by stomach acid or undergo extensive metabolism in the liver before reaching their target sites in the body (first-pass metabolism) [13]. This is also a serious issue in assessing their ADME features in animal experiments, whereas using very cheap but valid in silico tools is very useful. An additional challenge for exploring ADME properties in wet lab experiments consists in the low aqueous solubility of natural compounds [14] , which limits the ability to effectively deliver the compound to the biological system examined. Moreover, the growing need to minimize animal use in medical development and research highlights the increasing significance of in silico tools [15].

The prediction of physicochemical properties for understanding ADME characteristics has deep roots in computational chemistry. A pioneering development was the calculated partition coefficient (cLogP), introduced by Hansch, Iwasa, and Fujita in the 1950s and 1960s, which remains one of the most referenced physicochemical parameters. A major breakthrough came with the introduction of the Lipinski's "Rule of 5," which was based on the fact that straightforward molecular descriptors – including molecular weight and the number of oxygen, nitrogen, and hydrogen bonds – strongly influence a drug's absorption and membrane permeation. The field continued to evolve with Leeson and Springthorpe's criterion (cLogP > 4.5-5), which was later enhanced by new concepts such as lipophilicity ligand efficiency (LLE). Further advances included the development of the solubility forecast index (SFI) and the property forecast index (PFI). These developments were complemented by extensive analyses of clinical drug candidates, which helped refine and update Lipinski's original rules to better predict drug-like properties [16]. In the last decades substantial efforts have been deployed and marked progress has been made in the field of developing in silico methods and tools for the ADME assessment of new chemical compounds.

In this paper, we have proposed a synthesis of the scientific literature in order to understand the in silico methods and tools available for the prediction of ADME properties, with a focus on those that have been used to evaluate natural compounds. We were interested in understanding which of the available methods and tools have been most frequently used in this field, although such methods and tools have been developed for drug discovery purposes in general and not specifically for natural products. We focus here exclusively on ADME (Absorption, Distribution, Metabolism, and Excretion) rather than ADMET. While both abbreviations are commonly used, our emphasis on ADME is related

 doi:10.20944/preprints202501.2308.v1

to its direct relevance to pharmacokinetics. The 'T' in ADMET primarily addresses safety concerns (a rather distinct field of study) and this limited focus on ADME allows for a clearer focus on the pharmacokinetic aspects of drug development.

2. In silico methods available for ADME predictions

2.1. Quantum mechanics (QM) methods

Early studies of how drugs interact with receptors attempted to use quantum mechanics but these calculations were limited because they required intensive computational resources. Thanks to significant advances in computer speed and new software, quantum mechanics calculations are now used more regularly to study drug-related problems. In the past three decades, such calculations have become much more common in studying ADME properties [17].

Quantum chemistry states that electrons in atoms cannot have any energy level or be located anywhere in space. Their energy and location are determined by the Schrödinger equation. QM tries to solve this equation, often using approximations. The simplest of these approximations are called *semiempirical methods*. They were originally implemented in computer programs written in FORTRAN and today these semiempirical approaches are widely used and are included in many different computational chemistry software packages [16]. QM is capable of forecasting the molecular properties of a novel compound in the absence of any prior knowledge or parameterization [18]. Semiempirical methods can perform calculations very quickly on modern computers because they use a simplified approach to handle electron interactions and incorporate experimental data, such as ionization energy, bonding energy, and transition energy. These methods achieve their speed by replacing certain electron interaction calculations (e.g., two-electron repulsion integrals) with pre-determined empirical parameters (all QM methods use approximations, but these approximations are most pronounced in semiempirical methods). This simplification makes it possible to perform calculations on very large molecular systems, such as entire proteins, which would be computationally impossible with more detailed methods [16–18]. Among the most widely known semiempirical methods available in open-source and proprietary software are the following:

- MNDO (Modified Neglect of Diatomic Overlap) [19] and its improvements, AM1 (Austin Model 1) and PM3 (Parametric Method 3), the latter being an improvement of the former, with higher accuracy [18];
- PM6 and later PM7, which have provided corrections and improvements over AM1 and PM3 [20–22];
- the OMx family (orthogonalization-corrected methods - OM1, OM2, OM3), which have been shown to be roughly equivalent or slightly inferior in performance when compared with standard DFT methods (BPE, B3LYP), which will be discussed below [23];
- the density-functional tight-binding (DFTB) method which, with strong theoretical underpinnings, is closely connected to the DFT. However, it is as computationally efficient as empirical tight-binding methods and is used for big molecules, including those present in biological systems [24].

Semiempirical methods (MNDO, PM6) have been used up to date mostly to characterize the chemical stability and reactivity of natural compounds [25–27] or to optimize their chemical structures (e.g., for docking purposes) [28–31], but not to directly make inferences or predictions about specific ADME properties.

Ab initio *methods* (latin "ab initio": from the beginning represent a more advanced and accurate approach within QM calculations. Such methods provide a more complete mathematical description of how particles interact within molecules by using a special mathematical function called the Hamiltonian. The Hamiltonian calculates the total energy of a system by adding up two types of energy: the kinetic energy and potential energy for all electrons and nuclei in the molecule. Hartree-Fock (HF) is one of the most basic ab initio methods [16]. It is the foundation of molecular orbital (MO) theory and it is also referred to as the independent particle model or mean-field theory [18]. It uses a self-consistent field (SCF) approach, each electron being treated as moving within an average electrical field created by all the other electrons. However, it doesn't consider the specific, instantaneous positions of those other electrons. Because of this simplification, the repulsion between electrons is overestimated, because the averaging approach doesn't account for the fact that electrons naturally avoid each other in real systems [16].

 doi:10.20944/preprints202501.2308.v1

In 1934, Moller and Plesset published a short five-page paper describing how the Hartree-Fock (HF) method might be improved to account for electron-pair correlation using second-order perturbative theory (the interactions between electrons are called electronic correlations in quantum systems). This method is now called Møller-Plesset perturbation theory (MPPT or simply MP). It initially received little attention. Over the past 40 years, however, it has had a major impact on the development of ab initio quantum chemical methods [32,33].

Configuration interaction theory (CI) was the first method used in quantum chemistry to improve on the HF method. E.A. Hylleraas performed the first CI calculations for the helium atom in 1928. Although CI theory started early, it didn't develop much until the 1950s. This was due to a lack of computing power and the disruptions of World War II. In 1950, S.F. Boys did a small CI study on beryllium, which gave lower energy than the HF method and was computationally cheaper[34]. However, even today both MP and CI methods are still considered too computationally intensive [35]. Full CI is impractical for all but the smallest molecules. Therefore, truncated CI methods, which include only a few selected configurations, are more widely used. For example, CIS only includes configurations in which one electron is "excited" to a higher energy level, while CISD includes both single and double excitations from the H-F reference wave function. [36,37].

Coupled cluster (CC) methods are another improvement, that calculates electronic correlation energy by bulding upon the HF wave function and including increasing levels of electronic excitation (single, double, etc.). Higher-level excitations enhance accuracy but need substantially more computational power. CCSD(T), which contains single, double, and perturbative triple excitations, is regarded as the "gold standard" in computational quantum chemistry [38]. CC methods have very high accuracy but can only be applied to maximum 20 atoms [39], making them rarely suitable for ADME or drug development applications.

A variety of other ab initio methods have also been made available and are implemented in proprietary or open-source sofware, each with their own strengths, limitations and costs (e.g., CASSCF - complete active space self-consistent field, CASPT2 - complete active space perturbation theory, and other multireference methods [40], multi-configurational self-consistent field (MCSCF) [41], quantum Monte Carlo methods [42], explicitly correlated methods e.g .R12/F12 [43], the density matrix renormalization group method (DMRG) [44] , etc.).

Among ab initio methods for the study of large systems, such as those pertinent to ADME, density functional theory (DFT) remains the sole computationally viable approach. DFT's utility stems from ongoing enhancements in its approximation of complex electron interactions within the Kohn-Sham framework [45]. DFT directly calculates electron density instead of working with the complex many-electron. Due to the incomplete understanding of the theoretical framework for defining the energy function of electron density, Hartree-Fock approaches were deemed more dependable than Density Functional Theory. This, however, has lately altered, and DFT approaches are now gaining significant popularity [16]. DFT methods are not yet fast enough for extensive free energy calculations on systems beyond simple models. Nevertheless, they can serve as a benchmark for quicker methods and give important insights into how energy (enthalpy) and disorder (entropy) contribute to the overall stability of a system. For studying large molecules that interact through weak forces (non-covalent interactions), like proteins binding to drugs (protein/ligand complexes), a modified version of this method, named DFT-D (dispersion-corrected DFT) is now the standard method, but the method is still very slow [46]. Other widely used DFT approaches are those based on generalized gradient approximation (GGA and hybrid-GGA), such as EDF1, EDF2, BP, BP86, BLYP, M06, B97X-D [18], BPE [47] and B3LYP [48]. Among the extensive and expanding collection of DFT functionals, B3LYP has been the most frequently employed functional in the field of biochemical research for the past several decades [49].

Much like semiempirical methods, DFT (the most commonly used ab initio method for biological systems) appears to be the predominant choice for studying natural compounds. In the literature published up to date DFT has been used rather to assess the chemical stability and reactivity of certain compounds [50–55] or to optimize ligands before docking procedures [52,56], than for the evaluation proper of ADME properties.

Hybrid QM/MM (*quantum mechanics/molecular mechanics*) methods are an effective approach that creatively integrates quantum mechanical and classical molecular mechanical calculations. First introduced by Warshel and Levitt in 1976, this method has since been extensively used in computational molecular biology, drug design, nanotechnology, and quantum chemistry. It was also a key

 doi:10.20944/preprints202501.2308.v1

factor in the 2013 Nobel Prize in Chemistry [57]. In this approach, quantum mechanics is applied to a small region that requires great accuracy (usually the reactive site), while computationally efficient molecular mechanics modelling is used to simulate the adjacent environment. Since MM computations costs are virtually minimal in comparison to QM, the overall cost is mostly dictated by the size of the QM part. This makes QM/MM approaches especially useful for analyzing large biological systems such as enzymes or protein receptors, where just the active region requires quantum mechanical treatment and the protein framework can be addressed using molecular mechanics [16]. The hybrid QM/MM method enables researchers to simulate large macromolecular systems while maintaining the accuracy of detailed modeling. This approach helps overcome the size limitations of biological systems, which usually have a large or very large number of atoms (quantum mechanics methods are very accurate, but are feasible only for small chemical structures) [58].

The following factors must be addressed in order to utilize QM/MM calculations for biomacromolecules:

a) Defining the QM region: the accuracy of QM/MM calculations is generally enhanced when the QM region includes larger and less polar functional groups at the frontier between the QM and MM zones;
b) Handling the QM/MM interface: in biomacromolecular studies, covalent bonds at the QM/MM interface, such as those between protein residues in a receptor, require careful attention. Errors introduced by embedding the QM region into the classical force field of the bulk protein are minimized when the net atomic charges of the boundary atoms are low.
c) Addressing multiple local minima: the challenge of numerous local minima can be mitigated by performing extensive sampling of protein and ligand conformations and by ensuring the biomacromolecular system is accurately prepared and modeled [59].

While hybrid QM/MM methods are being used to some extent for studying the pharmacodynamics of natural products [60–63] and can be also employed for ADME assessment applications, we haven't found any reports of their use for this purpose.

2.2. Molecular docking

Since its emergence in the mid-1970s, docking has played a pivotal role in drug discovery and development, as well as in understanding how chemical compounds interact with their molecular targets. Over time, there has been a substantial rise in studies utilizing molecular docking to uncover the structural features required for effective ligand-receptor binding, alongside advancements in refining docking techniques for greater accuracy [64]. Molecular docking has emerged as the method of choice when researchers have access to the three-dimensional crystal structure of their target protein. It has seen widespread adoption thanks to two key developments: the dramatic increase in computing power and resources, coupled with the exponential growth of structural databases containing both small molecules and proteins. These factors have made docking both more accessible and more powerful as a research tool [65].

Molecular docking is a modeling technique that examines the interactions between a macromolecule (such as a protein, enzyme, or DNA) and a small molecule (ligand, i.e., the potential drug ingredient). Molecular docking primarily has three objectives: virtual screening, binding affinity calculation, and prediction of the specific spatial position, conformation and orientation of the ligand [66]. Sampling in molecular docking is challenging due to the vast conformational space, which includes rotational, translational, and internal flexibility of both the ligand and protein, as well as, in some cases, solvent effects. Exhaustively exploring all possible conformations and orientations within the computational time required for virtual screening is currently unfeasible. As a result, finding efficient sampling methods in docking studies is still an area of active research [65].

Docking studies can be categorized into three types based on molecular flexibility: rigid docking (both molecules fixed), flexible-rigid docking (one molecule flexible, the other fixed), and flexible docking (both molecules flexible) [67]. The first docking methods were based on the idea that molecules have fixed shapes, like a lock and key. This meant that both the receptor (a protein) and the ligand (a drug) were treated as rigid. However, it was later accepted that molecules can change their shapes when they interact (the "induced-fit" model). Because the backbone of a molecule influences the positions of many side chains, allowing both the receptor and ligand to be flexible makes the docking calculation much more complex. However, these flexible docking methods are superior

 doi:10.20944/preprints202501.2308.v1

because they more accurately predict not only how molecules fit together but also how strongly they bind to each other [68].

Docking is not regarded mainly as a tool particularly apt for ADME purposes, and there have been literature reviews that have not even considered ADME as a potential application of docking [64,67]. Molecular docking can, however, be used to evaluate the interaction between natural compounds (as well as synthetic or semisynthetic ones) with a variety of proteins involved in different steps of the ADME processes. Thus, a variety of natural compounds have already been evaluated through molecular docking and reviewed for their ability to interact with various transporter proteins relevant for the absorption, excretion or cell efflux [69–71]. The effects of 75 flavonoids on the P-glycoprotein have also been evaluated through in vitro and in vivo models, after which the authors have used molecular docking to better understand the nature of the interactions between the five flavonoids that demonstrated significant inhibition and the transporter protein [72]. In addition to in vitro uptake tests, molecular docking was used to examine the selectivity of 10 natural-origin inhibitors of the organic cation transporter 1 (OTC1), and benzoylpaeoniflorin was found to be the most selective [73].

The interactions between natural products and proteins involved in the drug distribution step, such as albumin, α1-acid glycoprotein or hemoglobin can also be investigated in silico using molecular docking. This is how, for instance, the interactions between enantiomers of stipuol (a polyacetylene from *Panax notoginseng* (Burkill) F.H.Chen) and human serum albumin have been evaluated, the authors finding that the natural compounds bind at subdomain III of albumin [74].

In a similar way, the interactions between natural compounds and metabolizing enzymes such as the cytochrome P-450 family of proteins can be evaluated using molecular docking and a review of this approach for drugs in general (not necessarily natural products) has also been published [75]. Virtual screening models based on machine learning have been used to identify selective CYP1B1 inhibitors (different libraries of compounds, including one of natural products), and the compounds thus identified have then been evaluated for their selectivity against other isoforms by using molecular docking [76]. Inhibitors of other CYP450 isoforms have also been evaluated using molecular docking [77–79]. Drug transporters in the kidney are particularly relevant for drug excretion and the interactions of numerous natural compounds with such transporters have been explored through molecular docking [69].

2.3. Pharmacophore modeling

The notion of „pharmacophore" was pioneered by Ehrlich in 1909, who described it as "a molecular framework that carries (Gr. *phoros*) the essential features responsible for a drug's (Gr. *pharmacon*) biological activity" Over the past century, the concept has remained, but its scope and significance have been remarkably broadened. Currently (according to IUPAC) a pharmacophore is understood as the complex combination of spatial and electronic features that are essential for optimal interactions with a specific biological target, ultimately initiating or inhibiting a biological response [80]. The pharmacophore model relies on the interaction between ligands and receptors. These interactions encompass all information pertaining to the structural, spatial, and chemical features that underlie specific pharmacological effects. The interaction mostly entails non-covalent bonding, including hydrogen bonding, pi-pi stacking, and ion-dipole interactions, among others [81].

Pharmacophore models are often developed using chemical features found in common among known ligands for a specific target. These are known as ligand-based pharmacophores. Obviously, when only one or several ligands are known for a specific target, or when the available ligands are devoid of structural diversity, ligand-based pharmacophores cannot be used. In such cases, the so-called (protein) structure-based pharmacophore (SBP) models are used [82]. SBP modeling has become increasingly prominent in recent years, driven by the significant growth in high-resolution protein structures. As of January 2025, the Protein Data Bank (PDB) houses approximately 123,000 three-dimensional structures of biological macromolecules, primarily proteins, providing an unprecedented foundation for understanding these molecular therapeutic targets [83]. If experimental structural data for a particular protein are not available, computational methods such as homology modeling can provide an alternative approach to generate a 3D model. Machine learning techniques have also proven effective in predicting protein structures, AlphaFold2 being probably the most widely known example [84].

 doi:10.20944/preprints202501.2308.v1

Developing an effective pharmacophore model demands a deep knowledge of the target protein's structure and binding characteristics. Proteins commonly feature multiple binding sites, and within each site, different ligands can adopt various binding orientations by interacting with distinct regions of the binding pocket. This complexity is further increased by protein subfamily variations, such as those seen among Src kinases, where subtle structural differences between related proteins can significantly impact pharmacophore selectivity [85]. The initial step in structure-based pharmacophore modeling involves selecting and preparing the target protein structure (a step similar in approach and importance for other computational approaches, too, such as molecular docking). This is followed by identifying potential binding sites on the protein. A detailed analysis is then conducted to determine the complementary chemical properties and spatial arrangement of the amino acids at the binding site. Based on this information, pharmacophore features are generated and refined using specialized tools available within the selected software. Finally, the critical pharmacophore features responsible for the biological activity are identified [86]. A detailed description of these steps can be found in the referenced publications and interested readers are encouraged to consult these sources for further information [83,84].

A pharmacophore model has been used recently in the study of a number of natural compounds (mainly flavonoids) to explore their ability to inhibit organic anion transporter 3 (OAT3) [87], organic cation transporter 1 [88] and organic cation transporter 2 [89]. Although the focus on those transporters was rather related to safety endpoint (prevention of nephrotoxicity or hepatotoxicity), the same approach can be equally relevant for ADME purposes, as biological transporters are involved in ADME processes and drug interactions related to the excretion of certain active substances [90]. Pharmacophore models were also used to evaluate the effects of a number of flavonoids on urate transporter 1 (URAT1) [91].

Pharmacophore models have also been used to explore the effects of certain natural products on specific CYP450 fractions, such as those of saponins [92] or flavonoids [93] on CYP3A4 or those of a variety of natural compounds on CYP1A1 [94,95].

2.4. Quantitative Structure-Activity Relationship (QSAR) models

QSAR modeling is a computer-based method that establishes empirical correlations between chemical descriptor values derived from molecular structures and experimentally obtained properties or bioactivities. These models are then used to predict or design new chemicals with specific properties [96]. In other words, the chemical structure of a substance can be described by a series of numeric features named generically "molecular descriptors" (e.g., molecular weight, number of double bonds, percentage of oxygen atoms, and others much more sophisticated) and QSAR establishes mathematical relationships between these numeric values and certain biological activities of interest (for instance, the inhibition or not of a certain CYP450 isoform, the percentage of oral bioavailability etc).

The use of QSAR has undergone significant changes in recent decades, which include: (1) using more detailed descriptions of molecules, from simple one-dimensional (1D) representations to more complex multi-dimensional (nD) ones; and (2) employing more sophisticated methods, including machine learning, to correlate molecular structure with biological properties. Initially, QSAR studies were limited to small groups of closely related chemicals and relied on basic regression analysis. Now, QSAR can handle much larger and more diverse datasets, using a wide range of machine learning techniques for both modeling and virtual screening [97].

Based on the assumed relationship between the chemical descriptors and the biological effect modelled, QSAR models can be classified into linear (e.g., based on multiple linear regression or partial least squares) and non-linear (e.g., support vector machines or artificial neural networks) [98,99].

Based on the mode in which descriptors are computed and models are built, QSAR models with different dimensions are distinguished: 2D-QSAR, 3D-QSAR, and more recently 4D-, 5D-, 6D-QSAR models have been proposed. The 2D-QSAR models are based on descriptors computed based on 2D representation of ligand structures (but ignoring completely the 3D shape and spatial arrangement of atoms within the ligand) [100]. 3D-QSAR correlates biological target properties with molecular descriptors derived from 3D chemical structures; it uses probe-based sampling within a molecular lattice to compute the 3D molecular descriptors [101]. In a paper published in 1997, Hopfinger et al. proposed for the first time the concept of 4D-QSAR [102]. The fourth dimension, known as 'ensemble sampling,' consists in a set of molecular configurations, describing ligands in different molecular and

 doi:10.20944/preprints202501.2308.v1

spatial forms, such as conformations, orientations, stereoisomers, and protonation states [100,103]. First proposed by Vedani and Dobler in 2002 [104], the 5D-QSAR approach addresses the induced-fit phenomenon in the ligand-protein interactions by considering multiple conformations that a ligand can adopt when binding to its target. This flavor of QSAR, which builds upon 4D-QSAR, creates virtual representations of these different binding scenarios to improve the predictivity of the model(s) [105]. 6D-QSAR builds upon previous QSAR versions by adding a sixth dimension: solvation. It allows the simultaneous evaluation of multiple solvation models to better understand how solvents influence molecular interactions. This can be carried out explicitly by mapping solvent properties onto the molecular surface or implicitly by adjusting the contribution of solvation terms (like ligand desolvation and solvent elimination) for different models within the analysis. During the simulation, weights associated to different conformations are adjusted through an evolution process [106].

Mechanistic QSAR is an approach that employs parameters/descriptors that have a consistent and unambiguous meaning, aiming to provide insights into the underlying mechanisms of the predicted biological effects. The alternative, named sometimes non-mechanistic QSAR, prioritizes the development of equations with predictive value, even if the parameters used are not readily interpretable. This approach focuses on statistical correlations between molecular descriptors and biological activity, but it does not necessarily offer mechanisti insights into the relationship between the descriptors and the biological effects (e.g., Comparative Molecular Field Analysis - CoMFA) [107,108].

A global QSAR model has been defined as a model characterized by a dataset that is highly structurally diverse, which in turn reflects a variety of mechanistic actions among the chemical compounds of the dataset. In such a model statistical methods are employed to attempt to identify structure/activity patterns that are independent of the mechanistic differences [109]. It has been stated that despite the fact that statistics might indicate that these models are robust, they tend to exhibit relatively low predictivity. Local models, of the contrary, are characterized by structural or mechanistic similarity [109]. The heterogenous definition of global QSAR models is somewhat problematic, and it could gain clarity by distinguishing between structurally global models (models built on a variety of structurally diverse compounds) and mechanistically global models (models built on chemical compounds acting through a variety of mechanisms and on distinct biological targets). There are theoretical reasons to believe that the latter have poorer performances than the former.

Sometimes QSAR models are qualified by the nature of the chemical descriptors used in their construction, for instance “topological QSAR” models (based on topological descriptors) [110,111] or “quantum mechanical QSAR” models (based on quantum mechanical descriptors) [112]. The “hologram QSAR” technique involves the decomposition of each molecule in the dataset into a molecular hologram. This hologram is primarily composed of linear, branched, and overlapping fragments, which are then arranged in an array of fixed length (ranging from 53 to 401 bins). These bin values serve as X variables in QSAR modeling, capturing both the composition and topology of molecules. Several factors affect how the hologram is generated and how the resulting HQSAR models perform statistically, including: the length of the hologram, the size of fragments, and how fragments are distinguished (based on atom types, bonds, connections, hydrogen atoms, chirality, and groups of hydrogen bond donors or acceptors) [113].

QSAR models can also be classified based on the categorical or continuous nature of the predicted outcome, distinguishing between classification and regression models. Classification models use chemical information of the training data set to make predictions about an outcome category, most often using a binary category (e.g., active/inactive, orally bioavailable/non-bioavailable etc). Often such a classification involves an arbitrary conversion of a continuous variable in two categories at a certain threshold. It is argued that in such a case, “the predictive reliability of the resultant QSAR model will increase at the expense of the resolution of the prediction” [114]. Regression QSAR models make predictions about the outcome variable on a continuous scale (predicting, for instance, a numerical value of ki, IC50 or of a percentage for the oral bioavailability and not a mere category) [115].

Biological effects that can be predicted using QSAR models can be vastly diverse, from the effects on a specific biological target to the prediction of a specific adverse effect or a specific ADME feature. Among the in silico methods for ADME predictions, probably QSAR models have been used the most extensively, and generally the same models can be applied for both natural and synthetic organic chemical compounds, although often the specific use of such models for natural compounds has not been explicitly mentioned. In the case of intestinal absorption prediction, the performance of QSAR

 doi:10.20944/preprints202501.2308.v1

models evaluated in a large benchmark study was only slightly inferior to that of in vitro methods (83% of QSAR predictions and 87% of in vitro method predictions fell within 2-fold of observed values [116]. Similarly, an earlier study indicated that a pair of computer models achieved reliability comparable to the Caco-2 and 2/4/A1 cell lines, with one model predicting the absorption of a collection of 65 medicines nearly as well as in vivo absorption experiments conducted in rats [117].

Kansy et al. introduced the parallel artificial membrane permeation assay (PAMPA) in 1998 [118]. PAMPA is composed of lipophilic filters that are coated with lecithin in an organic solvent solution. Transcellular permeation is assessed by a rapid in vitro assay system that is suitable for high-throughput screening. Akamatsu et al. (2009) constructed an in silico prediction model of human oral absorption for potentially transported compounds by developing a QSAR model that describes correlations between chemical structure and PAMPA permeability [119]. This study has not mentioned explicitly the use of the QSAR models for natural compounds, but similar QSAR models based on PAMPA have been developed and one has been used to predict the gastro-intestinal absorption of flavonoids from *Silybum marianum* (L.) Gaertn. [120].

Oral clearance refers to the rate at which a substance/drug ingredient is removed from the body following oral administration. Boik and Newman (2008) developed three QSAR classification models based on data from human oral clearance, rodent LD50, and in-vitro cytotoxicity studies. Then they used these models to analyze over 115,000 natural compounds, resulting in the prediction that hundreds of these substances exhibit low systemic toxicity, low to moderate oral clearance, and manageable cytotoxicity levels [121]. There are numerous other QSAR/QSPR models that have been developed to predict oral absorption of drugs, although they have not necessarily been applied to natural compound datasets (Table 1).

Table 1. QSAR models for the prediction of the oral absorption of pharmaceuticals.

 doi:10.20944/preprints202501.2308.v1

Type of model	Dataset size (training, test sets)	Performance (best model)	Outcome variable	Reference
Regression QSAR	86 (67, 9, 10*)	RMS - 9.4% HIA units (training), 19.7% HIA units (CV), 16.0% HIA units (external set)	Human intestinal absorption (%)	[122]
Hologram QSAR, regression	638 (50, 128)	R^2 – 0.79, Q^2 – 0.63	Human intestinal absorption (%)	[123]
Classification	272 (232, 40)	Accuracy (train set) – 71%, accuracy (test set): 60%.	Bioavailability data in healthy human subjects (4 classes of bioavailability: class 1 (<20%), class 2 (20-49%), class 3 (50-79%), class 4 (80-100%).	[124]
Regression and classification models	458	Regression: R^2 – 0.60 Classification: CCR – 0.88, MCC – 0.75 (10-fold cross-validation)	Human intestinal absorption (%). Three ordinal classes of absorption (class 1 - >80%, class 2 – 30% - <80%, class 3 - < 30%).	[125]
Regression models based on Abraham descriptors	169 (38 + 131; 31+138)	0.85 (train set); 0.78 (cross-validation)	Human intestinal absorption (%)	[126]
Regression and classification	96 (67 + 9+12*)	RMS – 6.5 (train set), 27.7 (test set), 22.8 (external prediction set). For classification, sensitivity 100%, specificity 50%.	Human intestinal absorption (%). For classification purposes, a 50% HIA threshold was used to define two classes.	[125,127]
Classification QSAR, using structural descriptors	1262 (899+362)	AAE – 0.12 (12%); Accuracy: 79-86%	Human intestinal absorption (%, divided in six classes of about 16% per class)	[128]
Regression models using five classes of descriptors	169 (113 + 56)	R^2 – 0.86 (training set), 0.73 (test set)s	Human intestinal absorption (%)	[129]
Regression model using descriptors computed based on DFT	241 (38 + 203)	RMSE – 12.8 (% HIA) (15 on the entire test set)	Human intestinal absorption (%)	[130]
Classification QSAR using a variety of descriptor classes	141 (+ an external data set of 27 compounds)	Accuracy: 88.9% (external data set), 65.71% (10-fold CV)	Human intestinal absorption (%) (5 classes)	[131]
MI-QSAR (QSAR based on "descriptors explicitly derived from simulations of solutes [drugs] interacting with phospholipid membrane models")	188 (164 + 24)	R^2 = 0.68 (train set), 0.65 (test set).	Human intestinal absorption (%) and	[132]
Regression and classification models (using a variety of descriptor classes)	553 (455 +98)	R^2 - 0.76** (train set), R^2 - 0.79** (test set), AME[a] – 7.3% (test set), Accuracy > 96.8%.	Human intestinal absorption (%)	[133]
Multiple regression models using a variety of descriptors	552 (380+172)	R^2 – 0.64** (train set), R2 – 0.79** (test set)	Human intestinal absorption (%)	[134]
Regression models using descriptors computed with two commercial products and predicted pKa	567 (+25 + 22***)	R^2 for log P_{eff}[b] - 0.72-0.84; RMSE - 0.35–0.45 log units (equivalent to 2.24-2.82%)	Human intestinal absorption (%)	[135]
Classification QSAR using multiple classification algorithms and 166 descriptors	225 (158+67)	Accuracy – 94% (training set), 91% (test set)[c]. κ statistic – 0.58	Human intestinal absorption (%). Two classes: high (> 30%) and low (< 30%).	[136]
Classification QSAR using FP4 and MACCS fingerprints	578 (480+98, (+634***)	Accuracy – 98.5% (training set), 98.8% (test set), 94% (validation set)	Human intestinal absorption (%). Two classes: high (> 30%) and low (< 30%).	[137]

 doi:10.20944/preprints202501.2308.v1

Type of model	Dataset size (training, test sets)	Performance (best model)	Outcome variable	Reference
Regression and classification QSAR using topological descriptors (computed with the CODES program)	367 (202 + 165)[d]	R^2 = 0.93 (train set), Q^2 = 0.92 (LOO cross-validation). Global accuracy: 74%.	Human intestinal absorption (%). Three classes (cut-offs: 30%, 50% and 70%).	[138]
Classification and regression QSAR models build with different descriptors and algorithms	577 (78+489)	Accuracy: 99.37%, 99.58% (train set), 95.92%, 94.90% (test set). RMSE – 6.39 (train set), 5.71 (test), R^2 – 0.972 (train set), 0.953 (test set)	Human intestinal absorption (%). Two classes, using a 30% threshold.	[139]
Regression QSPR models using 2D and 3D descriptors	1272 (1017+255)	R^2 = 0.97, Q^2= 0.83, RMSE CV = 0.31 (training test), R^2 = 0.81, RMSE T = 0.31 (test set)	Caco-2 cell permeability (permeability coefficient of Caco-2 monolayer cell - Papp)	[140]
Classification and regression QSAR/QSPR models	141 (98, +43)	Accuracy: 0.77 (10-fold CV), 0.70 (external data set) R^2: 0.38 (training set), 0.05 (external data set)	Human intestinal absorption (%). Two classes, using an 85% threshold.	[141]
Regression QSAR using a variety of descriptors computed with the Dragon software	160 (90 + 30 + 40)	R^2 – 0.771 (training set), 0.716 (test set). RMSE – 0.182 (training set), 0.189 (test set)	Human intestinal absorption (%) – more precisely, log10 (HIA% + 10).	[142]
Regression QSAR using artificial neural networks	86 (67 + 9 + 10)	R^2 – 0.802 (test set); RMS – 0.59 (train set), RMS – 0.42 (test set).	Human intestinal absorption (%).	[143]
Regression QSAR using mainly structural descriptors	467 (417+50)	R^2 – 0.79 (train set), 0.79 (test set), RMSE – 12.3% HIA	Human intestinal absorption (%).	[144]

* Train set, test set, external prediction set; **The authors have reported the r value (not r^2); *** External validation data set; [a]AME – absolute mean error; [b]*In vivo* human jejunal permeability coefficients; [c]The accuracy is relatively high, but the data set was highly imbalanced (balanced accuracy would have been preferable). [d]The best performing model selected was trained on only 37 compounds.

QSAR models have also been developed to predict plasma protein binding [145–148] or the unbound fraction [149], the volume of distribution in humans, rodents or other mammals [150–152], the blood-brain barrier (BBB) permeability [153–155], inhibition of specific CYP450 fractions [156–159], elimination half-life (149), rodent (150,151) and human clearance (152), effect of substances on key transporters such as P-glycoprotein [160–162], organic anion transporters [163–165], organic cation transporters [166] etc. Some of the models referenced here have been specifically developed for natural products (for instance, Li et al. (2018) explored the quantitative relationship between the structural characteristics of flavonoids and their ability to inhibit CYP3A4 [158]), but most of the published models have been developed for medicines or organic substances irrespective of their natural or synthetic origin.

Certain molecular descriptors can be computed/estimated using molecular dynamic simulations, case in which QSAR and MD simulations are integrated in a single final model. For instance, Iyer et al. (2007) have used MD to simulate the behaviour of different ligands across a dimyristoylphosphatidylcholine (DMPC) monolayer membrane and compute a number of descriptors used then to build QSAR models predicting the intestinal oral absorption of the drugs included in the models. The test and train compound datasets included both natural and synthetic substances [132]. Other relevant descriptors can be computed using the DFT theory, case in which QSAR and DFT are integrated in a single model (see, for instance, ref. [130]). Another hybrid approach has been recently (year 2024) reported for 48 compounds, mostly natural (including alkaloids, flavonoids and coumarins). The authors built a QSPR model to predict the apparent permeability coefficient (Papp) (measured on Caco-2 cells), then used MD simulation to evaluate the molecular mechanisms involved in the absorption differences observed for ligustrazine and EGCG, as well as

 doi:10.20944/preprints202501.2308.v1

molecular docking to understand the interactions between the compounds and the P-glycoprotein [167].

2.5. Molecular dynamics (MD) simulations

X-ray crystallography provides static snapshots of molecules because they are locked in a crystal. This led to an initial perception that large molecules like enzymes and receptors were rigid. However, combining multiple snapshots often reveals that these molecules exist in different states, indicating they are actually dynamic and fluctuate between these states [168]. Molecular dynamics simulations use physics-based modeling to predict how individual atoms move within proteins and other molecular systems over time. These highly precise simulations can track atomic positions down to the femtosecond level, allowing researchers to acquire knowledge about important biological processes like protein folding and ligand binding. The simulations provide valuable insights into how biomolecules respond at the atomic scale to various changes like mutations or chemical modifications. This computational approach has become a powerful tool for understanding molecular behavior and interactions [169]. The method was for the first time applied to proteins at the end of the 1970s [170]. The applications of MD simulations in the fields of molecular biology and drug discovery are vast, but we are here focused on their potential use in making ADME predictions. It seems likely that the most important such applications consist in improving molecular docking results for the interactions between different ligands and the proteins of interest, relevant for pharmacokinetics purposes.

While molecular docking can predict how a protein and a ligand interact, the initial prediction of how the ligand is positioned and oriented within the protein's binding site is often approximate. To get a more accurate picture of the interaction, further refinement is necessary. MD simulations can refine the docking results by minimizing steric clashes (where atoms are too close together) and adjusting the initial binding position of the ligand [171].

To enhance the effectiveness of virtual screening, MD simulations can be combined with binding free energy calculations to evaluate the binding energies of small molecules with a target (thus understanding how strongly a ligand binds to the target of interest). These energies can then be used to rank candidates and refine the scoring of generated poses. Methods like MM-PBSA (Molecular Mechanics Poisson-Boltzmann Surface Area) and MM-GBSA (Molecular Mechanics Generalized Born Surface Area) enable more precise calculations of the binding free energy between the ligand and receptor. Entropy changes are estimated through normal mode analysis, while enthalpy changes for the entire system are statistically derived from electrostatic, van der Waals, and solvation energies, all based on MD simulation trajectories. This approach allows for the accurate determination of binding free energies [171]. Whereas MM-PBSA and MM-GBSA are not themselves MD simulations, they use the MD-generated data to estimate more or less accurately the free energy and thus to better understand the ligand-target interaction.

All ligand-protein interactions that can be studied for ADME purposes using molecular docking can also be better understood using MD simulations, and the applications of the former are also applications of the latter. But MD simulations have also been used for certain processes relevant for ADME but less investigated with molecular docking. The most important such application is the study of membrane permeability. In 1994, Marrink and Berendsen performed the first atomistic simulation of water permeability in a lipid bilayer. Since then, hundreds of coarse-grained and atomistic simulations have been carried out to study the passive transport of permeants across different lipid bilayers and synthetic membranes [172]. Molecular simulations are useful for studying both synthetic and natural substances. Two studies specifically investigated how menthol enhances the skin's ability to absorb substances [173,174]. These studies showed that menthol readily penetrates the skin's protective outer layer (lipid bilayer), making its interior more fluid. However, at high concentrations, menthol can damage this layer. Borneol, another natural compound, at low concentrations helps substances penetrate the skin, but at high concentrations creates pores and unusual structures called reverse micelles, a mechanism similar to how some antimicrobial peptides disrupt cell membranes (the "carpet mechanism") [172,175,176].

2.6. Physiologically-Based Pharmacokinetics (PBPK) Modeling

PBPK models simulate how a drug and its metabolites change in concentration over time in the blood and target organs. Building PBPK models, however, requires more effort and data compared to simpler predictive methods. They necessitate estimating a larger number of parameters and

 doi:10.20944/preprints202501.2308.v1

demand extensive knowledge of the body's physiology and the specific drug's properties [177]. The model's parameters are obtained either from non-clinical experimental data, clinical trials or calculated based on appropriately validated formulas. Broadly, these parameters can be categorized into two distinct groups: drug-specific parameters and organism-specific parameters. Although a PBPK model is usually very complex, with hundreds of ordinary differential equations, a large number of parameters, particularly organism-specific ones, are already known (e.g., organ blood flows, tissue volumes, membrane permeabilities etc); therefore, for a new medicine, often less than five drug-parameters are needed (e.g., clearance, volume of distribution, binding parameters etc) [178].

PBPK models estimate drug kinetics in one or several organs, necessitating a submodel for each organ or tissue considered and integrate then the organ submodels in a whole-body, comprehensive model. Drug transport occurs via blood circulation as determined by anatomical structures. PBPK strive to replicate authentic, quantifiable physiological and/or pharmacokinetic processes instead of more abstract rate constants and volumes. Parameters used in PBPK models may be 'intrinsic' covariates, such as weight, body surface area or glomerular filtration rate, rather than covariates identified post-hoc to improve model fit (such as, for instance, the patient age, in the absence of a clear mechanism or explanation of why it would improve fit) [179]. A PBPK model for a specific medicine can be validated using both in vitro data and in vivo oral pharmacokinetic data from rats. To predict the human oral pharmacokinetic profile, the physiological and in vitro parameters from the rat model can be substituted with human-specific parameters, following a successful prediction of the rat's in vivo oral pharmacokinetics from in vitro data [180].

Compartmental models, widely used in the traditional pharmacokinetic research, are built using a "top-down" method, where all the model's information is derived from a specific set of pharmacokinetic and covariate data. Conversely, PBPK models tend to use a "bottom-up" approach, where the model's information is derived from in vitro data and pre-existing knowledge of physiological and pharmacological mechanisms involved in the drug pathways through the body, the model parameters being thus derived from mechanistic principles (the focus is not so much on merely fitting the data as to providing a mechanistic understanding of the drug behavior) [179].

PBPK models offer valuable predictions for various drug development purposes:

- At the clinical trial design stage, they can predict how drug formulation and food intake will influence pharmacokinetics, guiding initial human studies and forecast drug-drug interactions mediated by enzymes or transporters, informing inclusion/exclusion criteria, dose selection, and potentially waiving unnecessary clinical interaction studies or studies where enrolling subjects is anticipated to be difficult;
- They can predict appropriate dosing regimens for different pediatric subsets, from newborns to adolescents, by enabling informed selection of sampling timepoints and proposing suitable doses.
- They can predict exposure to the drug in patients with impaired renal or hepatic function, guiding organ impairment studies or supporting decisions to waive such studies.
- They can estimate the drug disposition in the mother and fetus, aiding in optimizing the therapeutic benefit-risk ratio during pregnancy.
- They can predict pH-mediated drug-drug interactions in patients receiving proton pump inhibitors or antacids, guiding formulation development and efforts to minimize food-drug interactions [181].

It is not feasible to cover technical details on building PBPK within the limited space of this paper, but they are comprehensively addressed in the literature, with several high-quality tutorials providing valuable guidance on different aspects of the modelling process [178,182,183].

A limited number of studies have used PBPK modelling up to date to approach ADME properties for natural products. For instance, a recent study reported its use to estimate the biodistribution of *Centella asiatica* administered orally in a rat model [184]. Another study used a PBPK model to simulate the drug-drug interaction between aconitine and the glycyrrhizic acid, using parameters derived from in vitro experiments. The model simulations were validated against experimental observations to evaluate their accuracy. The analysis focused on understanding how CYP3A and P-glycoprotein contribute to the reduction in aconitine toxicity induced by the glycyrrhizic acid [185]. Similarly, PBPK models were used to explore the interactions between silybin A and losartan [186], between schisandrin A, schisantherin A and cyclophosphamide [187], the drug-drug interactions involving bergamottin (from grapefruit juice), curcumin (from turmeric) and hyperforin (from St. John's wort) on the one hand, and anticancer drugs on the other [188], or the interaction between

 doi:10.20944/preprints202501.2308.v1

rifampicin and retrorsine (a pyrolizzidine alkaloid occurring occasionally as a contaminant in herbal teas) [189]. A PBPK model was used to show that cichoriin is capable of reaching high concentration levels in the lung (when administered intravenously) and thus, that it is a potential candidate for development against COVID-19 [190] and another to predict the plasma concentration variation for schaftoside, subsequent to the ingestion of a capsule containing total flavonoids extracted from *Desmodium styracifolium* [191].

3. Comprehensive ADME Tools

Whereas many of the approaches discussed above involve the building and validation of a single model or a family of models that are then used "inhouse" for the virtual screening of natural or synthetic compounds, there are a number of software tools that have been developed specifically for in silico ADME predictions. They are either available as server services or as standalone products, for free or for a fee. Although a larger number of commercially or freely available software is available for ADME predictions, our search of PubMed using appropriate keywords ("natural" OR "herbal" plus the name of each software product, and when not very specific, also the name of the manufacturing company) has shown that for natural products only a part of the available software has been used, and only a few have been applied in multiple papers in the field of natural products: SwissADME webserver[192] (> 250 papers), pkCSM [193] (> 60 papers), ProTox-II/Protox 3.0 [194,195] (> 45 papers), QikProp [196] (part of the Schrödinger Suite – > 40 papers), ADMETSAR [197] (> 39 papers), ADMETlab 2.0/3.0 [198,199] (> 30 papers), ADMET Predictor [200] (> 12 papers), FAF-Drugs4 [201] (6 papers), ACD/Percepta [202] (4 papers), ADMET module of Biovia Discovery Studio (2 paper [71]). Admetboost [203] and Interpretable-ADMET [204] have been relatively recently published and made available for free, but no paper has apparently used them as yet in the assessment of natural products. We have limited the discussion to general-purpose ADMET tools, but specialized tools for specific ADMET are also available; for instance, a review of software applications used to specifically predict CYP450-mediated drug metabolism is available elsewhere [205].

In what follows we present in more details the most widely used such free application (SwissADME) and the most widely used commercial application (QikProp). We will then provide concise overviews of other commonly utilized tools, including pkCSM, Protox 3.0, ADMETSAR, and ADMETlab.

3.1. SwissADME

The SwissADME web application (available http://www.swissadme.ch/) is a free and user-friendly platform used to predict ADME properties of potential drug candidates. No computer-aided drug design (CADD) expertise is required. Its key features include:

- Simple Submission: single or multiple molecules can be uploaded for analysis with ease.
- Clear Results: straightforward visualizations are available for the predictions made for each molecule.
- Data Sharing: the results for individual molecules can be saved and shared, or can be used to generate comprehensive, interactive graphs for deeper analysis.
- Integrated CADD Tools: the application seamlessly connects to the SwissDrugDesign workspace, granting the user access to a suite of other useful applications developed by the SIB Swiss Institute of Bioinformatics, to identify promising drug candidates based on their similarity to known active compounds (ligand-based virtual screening), to predict potential biological targets for a specific ligand, to perform ligand-protein docking, to create bioisosteric replacement of functional groups, or to perform molecular mechanics (e.g., to analyze the 3D structure and energy of a molecule) [192].

The following properties and values are predicted by SwissADME:

3.1.1. Biovailability radar

A bioavailability radar map is generated by the server application for a swift evaluation of a molecule drug-likeness. For this purpose, six physicochemical features are used: size, flexibility, polarity, lipophilicity, solubility, and saturation [192].

3.1.2. Physicochemical properties

 doi:10.20944/preprints202501.2308.v1

A small number of physicochemical properties are computed by the application using OpenBabel (molecular weight, molecular refractivity, counts of heavy and aromatic atoms, counts of H-bonds donors and acceptors, as well as the polar surface area, PSA, the latter being known to be often a good predictor of oral absorption or of the blodd-brain barrier permeability) [192].

3.1.3. Lipophilicity

The n-octanol - water partition coefficient (log Po/w) is a classic measure of lipophilicity, playing a crucial role in the pharmacokinetic behavior of drugs and drug candidates, being regarded as "the single most informative and successful physicochemical property in medicinal chemistry" [206]. SwissADME offers five predictive models to calculate this parameter and a "consensus log Po/w" is derived by taking the arithmetic mean of the predictions generated by these five approaches, but the five individual results are also returned [192].

3.1.4. Water solubility

Three methods are used by the application to predict water solubility, presenting the results in a variety of formats: logarithmic scale (the decimal logarithm of the molar solubility - log S -, allowing comparisons of values across a wide range), molar and mass concentrations (mol/l and mg/ml), and qualitative classes (e.g., "soluble", "moderately soluble", "poorly soluble" etc) [192]

3.1.5. The skin permeability coefficient

Skin permeability is crucial in a variety of applications, such as the risk evaluation of hazardous compounds, the design of skin cream products, and transdermal medication delivery [207]. The assessment of skin permeability often involves measuring the skin permeation coefficient (Kp) of a chemical through the stratum corneum, which is the top layer of the epidermis [208]. SwissADME uses a model adapted from Potts and Guy (1992), based primarily on the molecule lipophilicity and its size [192,209]. For this model, the authors have claimed an r^2=0.67, but in a benchmark study on 30 molecules the correlation coefficient was only 0.544, whereas QikProp predictions achieved a correlation coefficient slightly higher (0.612) [210].

3.1.6. Passive HIA and BBB permeability

Both these PK variables are predicted by SwissADME using the BOILED-Egg model, which is not based on regression, but rather on a classification algorithm, providing results in a binary form (GI absorption high vs. low, and permeant vs. non-permeant, respectively) [211].

3.1.7. Pgp substrate property

P-glycoprotein (ABCB1) is an ATP-dependent efflux pump that transports a large variety of structurally diverse hydrophobic and amphipathic compounds. These substances include pharmaceuticals, peptides and lipid-type molecules. The pump plays a critical role in the efflux of these compounds from cells due to its broad substrate specificity [212]. SwissADME uses a machine-learning algorithm based on support-vector machines (SVM) to classify the evaluated compounds as substrates of Pgp or non-substrates [213].

3.1.8. CYP fraction substrate

The different fractions of CYP450 enzyme system metabolizes over two-thirds of xenobiotics, therefore being able to predict whether a potential drug or toxic will be a substrate of one or several of those fractions is important and useful [205]. Similarly to the Pgp substrate predictions, SwissADME uses several classification models to predict whether a substance is or not a potential inhibitor of five CYP450 fractions: CYP1A2, CYP2C19, CYP2C9, CYP2D6, and CYP3A4 [192].

3.1.9. Drug likeness and medicinal chemistry properties

Besides the ADME properties mentioned above, the web server application also computes several drug-likeness scores widely used in the field of drug discovery (Lipinski [214], Ghose [215], Veber [216], Egan [217], Muegge [218]), as well as the Abbot bioavailability score [219]. Such scores are relevant in the understanding ADME properties of a substance, because compounds with appropriate scores of drug-likenss/druggability also have desirable ADME characteristics [11]. Several

 doi:10.20944/preprints202501.2308.v1

medicinal chemistry alerts, filters or scores are also implemented: PAINS alerts [220], structural alerts (e.g., potentially mutagenic groups, other groups with potential negative effects on PK or very reactive groups) [221], a lead-like filter [222], and a synthetic accessibility score (ranging between 1 and 10, from very easy to very difficult).

3.2. QikProp

Schrödinger's proprietary software, QikProp, is widely used to generate ADME predictions for various drug candidates, including natural products [196], but its use seems to become outpaced (at least in the field of natural compounds) by free webservers such as SwissADME or pkCSM. The software estimates a total of 46 molecular descriptors relevant to understanding ADME properties. One of the descriptors calculated by QikProp is a "drug-likeness parameter" represented by a "star" rating system. This parameter serves as an overall ADME compliance score, where the number of "stars" assigned corresponds to the number of descriptors that that are outside the range of values that are considered optimal for 95% of known medications. The higher the number of "stars," the greater the deviation from the typical drug-like profile, potentially indicating poorer ADME characteristics for the compound [196]. Jorgensen and Duffy [223–225] are credited as the creators of this approach [196]. This global ADME score generated by QikProp is used to help prioritize and select promising drug candidates for further development, based on their predicted ADME suitability. QikProp predictions are calculation-based, as opposed to fragment-based. Fragment-based methods may be problematic when they fail to identify specific components of a structure or encounter unfamiliar fragment interactions. In contrast, QikProp will calculate properties based on the entire molecule. This allows QikProp to make its predictions on scaffolds that are new and unfamiliar [226].

Among the calculated descriptors are the following [196]:

3.2.1. Total Solvent-Accessible Molecular Surface (Smol or SASA, solvent-accessible surface area)

It is expressed in $Å^3$ and is calculated using a probe radius of 1.4 Å (the approximate radius of a water molecule). Approximately 95% of pharmaceutical compounds exhibit S mol values between 300–1000 $Å^2$. It quantifies the surface area of a molecule accessible to a solvent (usually water) and reflects the molecule's hydrogen bonding capacity. It is an important property useful and used in predicting solubility and permeability of a potential drug [227,228].

3.2.2. Hydrophobic Portion of the Solvent-Accessible Molecular Surface (Smol, hfob)

It is expressed in $Å^3$ and is calculated using a probe radius of 1.4 Å. Approximately 95% of pharmaceutical compounds exhibit S mol values between 0 and 750 $Å^2$. Also known as FOSA, it describes the nonpolar interactions of the ligand, being useful in estimating the permeability capability of a small molecule [229].

3.2.3. Total Volume of Molecule Enclosed by Solvent-Accessible Molecular Surface (Vmol, hfob).

It is a measure used to quantify the volume occupied by a molecule within its solvent-accessible surface. It is expressed in $Å^3$ and computed using a probe radius of 1.4 Å. Approximately 95% of pharmaceutical compounds exhibit V mol, hfob values between 500 and 2000 $Å^3$ [196,230].

3.2.4. Logarithm of Aqueous Solubility (log Swat)

It is the logarithm to base 10 of a molecule's solubility, measured in mol/L. In drug development, the ability to dissolve in water (aqueous solubility) is a key physical and chemical property that needs to be optimized, particularly for aspects 'A' and 'D' in ADME-Tox. This solubility, along with the capacity to penetrate membranes (membrane permeability), plays a major role in determining how well a drug is absorbed into the body when taken orally (oral bioavailability) [231]. Approximately 95% of pharmaceutical compounds exhibit log S wat values between -6.0 and 0.5 [196,230].

3.2.5. Predicted octanol / water partition coefficient (QPlogPo/w).

Recommended values for this parameter (see als section 3.1.3), as satisfied by 95% of pharmaceutical compounds, are between –2.0 and – 6.5 [232].

 doi:10.20944/preprints202501.2308.v1

3.2.6. Logarithm of Predicted Binding Constant to Human Serum Albumin (log K HSA or logKhsa)

Approximately 95% of pharmaceutical compounds exhibit logKhsa values between -1.5 and 1.2. It is used to predict a compound's binding affinity to plasma proteins such globulins, human serum albumin, lipoproteins, and glycoproteins [229].

3.2.7. Logarithm of Predicted Blood/Brain Barrier Partition Coefficient (log B/B).

In pharmacokinetics, BBB distribution is quantitatively evaluated using the blood-brain partition coefficient, mathematically represented as the logarithm of the concentration ratio between brain tissue and blood (and referred to as log B/B) [233]. A positive logB/B signifies a greater concentration of the compound in brain tissue compared to blood, suggesting effective brain penetration, whereas a negative logB/B indicates limited brain accessibility. Approximately 95% of pharmaceutical compounds exhibit log B/B values between -3.0 and 1.0 [229] (this means that drugs with excellent blood-brain barrier penetration can reach levels about 10 times higher in the brain than in the blood, whereas for drugs with low blood-brain barrier penetration, the concentration in the brain is only 1/1000th of the concentration in the blood).

3.2.8. Predicted Apparent Caco-2 Cell Membrane Permeability (BIP caco-2)

Permeability assays are an essential part of ADME profiling, as adequate permeability across the intestinal membranes can improve oral absorption, which is a highly sought after property in most drug discovery initiatives. The Caco-2 cell permeability model is useful because it is human-derived and expresses transporters common to the human gastrointestinal tract [234]. The model can predict how well a drug can pass through the gut wall into the bloodstream. This parameter reflects (mainly) non-active transport across gut barrier, and is influenced by factors such as the drug's solubility and its ability to cross cell membranes. However, the absorption of drugs can also be affected by interactions with transporters and metabolizing enzymes present in the gut wall [235]. Approximately 95% of pharmaceutical compounds exhibit BIP caco-2 values between less than 5 (low) and higher than 100 (high). It has been stated that if BIP caco-2 is > 22 nm/s, there is a higher probability that a drug will be orally bioavailable [196]. However, this descriptor, even if experimentally measured (computational methods may have larger margins of error), is only of a limited use in predicting oral bioavailability. Broadly speaking, it seems that once the permeation rate exceeds a certain threshold—approximately 100 nm/s—it is no longer the primary factor limiting oral bioavailability. It has been estimated that compounds with permeation rates below 100 nm/s exhibit an average oral bioavailability of 13%, whereas those with rates above 100 nm/s have an average of 35% [236]. In vitro apparent permeability (Papp) measurements done using Caco-2 or MDCK cells (discussed below) cannot be directly translated to in situ effective permeability (Peff) values due to inherent methodological and physiological differences (but mathematical models can use those values in predicting relatively well the human absorption properties) [237].

3.2.9. Predicted Apparent Madin-Darby Canine Kidney (MDCK) Cell Permeability (QPMDCK)

Similar to the BIP Caco-2, MDCK cell permeability serves as a valuable tool for predicting oral drug absorption. This is attributed to the elevated expression of transporter proteins and the scarcity of metabolizing enzymes in these cells, making them an effective model for assessing drug permeability [238]. QPMDCK is measured in nm s^{-1} and about 95% of pharmaceutical compounds exhibit QPMDCK values between less than 25 (poor) to over 500 (great) [196,239].

3.2.10. Index of Cohesion Interaction in Solids (Indcoh)

Calculated from the number of hydrogen bond acceptors (HBA), donors (HBD), and the surface area accessible to the solvent (SASA, Smol) by the relation Indcoh = HBA × $HBD^{1/2}$/Smol. Approximately 95% of pharmaceutical compounds exhibit Indcoh values between 0.00 to 0.05 [196,240]. The descriptor in question appears to have gotten little attention in the current body of scientific literature. There have been few in-depth debates or studies of its significance, applicability, or implications in a variety of fields. This scarcity indicates a potential research gap, which may allow for additional analysis and exploration to better understand the descriptor's utility and relevance.

 doi:10.20944/preprints202501.2308.v1

3.2.11. Globularity Descriptor (Glob)

Calculated as $(4\pi r^2) / S$ mol, where r is the radius of the sphere whose volume is equal to the molecular volume. Approximately 95% of pharmaceutical compounds exhibit Glob values between 0.75 and 0.95 [241]. The $4\pi r^2$ term represents the surface area of a perfect sphere with the same volume as the molecule. The denominator represents the actual surface area of the molecule, considering its real shape and any irregularities. This descriptor quantifies the molecular shape's conformational compactness relative to an ideal spherical configuration, i.e., is a measure of the molecule bulkiness [242]. A value close to 1 indicates a spherical and compact molecule, while lower values suggest a more elongated, irregular, or flexible shape. Flexible molecules typically have lower globularity values as they can easily change their shape and become more elongated [243].

3.2.12. Predicted Polarizability (QPpolrz)

Polarizability quantifies how easily the cloud of electrons around an atom or molecule can be distorted or displaced from its equilibrium configuration when subjected to an external electric field. For electric fields significantly weaker than the intrinsic electrostatic interactions within the atom or molecule, the energy associated with this distortion increases with the square of the field's strength. While polarizability can be fully described using a tensor (more precisely, a 3D tensor), in cases where the charge distribution around the molecule is spherically symmetric (or approximated so), a single number (a scalar value) can effectively represent polarizability. In many instances, using an average polarizability is sufficient for calculations [244]. In pharmaceutical research, polarizability has been used as a key descriptor for a variety of QSPR and QSAR models, *inter alia* showing a moderate relationship with the n-octanol/water partition coefficient logarithm (log P, explaining approximately 21% of the variance), and a more substantial connection to aqueous solubility, (accounting for roughly 44% of the observed variability) [244]. Other aspects of pharmacokinetic relevance seem also to be related to polarizability, e.g., the van der Waals interactions between different drugs or metabolites and body albumin or lipids, are favored by the polarizability of those drugs or molecules [245]. The way chemicals bind with body fluids or cells, their movement, and other interactions are influenced by the polarizability of both the chemical and the system. Despite its importance, most QSAR studies up to date have not utilized this descriptor [246]. Whereas it seems to have great potential in offering insight in the pharmacokinetics of a molecule, QPpolrz is not a PK parameter as such, but rather an important descriptor that can be used in predicting various PK properties. Approximately 95% of pharmaceutical compounds exhibit polarizability values between 13.0 to 70.0 [196].

3.2.13. Predicted Skin Permeability (log Kp)

A benchmarking study comparing experimental Kp values for 30 compounds found that QikProp predictions achieved a correlation coefficient of 0.612 with the experimental data. In contrast, the simpler Potts and Guy model, which uses linear regression based solely on logP and molecular weight (MW), had a lower correlation coefficient of 0.544 [210]. While QikProp demonstrated better predictive accuracy compared to the Potts and Guy model, a correlation coefficient of 0.61 is still relatively modest and it indicates that its predictions are not highly accurate. Approximately 95% of pharmaceutical compounds exhibit log Kp values between -8.0 and -1.0, this being therefore considered "the recommended range" [196,247].

3.2.14. Number of Likely Metabolic Reactions (#metab)

The number of metabolic steps a molecule undergoes before reaching its target site influences its ability to access that site after entering the bloodstream. Fewer metabolic steps generally mean easier access to the target site [248]. Molecules that experience numerous metabolic processes may be destroyed or altered prior to arriving at the target site, hence diminishing their likelihood of efficacy. Drug developers are interested in identifying or creating compounds that exhibit metabolic stability and can effectively reach their target locations. Most pharmaceutical compounds (95%) undergo between 1–8 metabolic reactions [196].

 doi:10.20944/preprints202501.2308.v1

3.3. pkCSM

Similarly to SwissADME, this is a freely available web server, using machine learning algorithms to model the relationships between chemical structure and different ADME values. It incorporates two distinct types of predictive models:

- Quantitative models: 14 regression models are available to forecast numerical values for various pharmacokinetic and toxicity properties.
- Qualitative models: 16 classification models are available to predict the likelihood of a specific outcome falling into one of two categories.

When published in 2015, the authors claimed equality or superiority over other available models published in the literature [193].

The 30 models available in pkCSM are grouped under several headings:

a) **Absorption**: numeric predictions are available for water solubility, Caco2 permeability, HIA (%), skin permeability (log Kp), and binary categorical predictions are made for P-glycoprotein substrates (yes/no), P-glycoprotein I inhibitors and I P-glycoprotein II inhibitors (inhibitors of P-glycoprotein and P-glycoprotein transport, respectively).

b) **Distribution**: numeric predictions are available for VDss (human), fraction unbound (human), BBB permeability and CNS permeability. BBB permeability predictions are derived from in vivo measurements conducted in animal models. In contrast, CNS permeability predictions are based on data obtained through direct carotid artery injection, eliminating the influence of systemic distribution effects [193].

c) **Metabolism**: only binary categorical predictions are available under this heading. The application predicts whether a substance is a substrate for the CYP2D6 or CYP3A4 fractions, and whether it is an inhibitor of one or all of the CYP1A2, CYP2C19, CYP2C9, CYP2D6, and CYP3A4 fractions.

d) **Excretion**: total clearance is predicted numerically, whereas the property of a substance to be a renal OCT2 substrate is predicted categorically (yes/no).

Toxicity – a number of ten toxicity endpoints are predicted under this heading, but those are not really ADME features, but safety features, and we will not enter into details here. There is also an option to make all available predictions in a single click.

3.4. ProTox-II

ProTox-II was initially launched as a web-based platform for predicting the toxicity of chemical substances [249], and it has been used in at least 45 scholarly articles focusing on the ADME evaluation of natural compounds. However, in the meantime its successor Protox-3 has been developed and made available [195]. Protox-3 makes a relatively large number of predictions relevant from the toxicity standpoint, but with respect to ADME, it only features a number of six models predicting whether the compund is metabolized by the CYP fractions CYP1A2, CYP2C19, CYP2C9, CYP2D6, CYP3A4, and CYP2E1. The prediction is based on classification models (active/inactive), and the predicted class is accompanied by the probability of that prediction, thus offering some additional quantiative information on the extent of the trust one should have in that prediction. The paper describing the platform [249] has not reported on the calibration validation of those predictors, therefore it is not clear to what extent the probability output of those models are well aligned with true class probability (in other words, the authors have not reported on how accurate are those probabilities).

3.5. admetSAR

Similarly to ProTox, admetSAR has also recently reached a third version [197]. In our experience, the access to the website and the speed of the calculations was slower than for the other web-servers mentioned above (but the experience might be different for users from other countries). It also groups its different predictions under several headings:

a) **Physicochemical properties -** molecular weight, nAtom (number of atoms), nHet (number of heteroatoms), nRing (number of rings), nRot (number of rotable bounds), HBA (hydrogen bound acceptors), HBD (hydroen bound donors), TPSA, SlogP, „application domain" (applicability domain, i.e., the range of chemical space or feature space where a predictive model is reliable and valid, in

 doi:10.20944/preprints202501.2308.v1

other words, a measure about how trustful should be one in the validity of the predictions performed).

b) **Absorption** – logS, logP, pka, acidic pKa, basic pKa, Caco-2 (two predictions, one for the permeability value, the other to classify a substance as of high or low permeability), HIA, MDCK, F50%, F30%, F20% (oral bioavailability, a compound will be classified as belonging to one of the three and not belonging to the other two).

c) **Distribution** – BBB; inhibition of OATP1B1, OATP1B3, OATP2B1, OCT1, OCT2, BCRP, BSEP, MATE1, and Pgp; property of being a Pgp substrate; plasma protein binding ratio and VDss (only the last two are numeric values, all others are based on categorical models).

d) **Metabolism** – inhibition of different CYP fractions (CYP1A2, CYP3A4, CYP2B6, CYP2C9, CYP2C19, CYP2D6, CYP1A2, CYP3A4, CYP2B6, CYP2C9, CYP2C19, and CYP2D6), human liver microsomal stability (HLM), rat liver microsomal stability (RLM), and being a UGT substrate (all are binary predictions).

e) **Excretion** – plasma clearance (CLp), renal clearance (CLr), half-life (t1/2), and mean retention time (MRT).

The application leverages the CLMGraph framework, employing a deep-learning approach that combines regression and classification for predictions. The average AUC (Area Under the Curve) across 90 classification endpoints, as reported by the model developers, is 0.870, indicating high predictive accuracy for most endpoints. However, a few endpoints, notably renal clearance, demonstrate lower predictive performance. Over 82% of regression endpoints exhibit moderate or strong predictive power, with Pearson correlation coefficients exceeding 0.70. Certain parameters, though, such as half-life and mean residence time, are affected by a limited predictive accuracy [197].

3.6. ADMETlab

First launched in 2018, ADMETlab has reached its third version in 2024 [198]. It has been developed based on 400,000 data points of high quality and makes predictions for 119 endpoints (an increase of 31 additional endpoints as compared with the previous version). ADMTlab is based on what is in essence a QSAR model built with multi-task deep message passing neural networks (DMPNN) based on a variety of chemical descriptors.

The ADMET 3.0 platform encompasses a comprehensive set of 119 endpoints, categorized as follows:

a) **Physicochemical Properties** (21 endpoints): molecular weight, van der Waals volume, density, nHA (number of hydrogen acceptors), nHD (number of hydrogen donors), nRot (number of rotatable bonds), nRing (number of rings), MaxRing (number of atoms in the largest ring), nHet (number of heteroatoms), fChar (formal charge), nRig (number of rigid bonds), Flexibility (ratio of rotatable bonds and rigid bonds), Stereo Centers, TPSA, logS, logD7.4, logP, melting point, boiling point, pKa (acidic), pKa (basic).

b) **Medical Chemistry Properties** (20 endpoints): QED (*Quantitative Estimate of Drug-likeness*, a metric to assess how likely a compound is to have the necessary features to become a drug), SAscore (synthetic accessibility score, estimating the easiness or difficulty of synthesizing the chemical compound), GASA (*Graph Attention-based assessment of Synthetic Accessibility*), Fsp^3 (the ratio between the count of sp^3 hybridized carbon atoms and all atom carbons in the molecule), MCE-18 (Medicinal Chemistry Evolution, a descriptor evaluating the novelty of organic molecules through their tetrahedral carbon atoms), NPscore (Natural product likeness score), Lipinski Rule, Pfizer Rule, GSK Rule, GoldenTriangle, PAINS, Alarm_NMR Rule (used to predict thiol reactive compounds), BMS Rule, Chelating Rule, Colloidal aggregators, FLuc inhibitors (predict inhibitors of firefly luciferase), Blue fluorescence, Green fluorescence, Reactive compounds, Promiscuous compounds.

c) **Absorption Properties** (9 endpoints): Caco-2 Permeability, MDCK Permeability, PAMPA, Pgp inhibitor, Pgp substrate, HIA, F20%, F30%, and F50%.

d) **Distribution Properties** (9 endpoints): PPB (Plasma Protein Binding), VDss, BBB, Fu (fraction unbound in the plasm), OATP1B1 inhibitor, OATP1B3 inhibitor, BCRP inhibitor, MRP1 inhibitor, BSEP inhibitor.

 doi:10.20944/preprints202501.2308.v1

e) **Metabolism Properties** (14 endpoints): the property of being an inhibitor or substrate of CYP1A2, CYP2C19, CYP2C9, CYP2D6, CYP3A4, CYP2B6, CYP2C8, as well as HLM stability.

f) **Excretion Properties** (2 endpoints): CLplasma (plasma clearance) and T1/2.

36 toxicity properties and 8 toxicophore rules are additionally predicted for safety evaluation properties.

A synthetic comparison of the four main free ADME web-server applications is shown in Table 2.

Table 2. Synthetic comparison of the four main free ADME web-server applications.

Features	SwissADME	pkCSM	ADMETlab 3.0	admetSAR 3.0
Physicochemical properties	12	7*	21	10
Medicinal chemistry endpoints	10**	0	20	4**
Absorption*** endpoints	3 (C)	3 (N)	9 (2N, 7C)	14 (6N, 6C)
Distribution endpoints	1 (C)	4 (N)	9 (3N, 6 C)	11 (1N, 12C)
Metabolism endpoints	5 (C)	7 (C)	14 (C)	15 (C)
Excretion endpoints	0	2 (1N, 1C)	2 (N)	4 (2N, 2C)
PAINS included	Yes	No	Yes	No
Batch evaluation/API support	Multiple smiles allowed	Limit to 100 smiles	Input limited to one smile, but API available	Batch prediction allowed for 1000 molecules.
Interpretation help	++	++	+++	++
Uncertainty estimation	No	No	Yes (prediction probabilities for categorical predictions converted into six symbols)	Yes (prediction probabilities for categorical predictions)
Availability	Free	Free	Free	Free

*Water solubility, although included by pkCSM as an "absorption" endpoint, is in fact a physicochemical property. **Drug likeness metrics have also been included here. ***Some applications includes Pgp under Absorption, others under Distribution. In this table Pgp endpoints were included under absorption. N - numeric, C- categorical.

4. Conclusions

The application of in silico ADME-PK methods in pharmaceutical research became widespread following Lipinski's rule of 5, though it has historically been viewed as an auxiliary component of disciplines like drug metabolism, pharmacokinetics, and computational chemistry. Most practitioners in this field come from backgrounds in computer science, chemistry, or drug metabolism and PK rather than having specialized education in in silico ADME-PK techniques. While this lack of formal training might seem limiting, it has actually fostered a collaborative environment where experts from diverse fields - including in vitro research, statistics, analytical chemistry, pharmacokinetics, structural biology, medicinal chemistry, and machine learning - can work together to drive innovation [250].

Initially met with some hesitation, the application of in silico ADME models and tools to natural compounds, particularly those of herbal origin, has witnessed a significant surge in recent years. While this field is relatively young, having emerged within the last three decades, its growth has been remarkable. The number of research publications utilizing in silico approaches for assessing natural compounds has surpassed 3000 in our estimation, with a substantial majority appearing within the last five years. This rapid growth is evident in the publication trends:

- Pre-2020: Fewer than 100 publications containing the phrase "in silico" in the title or abstract.
- 2021: Approximately 200 such publications.
- 2023-2024: Over 270 such publications annually.

Similarly, if only seven papers indexed by PubMed seems to have included "natural" and "docking" in their title or abstract before 2010, in 2024 alone the number of this kind of papers has exceeded 300. This exponential growth underscores the increasing recognition and adoption of in silico methods in natural product research.

As seen in this review, there are multiple ways of approaching ADME predictions using computational methods, each with their own challenges, costs and limitations. Quantum mechanics methods have been less used, but it is likely that their applications will also increase more and more, as already evidenced by the trend in numbers in the last ten years. The usefulness of molecular docking in speeding drug development remains a topic of debate, despite significant refinement and

 doi:10.20944/preprints202501.2308.v1

improvement of this method over the years. It is widely known that the predictive accuracy of docking often falls short in complex biological systems and it is expected to improve the currently available docking tools to make them more reliable. Even so, molecular docking continues and will probably continue to play a valuable role, particularly in the field of natural products, where resources for wet lab experiments are scarcer. One would expect that pharmacophore modeling will also improve moving beyond the conventional approach based on key interaction points to more sophisticated approaches, for instance integrating quantum mechanics calculations or and machine learning tools. This will allow a better understanding of pharmacophoric features particular to natural compounds and better predictions not only of pharmacodynamic interactions, but also of their ADME profiles.

The increasing computational power and refinement of force fields is likely to allow more accurate predictions of protein-ligand interactions, membrane permeability, and metabolic transformations for natural compounds in the future. The extension of these techniques and improvement with the help of machine learning approaches, will probably contribute even more to a better understanding of protein interactions with natural compounds. They might even be successful where traditional QSAR methods may fall short due to the structural complexity, novelty and diversity. However, QSAR methods, particularly in implementations based on deep learning approaches seem to be the de facto workhorse for ADME predictions, particularly in the case of the ADME(T) prediction servers. An impressive number of such servers have become available for free in the last decade and they seem to be particularly valued by the natural compound researchers, who tend to use them more and more in their research.

Finally, the continued advancement of *-omics* technologies, coupled with the generation of additional wet-lab experimental data, will allow a refinement and improvement of predictions of natural compound ADME properties using the PBPK modelling. Most likely all of these methods will continue to evolve and improve, and they will allow a better understanding of the ADME profiles and druggability not only for natural compounds, but also for various chemical modifications of such compounds, intended specifically to make them more appropriate for drug development purposes. The future will most likely be brighter, but for the computational chemist, the drug developer, as well as for the phytochemist, these are exciting times to be alive.

Author Contributions: Conceptualization, R.A. and D.D.; methodology, R.A.; software, R.A.; validation, B.E.L.., D.D. and M.D.; formal analysis, R.A. and M.D.; investigation, R.A. and D.D.; resources, D.D.; data curation, B.E.L..; writing—original draft preparation, R.A., B.E.L. and M.D.; writing—review and editing, R.A. and D.D.; visualization, R.A.; supervision, D.D.; project administration, M.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data is contained within the article.

Acknowledgments: Not applicable.

Conflicts of Interest: R.A. received consultancy or speakers' fees in the past from Merck, UCB, Sandoz, Abbvie, Zentiva, Teva, Laropharm, CEGEDIM, Angelini, Biessen Pharma, Hofigal, AstraZeneca, and Stada. All other authors report no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ADMET	Asorption, distribution, metabolism, excretion, and toxicity
ADME	Asorption, distribution, metabolism, and excretion
QSAR	Quantitative structure-activity relationship
PBPK	Physiologically-based pharmacokinetics
PAINS	Pan-assay interference compounds

 doi:10.20944/preprints202501.2308.v1

cLogP	Calculated partition coefficient
LLE	Lipophilicity ligand efficiency
SFI	Solubility forecast index
PFI	Property forecast index
QM	Quantum mechanics
MNDO	Modified neglect of diatomic overlap
AM1	Austin model 1
PMn	Parametric method n
OMn	Orthogonalization-corrected method n
DFTB	Density-functional tight-binding
HF	Hartree-Fock
SCF	Self-consistent field
MPPT	Møller-Plesset perturbation theory
MP	Møller-Plesset perturbation theory
CI	Configuration interaction theory
CC	Coupled cluster
CASSCF	Complete active space self-consistent field
CASPT2	Complete active space perturbation theory
MCSCF	Multi-configurational self-consistent field
DMRG	Density matrix renormalization group method
DFT	Density functional theory
DFT-D	Dispersion-corrected DFT
GGA	Generalized gradient approximation
MM	Molecular mechanics
OTC	Organic cation transporter
IUPAC	International Union of Pure and Applied Chemistry
SBP	Structure-based pharmacophore
PDB	Protein Data Bank
OATn	Organic anion transporter n
URAT1	Urate transporter 1
CoMFA	Comparative Molecular Field Analysis
HQSAR	Hologram QSAR
PAMPA	Parallel artificial membrane permeation assay
QSPR	Quantitative structure-property relationship
RMS	Root mean square
HIA	Human intestinal absorption
CV	Cross-validation
CCR	Correct classification rate
MCC	Matthews correlation coefficient
AAE	Average Absolute Error
RMSE	Root mean square error
AME	Absolute mean error
BBB	Blood - brain barrier
DMPC	Dimyristoylphosphatidylcholine
EGCG	Epigallocatechin gallate
MD	Molecular dynamics
MM	PBSA - Molecular Mechanics Poisson - Boltzmann Surface Area
MM	GBSA - Molecular Mechanics Generalized Born Surface Area
CADD	Computer - aided drug design
Kp	Skin permeation coefficient
PK	Pharmacokinetics
Smol	Solvent - Accessible Molecular Surface
SASA	Solvent - Accessible Molecular Surface

 doi:10.20944/preprints202501.2308.v1

Vmol, hfob	Total Volume of Molecule Enclosed by Solvent - Accessible Molecular Surface
log Swat	Logarithm of Aqueous Solubility
QPlogPo/w	Predicted octanol / water partition coefficient
logKhsa	Logarithm of Predicted Binding Constant to Human Serum Albumin
log B/B	Logarithm of Predicted Blood/Brain Barrier Partition Coefficient
BIP caco2	Predicted Apparent Caco - 2 Cell Membrane Permeability
MDCK	Madin - Darby Canine Kidney
QPMDCK	Apparent MDCK Cell Permeability
Indcoh	Index of Cohesion Interaction in Solids
Glob	Globularity Descriptor
QPpolrz	Predicted Polarizability
VDss	Volume of distribution at steady state
HLM	Human liver microsomal stability
RLM	Rat liver microsomal stability
CLp	Plasma clearance
CLr	Renal clearance
MRT	Mean retention time
AUC	Area Under the Curve
DMPNN	Deep message passing neural networks
nHA	Number of hydrogen acceptors
nHD	Number of hydrogen donors
nRot	Number of rotatable bonds
nRing	Number of rings
MaxRing	Number of atoms in the largest ring
nHet	Number of heteroatoms
fChar	Formal charge
nRig	Number of rigid bonds
FLuc	Firefly luciferase
PPB	Plasma Protein Binding

References

1. Lin, J.; Sahakian, D.; De Morais, S.; Xu, J.; Polzer, R.; Winter, S. The Role of Absorption, Distribution, Metabolism, Excretion and Toxicity in Drug Discovery. *CTMC* **2003**, *3*, 1125–1154. https://doi.org/10.2174/1568026033452096.
2. Ferreira, L.L.G.; Andricopulo, A.D. ADMET Modeling Approaches in Drug Discovery. *Drug Discovery Today* **2019**, *24*, 1157–1165. https://doi.org/10.1016/j.drudis.2019.03.015.
3. Baell, J.B.; Nissink, J.W.M. Seven Year Itch: Pan-Assay Interference Compounds (PAINS) in 2017—Utility and Limitations. *ACS Chem. Biol.* **2018**, *13*, 36–44. https://doi.org/10.1021/acschembio.7b00903.
4. Atanasov, A.G.; Zotchev, S.B.; Dirsch, V.M.; the International Natural Product Sciences Taskforce; Orhan, I.E.; Banach, M.; Rollinger, J.M.; Barreca, D.; Weckwerth, W.; Bauer, R.; et al. Natural Products in Drug Discovery: Advances and Opportunities. *Nat Rev Drug Discov* **2021**, *20*, 200–216. https://doi.org/10.1038/s41573-020-00114-z.
5. Feher, M.; Schmidt, J.M. Property Distributions: Differences between Drugs, Natural Products, and Molecules from Combinatorial Chemistry. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 218–227. https://doi.org/10.1021/ci0200467.
6. Chaachouay, N.; Zidane, L. Plant-Derived Natural Products: A Source for Drug Discovery and Development. *DDC* **2024**, *3*, 184–207. https://doi.org/10.3390/ddc3010011.
7. Simoben, C.V.; Babiaka, S.B.; Moumbock, A.F.A.; Namba-Nzanguim, C.T.; Eni, D.B.; Medina-Franco, J.L.; Günther, S.; Ntie-Kang, F.; Sippl, W. Challenges in Natural Product-Based Drug Discovery Assisted with in silico -Based Methods. *RSC Adv.* **2023**, *13*, 31578–31594. https://doi.org/10.1039/D3RA06831E.

8. Barnes, E.C.; Kumar, R.; Davis, R.A. The Use of Isolated Natural Products as Scaffolds for the Generation of Chemically Diverse Screening Libraries for Drug Discovery. *Nat. Prod. Rep.* **2016**, *33*, 372–381. https://doi.org/10.1039/C5NP00121H.
9. Patil, B.S.; Jayaprakasha, G.K.; Chidambara Murthy, K.N.; Vikram, A. Bioactive Compounds: Historical Perspectives, Opportunities, and Challenges. *J. Agric. Food Chem.* **2009**, *57*, 8142–8160. https://doi.org/10.1021/jf9000132.
10. Atanasov, A.G.; Waltenberger, B.; Pferschy-Wenzig, E.-M.; Linder, T.; Wawrosch, C.; Uhrin, P.; Temml, V.; Wang, L.; Schwaiger, S.; Heiss, E.H.; et al. Discovery and Resupply of Pharmacologically Active Plant-Derived Natural Products: A Review. *Biotechnology Advances* **2015**, *33*, 1582–1614. https://doi.org/10.1016/j.biotechadv.2015.08.001.
11. Jia, C.-Y.; Li, J.-Y.; Hao, G.-F.; Yang, G.-F. A Drug-Likeness Toolbox Facilitates ADMET Study in Drug Discovery. *Drug discovery today* **2020**, *25*, 248–258.
12. Labuschagne, P. Impact of Wall Material Physicochemical Characteristics on the Stability of Encapsulated Phytochemicals: A Review. *Food Research International* **2018**, *107*, 227–247. https://doi.org/10.1016/j.foodres.2018.02.026.
13. Bilia, A.R.; Piazzini, V.; Risaliti, L.; Vanti, G.; Casamonti, M.; Wang, M.; Bergonzi, M.C. Nanocarriers: A Successful Tool to Increase Solubility, Stability and Optimise Bioefficacy of Natural Constituents. *CMC* **2019**, *26*, 4631–4656. https://doi.org/10.2174/0929867325666181101110050.
14. Alshamrani, M.; Khan, M.K.; Khan, B.A.; Salawi, A.; Almoshari, Y. Technologies for Solubility, Dissolution and Permeation Enhancement of Natural Compounds. *Pharmaceuticals* **2022**, *15*, 653. https://doi.org/10.3390/ph15060653.
15. Moroy, G.; Martiny, V.Y.; Vayer, P.; Villoutreix, B.O.; Miteva, M.A. Toward in silico Structure-Based ADMET Prediction in Drug Discovery. *Drug Discovery Today* **2012**, *17*, 44–55. https://doi.org/10.1016/j.drudis.2011.10.023.
16. Pozzan, A. QM Calculations in ADMET Prediction. In *Quantum Mechanics in Drug Discovery*; Heifetz, A., Ed.; Methods in Molecular Biology; Springer US: New York, NY, 2020; Vol. 2114, pp. 285–305 ISBN 978-1-07-160281-2.
17. Bowen, J.P.; Guner, O.F. A Perspective on Quantum Mechanics Calculations in ADMET Predictions. *Current Topics in Medicinal Chemistry* **2013**, *13*, 1257–1272.
18. F. Silva-Júnior, E.; M. Aquino, T.; X. Araujo-Junior, J. Quantum Mechanical (QM) Calculations Applied to ADMET Drug Prediction: A Review. *Current Drug Metabolism* **2017**, *18*, 511–526.
19. Voityuk, A. Application of the MNDO Method to Investigation of Properties and Reactivity of Molecules. *Journal of Structural Chemistry* **1988**, *29*, 120–146.
20. Stewart, J.J.P. Optimization of Parameters for Semiempirical Methods V: Modification of NDDO Approximations and Application to 70 Elements. *J Mol Model* **2007**, *13*, 1173–1213. https://doi.org/10.1007/s00894-007-0233-4.
21. Sulimov, A.V.; Kutov, D.C.; Katkova, E.V.; Sulimov, V.B. Combined Docking with Classical Force Field and Quantum Chemical Semiempirical Method PM7. *Advances in Bioinformatics* **2017**, *2017*, 1–6. https://doi.org/10.1155/2017/7167691.
22. Stewart, J.J. Optimization of Parameters for Semiempirical Methods VI: More Modifications to the NDDO Approximations and Re-Optimization of Parameters. *Journal of molecular modeling* **2013**, *19*, 1–32.
23. Korth, M.; Thiel, W. Benchmarking Semiempirical Methods for Thermochemistry, Kinetics, and Noncovalent Interactions: OMx Methods Are Almost As Accurate and Robust As DFT-GGA Methods for Organic Molecules. *J. Chem. Theory Comput.* **2011**, *7*, 2929–2936. https://doi.org/10.1021/ct200434a.
24. Seifert, G.; Joswig, J. Density-functional Tight Binding—an Approximate Density-functional Theory Method. *WIREs Comput Mol Sci* **2012**, *2*, 456–465. https://doi.org/10.1002/wcms.1094.
25. Estácio, S.P.; Mendes, F.R.D.S.; Marinho, E.; Campos, O.; Marinho, M.M.; Marinho, E.S. Electronic, Structural and Pharmacocinetic Characterization of Tricyclic Alkaloid Alternamide a: A Semi-Empirical Quantum Study and Admet. *Iinternational J. Res.-Granthaalayah* **2019**, *12*, 429–447.
26. Marcucci, M.C.; Oliveira, C.R.; Spindola, D.; Antunes, A.A.; Santana, L.Y.K.; Cavalaro, V.; Costa, I.B.; De Carvalho, A.C.; Veiga, T.A.M.; Medeiros, L.S.; et al. Molecular Dereplication and In Vitro and In silico

Pharmacological Evaluation of Coriandrum Sativum against Neuroblastoma Cells. *Molecules* **2022**, *27*, 5389. https://doi.org/10.3390/molecules27175389.

27. Ali, H.M.; Elgat, W.A.A.A.; EL-Hefny, M.; Salem, M.Z.M.; Taha, A.S.; Al Farraj, D.A.; Elshikh, M.S.; Hatamleh, A.A.; Abdel-Salam, E.M. New Approach for Using of Mentha Longifolia L. and Citrus Reticulata L. Essential Oils as Wood-Biofungicides: GC-MS, SEM, and MNDO Quantum Chemical Studies. *Materials* **2021**, *14*, 1361. https://doi.org/10.3390/ma14061361.
28. Taher, M.A.; Kundu, R.; Laboni, A.A.; Shompa, S.A.; Moniruzzaman, Md.; Hasan, M.M.; Hasnat, H.; Hasan, Md.M.; Khan, M. Unlocking the Medicinal Arsenal of *Cissus Assamica* : GC-MS/MS, FTIR, and Molecular Docking Insights. *Health Science Reports* **2024**, *7*, e70091. https://doi.org/10.1002/hsr2.70091.
29. Hung, N.H.; Quan, P.M.; Satyal, P.; Dai, D.N.; Hoa, V.V.; Huy, N.G.; Giang, L.D.; Ha, N.T.; Huong, L.T.; Hien, V.T.; et al. Acetylcholinesterase Inhibitory Activities of Essential Oils from Vietnamese Traditional Medicinal Plants. *Molecules* **2022**, *27*, 7092. https://doi.org/10.3390/molecules27207092.
30. Apeh, V.O.; Okafor, K.C.; Chukwuma, I.F.; Uzoeto, H.O.; Chinebu, T.I.; Nworah, F.N.; Edache, E.I.; Okafor, I.P.; Anthony, O.C. Exploring the Potential of Aqueous Extracts of *ARTEMISIA ANNUA* ANAMED (A3) for Developing New Anti-malarial Agents: In vivo and Silico Computational Approach. *Engineering Reports* **2024**, *6*, e12831. https://doi.org/10.1002/eng2.12831.
31. Barhouchi, B.; Menacer, R.; Bouchkioua, S.; Mansour, A.; Belattar, N. Compounds from Myrtle Flowers as Antibacterial Agents and SARS-CoV-2 Inhibitors: In-Vitro and Molecular Docking Studies. *Arabian Journal of Chemistry* **2023**, *16*, 104939. https://doi.org/10.1016/j.arabjc.2023.104939.
32. Cremer, D. Møller–Plesset Perturbation Theory: From Small Molecule Methods to Methods for Thousands of Atoms. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2011**, *1*, 509–530.
33. Ramachandran, K.I.; Deepa, G.; Namboori, K. The Ab Initio Method. In *Computational Chemistry and Molecular Modeling*; Springer Berlin Heidelberg: Berlin, Heidelberg, 2008; pp. 155–170 ISBN 978-3-540-77302-3.
34. Cremer, D. From Configuration Interaction to Coupled Cluster Theory: The Quadratic Configuration Interaction Approach. *WIREs Comput Mol Sci* **2013**, *3*, 482–503. https://doi.org/10.1002/wcms.1131.
35. Ouma, R.B.O.; Ngari, S.M.; Kibet, J.K. A Review of the Current Trends in Computational Approaches in Drug Design and Metabolism. *Discov Public Health* **2024**, *21*, 108. https://doi.org/10.1186/s12982-024-00229-3.
36. Austin, B.M.; Zubarev, D.Yu.; Lester, W.A. Quantum Monte Carlo and Related Approaches. *Chem. Rev.* **2012**, *112*, 263–288. https://doi.org/10.1021/cr2001564.
37. Dreuw, A.; Head-Gordon, M. Single-Reference Ab Initio Methods for the Calculation of Excited States of Large Molecules. *Chem. Rev.* **2005**, *105*, 4009–4037. https://doi.org/10.1021/cr0505627.
38. Von Lilienfeld, O.A.; Müller, K.-R.; Tkatchenko, A. Exploring Chemical Compound Space with Quantum-Based Machine Learning. *Nat Rev Chem* **2020**, *4*, 347–358. https://doi.org/10.1038/s41570-020-0189-9.
39. Arodola, O.A.; Soliman, M.E. Quantum Mechanics Implementation in Drug-Design Workflows: Does It Really Help? *Drug design, development and therapy* **2017**, 2551–2564.
40. Park, J.W.; Al-Saadon, R.; MacLeod, M.K.; Shiozaki, T.; Vlaisavljevich, B. Multireference Electron Correlation Methods: Journeys along Potential Energy Surfaces. *Chem. Rev.* **2020**, *120*, 5878–5909. https://doi.org/10.1021/acs.chemrev.9b00496.
41. Ishikawa, K.L.; Sato, T. A Review on Ab Initio Approaches for Multielectron Dynamics. *IEEE Journal of Selected Topics in Quantum Electronics* **2015**, *21*, 1–16.
42. Lüchow, A. Quantum Monte Carlo Methods. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2011**, *1*, 388–402.
43. Kong, L.; Bischoff, F.A.; Valeev, E.F. Explicitly Correlated R12/F12 Methods for Electronic Structure. *Chem. Rev.* **2012**, *112*, 75–107. https://doi.org/10.1021/cr200204r.
44. Yanai, T.; Kurashige, Y.; Mizukami, W.; Chalupskỳ, J.; Lan, T.N.; Saitow, M. Density Matrix Renormalization Group for Ab Initio Calculations and Associated Dynamic Correlation Methods: A Review of Theory and Applications. *International Journal of Quantum Chemistry* **2015**, *115*, 283–299.
45. Zhang, I.Y.; Xu, X. On the Top Rung of Jacob's Ladder of Density Functional Theory: Toward Resolving the Dilemma of SIE and NCE. *WIREs Comput Mol Sci* **2021**, *11*, e1490. https://doi.org/10.1002/wcms.1490.

46. Yilmazer, N.; Korth, M. Recent Progress in Treating Protein–Ligand Interactions with Quantum-Mechanical Methods. *IJMS* **2016**, *17*, 742. https://doi.org/10.3390/ijms17050742.
47. Perdew, J.P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868. https://doi.org/10.1103/PhysRevLett.77.3865.
48. Tirado-Rives, J.; Jorgensen, W.L. Performance of B3LYP Density Functional Methods for a Large Set of Organic Molecules. *J. Chem. Theory Comput.* **2008**, *4*, 297–306. https://doi.org/10.1021/ct700248k.
49. Sousa, S.F.; Ribeiro, A.J.M.; Neves, R.P.P.; Brás, N.F.; Cerqueira, N.M.F.S.A.; Fernandes, P.A.; Ramos, M.J. Application of Quantum Mechanics/Molecular Mechanics Methods in the Study of Enzymatic Reaction Mechanisms. *WIREs Comput Mol Sci* **2017**, *7*, e1281. https://doi.org/10.1002/wcms.1281.
50. Khanal, M.; Acharya, A.; Maharjan, R.; Gyawali, K.; Adhikari, R.; Mulmi, D.D.; Lamichhane, T.R.; Lamichhane, H.P. Identification of Potent Inhibitors of HDAC2 from Herbal Products for the Treatment of Colon Cancer: Molecular Docking, Molecular Dynamics Simulation, MM/GBSA Calculations, DFT Studies, and Pharmacokinetic Analysis. *PLoS ONE* **2024**, *19*, e0307501. https://doi.org/10.1371/journal.pone.0307501.
51. Hossain, A. Molecular Docking, Drug-Likeness and ADMET Analysis, Application of Density Functional Theory (DFT) and Molecular Dynamics (MD) Simulation to the Phytochemicals from Withania Somnifera as Potential Antagonists of Estrogen Receptor Alpha (ER- α). *Curr Comput Aided Drug Des* **2021**, *17*, 797–805. https://doi.org/10.2174/1573409916999200730181611.
52. Anza, M.; Endale, M.; Cardona, L.; Cortes, D.; Eswaramoorthy, R.; Zueco, J.; Rico, H.; Trelis, M.; Abarca, B. Antimicrobial Activity, in silico Molecular Docking, ADMET and DFT Analysis of Secondary Metabolites from Roots of Three Ethiopian Medicinal Plants. *Advances and Applications in Bioinformatics and Chemistry* **2021**, 117–132.
53. Eswaramoorthy, R.; Hailekiros, H.; Kedir, F.; Endale, M. In silico Molecular Docking, DFT Analysis and ADMET Studies of Carbazole Alkaloid and Coumarins from Roots of Clausena Anisata: A Potent Inhibitor for Quorum Sensing. *Advances and Applications in Bioinformatics and Chemistry* **2021**, 13–24.
54. Qaddir, I.; Rasool, N.; Hussain, W.; Mahmood, S. Computer-Aided Analysis of Phytochemicals as Potential Dengue Virus Inhibitors Based on Molecular Docking, ADMET and DFT Studies. *J Vector Borne Dis* **2017**, *54*, 255. https://doi.org/10.4103/0972-9062.217617.
55. Bitew, M.; Desalegn, T.; Demissie, T.B.; Belayneh, A.; Endale, M.; Eswaramoorthy, R. Pharmacokinetics and Drug-Likeness of Antidiabetic Flavonoids: Molecular Docking and DFT Study. *Plos one* **2021**, *16*, e0260853.
56. Macalalad, M.A.B.; Gonzales, A.A. In silico Screening and Identification of Antidiabetic Inhibitors Sourced from Phytochemicals of Philippine Plants against Four Protein Targets of Diabetes (PTP1B, DPP-4, SGLT-2, and FBPase). *Molecules* **2023**, *28*, 5301. https://doi.org/10.3390/molecules28145301.
57. Zhang, Y.-J.; Khorshidi, A.; Kastlunger, G.; Peterson, A.A. The Potential for Machine Learning in Hybrid QM/MM Calculations. *The Journal of Chemical Physics* **2018**, *148*, 241740. https://doi.org/10.1063/1.5029879.
58. Kar, R.K. Benefits of Hybrid QM/MM over Traditional Classical Mechanics in Pharmaceutical Systems. *Drug Discovery Today* **2023**, *28*, 103374. https://doi.org/10.1016/j.drudis.2022.103374.
59. Kollar, J.; Frecer, V. How Accurate Is the Description of Ligand–Protein Interactions by a Hybrid QM/MM Approach? *J Mol Model* **2018**, *24*, 11. https://doi.org/10.1007/s00894-017-3537-z.
60. Hou, J.; Zou, Q.; Wang, Y.; Gao, Q.; Yao, W.; Yao, Q.; Zhang, J. Screening for the Selective Inhibitors of MMP-9 from Natural Products Based on Pharmacophore Modeling and Molecular Docking in Combination with Bioassay Experiment, Hybrid QM/MM Calculation, and MD Simulation. *Journal of Biomolecular Structure and Dynamics* **2019**, *37*, 3135–3149. https://doi.org/10.1080/07391102.2018.1509019.
61. Khodarahmi, G.; Asadi, P.; Farrokhpour, H.; Hassanzadeh, F.; Dinari, M. Design of Novel Potential Aromatase Inhibitors via Hybrid Pharmacophore Approach: Docking Improvement Using the QM/MM Method. *RSC Adv.* **2015**, *5*, 58055–58064. https://doi.org/10.1039/C5RA10097F.
62. Phonglo, A.; Dowerah, D.; Sarma, S.; Ahmed, N.; Dutta, P.; Basumutary, M.; Deka, R.C. Essential Oil Constituents of Regional Ethnomedicinal Plants as Potential Inhibitors of SARS-CoV-2 Mpro: An Integrated Molecular Docking, Molecular Dynamics and QM/MM Study. *Journal of Biomolecular Structure and Dynamics* **2024**, 1–23.
63. Bharadwaj, S.; Rao, A.K.; Dwivedi, V.D.; Mishra, S.K.; Yadava, U. Structure-Based Screening and Validation of Bioactive Compounds as Zika Virus Methyltransferase (MTase) Inhibitors through First-Principle

 doi:10.20944/preprints202501.2308.v1

Density Functional Theory, Classical Molecular Simulation and QM/MM Affinity Estimation. *Journal of Biomolecular Structure and Dynamics* **2021**, *39*, 2338–2351.

64. Pinzi, L.; Rastelli, G. Molecular Docking: Shifting Paradigms in Drug Discovery. *IJMS* **2019**, *20*, 4331. https://doi.org/10.3390/ijms20184331.
65. Stanzione, F.; Giangreco, I.; Cole, J.C. Use of Molecular Docking Computational Tools in Drug Discovery. In *Progress in Medicinal Chemistry*; Elsevier, 2021; Vol. 60, pp. 273–343 ISBN 978-0-323-85056-8.
66. Asiamah, I.; Obiri, S.A.; Tamekloe, W.; Armah, F.A.; Borquaye, L.S. Applications of Molecular Docking in Natural Products-Based Drug Discovery. *Scientific African* **2023**, *20*, e01593. https://doi.org/10.1016/j.sciaf.2023.e01593.
67. Mohanty, M.; Mohanty, P.S. Molecular Docking in Organic, Inorganic, and Hybrid Systems: A Tutorial Review. *Monatsh Chem* **2023**, *154*, 683–707. https://doi.org/10.1007/s00706-023-03076-1.
68. Pagadala, N.S.; Syed, K.; Tuszynski, J. Software for Molecular Docking: A Review. *Biophys Rev* **2017**, *9*, 91–102. https://doi.org/10.1007/s12551-016-0247-1.
69. Zou, W.; Shi, B.; Zeng, T.; Zhang, Y.; Huang, B.; Ouyang, B.; Cai, Z.; Liu, M. Drug Transporters in the Kidney: Perspectives on Species Differences, Disease Status, and Molecular Docking. *Front. Pharmacol.* **2021**, *12*, 746208. https://doi.org/10.3389/fphar.2021.746208.
70. Ren, T.; Li, M.; Zheng, H.; Liu, W.; Zhang, J. Microdialysis Combined with RRLC-MS/MS for the Pharmacokinetics of Two Major Alkaloids of Bi Qi Capsule and the Potential Roles of P-Gp and BCRP on Their Penetration. *J Chromatogr B Analyt Technol Biomed Life Sci* **2018**, *1092*, 72–81. https://doi.org/10.1016/j.jchromb.2018.05.048.
71. Saeed, M.E.M.; Boulos, J.C.; Elhaboub, G.; Rigano, D.; Saab, A.; Loizzo, M.R.; Hassan, L.E.A.; Sugimoto, Y.; Piacente, S.; Tundis, R.; et al. Cytotoxicity of Cucurbitacin E from Citrullus Colocynthis against Multidrug-Resistant Cancer Cells. *Phytomedicine* **2019**, *62*, 152945. https://doi.org/10.1016/j.phymed.2019.152945.
72. Bai, J.; Zhao, S.; Fan, X.; Chen, Y.; Zou, X.; Hu, M.; Wang, B.; Jin, J.; Wang, X.; Hu, J.; et al. Inhibitory Effects of Flavonoids on P-Glycoprotein in Vitro and in vivo: Food/Herb-Drug Interactions and Structure–Activity Relationships. *Toxicology and Applied Pharmacology* **2019**, *369*, 49–59. https://doi.org/10.1016/j.taap.2019.02.010.
73. Bi, Y.; Wang, X.; Han, L.; Tian, Y.; Bo, T.; Li, C.; Shi, B.; Gui, C.; Zhang, Y. Selective Inhibition of Organic Cation Transporter 1 by Benzoylpaeoniflorin Attenuates Hepatic Lipid Accumulation through AMPK Activation. *J Nat Prod* **2023**, *86*, 191–198. https://doi.org/10.1021/acs.jnatprod.2c00927.
74. Feng, J.; Zhao, X.; Yan, Y.; Chen, H.; Liu, J.; Li, X.; Na, R.; Li, Q.X. Interactions between Stipuol Enantiomers and Human Serum Albumin. *Food Chem* **2022**, *385*, 132686. https://doi.org/10.1016/j.foodchem.2022.132686.
75. Ai, N.; Fan, X.; Ekins, S. In silico Methods for Predicting Drug-Drug Interactions with Cytochrome P-450s, Transporters and Beyond. *Adv Drug Deliv Rev* **2015**, *86*, 46–60. https://doi.org/10.1016/j.addr.2015.03.006.
76. Raju, B.; Verma, H.; Narendra, G.; Sapra, B.; Silakari, O. Multiple Machine Learning, Molecular Docking, and ADMET Screening Approach for Identification of Selective Inhibitors of CYP1B1. *J Biomol Struct Dyn* **2022**, *40*, 7975–7990. https://doi.org/10.1080/07391102.2021.1905552.
77. Lu, J.; Liu, D.; Zhou, X.; Chen, A.; Jiang, Z.; Ye, X.; Liu, M.; Wang, X. Plant Natural Product Plumbagin Presents Potent Inhibitory Effect on Human Cytochrome P450 2J2 Enzyme. *Phytomedicine* **2018**, *39*, 137–145. https://doi.org/10.1016/j.phymed.2017.12.026.
78. Shukla, R.; Chetri, P.B.; Sonkar, A.; Pakharukova, M.Y.; Mordvinov, V.A.; Tripathi, T. Identification of Novel Natural Inhibitors of Opisthorchis Felineus Cytochrome P450 Using Structure-Based Screening and Molecular Dynamic Simulation. *J Biomol Struct Dyn* **2018**, *36*, 3541–3556. https://doi.org/10.1080/07391102.2017.1392897.
79. Wang, Q.; Kuang, Y.; He, J.; Li, K.; Song, W.; Jin, H.; Qiao, X.; Ye, M. The Prenylated Phenolic Natural Product Isoglycycoumarin Is a Highly Selective Probe for Human Cytochrome P450 2A6. *Eur J Pharm Sci* **2017**, *109*, 472–479. https://doi.org/10.1016/j.ejps.2017.08.035.
80. Yang, S.-Y. Pharmacophore Modeling and Applications in Drug Discovery: Challenges and Recent Advances. *Drug Discovery Today* **2010**, *15*, 444–450. https://doi.org/10.1016/j.drudis.2010.03.013.

81. Kumar, S.; Deepika, D.; Kumar, V. Pharmacophore Modeling Using Machine Learning for Screening the Blood-Brain Barrier Permeation of Xenobiotics. *Int J Environ Res Public Health* **2022**, *19*, 13471. https://doi.org/10.3390/ijerph192013471.
82. Szwabowski, G.L.; Cole, J.A.; Baker, D.L.; Parrill, A.L. Structure-Based Pharmacophore Modeling 1. Automated Random Pharmacophore Model Generation. *J Mol Graph Model* **2023**, *121*, 108429. https://doi.org/10.1016/j.jmgm.2023.108429.
83. Sanders, M.P.A.; McGuire, R.; Roumen, L.; De Esch, I.J.P.; De Vlieg, J.; Klomp, J.P.G.; De Graaf, C. From the Protein's Perspective: The Benefits and Challenges of Protein Structure-Based Pharmacophore Modeling. *Med. Chem. Commun.* **2012**, *3*, 28–38. https://doi.org/10.1039/C1MD00210D.
84. Giordano, D.; Biancaniello, C.; Argenio, M.A.; Facchiano, A. Drug Design by Pharmacophore and Virtual Screening Approach. *Pharmaceuticals* **2022**, *15*, 646. https://doi.org/10.3390/ph15050646.
85. Temml, V.; Kutil, Z. Structure-Based Molecular Modeling in SAR Analysis and Lead Optimization. *Comput Struct Biotechnol J* **2021**, *19*, 1431–1444. https://doi.org/10.1016/j.csbj.2021.02.018.
86. Muhammed, M.T.; Aki-Yalcin, E. Pharmacophore Modeling in Drug Discovery: Methodology and Current Status. *Journal of the Turkish Chemical Society Section A: Chemistry* **2021**, *8*, 749–762. https://doi.org/10.18596/jotcsa.927426.
87. Wang, F.-H.; Tan, H.-X.; Hu, J.-H.; Duan, X.-Y.; Bai, W.-T.; Wang, X.-B.; Wang, B.-L.; Su, Y.; Hu, J.-P. Inhibitory Interaction of Flavonoids with Organic Anion Transporter 3 and Their Structure-Activity Relationships for Predicting Nephroprotective Effects. *J Asian Nat Prod Res* **2024**, *26*, 353–371. https://doi.org/10.1080/10286020.2023.2240722.
88. Bai, W.; Tan, H.; Duan, X.; Hu, J.; Wang, F.; Wu, J.; Bai, J.; Hu, J. Inhibitory Effects of Flavonoids on Organic Cation Transporter 1: Implications for Food/Herb-Drug Interactions and Hepatoprotective Effects. *Food Chem Toxicol* **2024**, *193*, 114983. https://doi.org/10.1016/j.fct.2024.114983.
89. Tan, H.; Wang, F.; Hu, J.; Duan, X.; Bai, W.; Wang, X.; Wang, B.; Su, Y.; Hu, J. Inhibitory Interaction of Flavonoids with Organic Cation Transporter 2 and Their Structure-Activity Relationships for Predicting Nephroprotective Effects. *J Appl Toxicol* **2023**, *43*, 1421–1435. https://doi.org/10.1002/jat.4474.
90. Komaniecka, N.; Maroszek, S.; Drozdzik, M.; Oswald, S.; Drozdzik, M. Transporter Proteins as Therapeutic Drug Targets-With a Focus on SGLT2 Inhibitors. *Int J Mol Sci* **2024**, *25*, 6926. https://doi.org/10.3390/ijms25136926.
91. Du, C.; Yin, H.; Xie, A.; Yu, J.; Wang, Y.; Yao, F.; Zhang, S.; Zhang, Y.; Liu, L.; Wang, P.; et al. Virtual Screening and Biological Evaluation of Natural Products as Urate Transporter 1 (URAT1) Inhibitors. *J Biomol Struct Dyn* **2024**, 1–14. https://doi.org/10.1080/07391102.2024.2331101.
92. Hao, M.; Zhao, Y.; Chen, P.; Huang, H.; Liu, H.; Jiang, H.; Zhang, R.; Wang, H. Structure-Activity Relationship and Substrate-Dependent Phenomena in Effects of Ginsenosides on Activities of Drug-Metabolizing P450 Enzymes. *PLoS One* **2008**, *3*, e2697. https://doi.org/10.1371/journal.pone.0002697.
93. Bai, J.; Li, L.; Zhao, S.; Fan, X.; Zhang, J.; Hu, M.; Chen, Y.; Sun, Y.; Wang, B.; Jin, J.; et al. Heterotropic Activation of Flavonoids on Cytochrome P450 3A4: A Case Example of Alleviating Dronedarone-Induced Cytotoxicity. *Toxicol Lett* **2020**, *319*, 187–196. https://doi.org/10.1016/j.toxlet.2019.11.016.
94. Tahir, R.A.; Hassan, F.; Kareem, A.; Iftikhar, U.; Sehgal, S.A. Ligand-Based Pharmacophore Modeling and Virtual Screening to Discover Novel CYP1A1 Inhibitors. *Curr Top Med Chem* **2019**, *19*, 2782–2794. https://doi.org/10.2174/1568026619666191112104217.
95. Goyal, N.; Liu, J.; Lovings, L.; Dupart, P.; Taylor, S.; Bellow, S.; Mensah, L.; McClain, E.; Dotson, B.; Sridhar, J.; et al. Ethynylflavones, Highly Potent, and Selective Inhibitors of Cytochrome P450 1A1. *Chem Res Toxicol* **2014**, *27*, 1431–1439. https://doi.org/10.1021/tx5001865.
96. Muratov, E.N.; Bajorath, J.; Sheridan, R.P.; Tetko, I.V.; Filimonov, D.; Poroikov, V.; Oprea, T.I.; Baskin, I.I.; Varnek, A.; Roitberg, A.; et al. QSAR without Borders. *Chem. Soc. Rev.* **2020**, *49*, 3525–3564. https://doi.org/10.1039/D0CS00098A.
97. Neves, B.J.; Braga, R.C.; Melo-Filho, C.C.; Moreira-Filho, J.T.; Muratov, E.N.; Andrade, C.H. QSAR-Based Virtual Screening: Advances and Applications in Drug Discovery. *Frontiers in Pharmacology* **2018**, *9*. https://doi.org/10.3389/fphar.2018.01275.

98. Zhang, S.; Wang, N.; Su, L.; Xu, X.; Li, C.; Qin, W.; Zhao, Y. MOA-Based Linear and Nonlinear QSAR Models for Predicting the Toxicity of Organic Chemicals to Vibrio Fischeri. *Environ Sci Pollut Res* **2020**, *27*, 9114–9125. https://doi.org/10.1007/s11356-019-06681-y.
99. Ramandi, M.; Riahi, S.; Rahimi, H.; Mohammadi-Khanaposhtani, M. Molecular Docking, Linear and Nonlinear QSAR Studies on Factor Xa Inhibitors. *Struct Chem* **2020**, *31*, 2023–2040. https://doi.org/10.1007/s11224-020-01535-7.
100. Zhao, Y.; Xia, Y.; Yu, Y.; Liang, G. QSAR in Natural Non-Peptidic Food-Related Compounds: Current Status and Future Perspective. *Trends in Food Science & Technology* **2023**, *140*, 104165. https://doi.org/10.1016/j.tifs.2023.104165.
101. Olotu, F.A.; Agoni, C.; Soremekun, O.; Soliman, M.E.S. The Recent Application of 3D-QSAR and Docking Studies to Novel HIV-Protease Inhibitor Drug Discovery. *Expert Opinion on Drug Discovery* **2020**, *15*, 1095–1109. https://doi.org/10.1080/17460441.2020.1773428.
102. Hopfinger, A.J.; Wang, S.; Tokarski, J.S.; Jin, B.; Albuquerque, M.; Madhav, P.J.; Duraiswami, C. Construction of 3D-QSAR Models Using the 4D-QSAR Analysis Formalism. *J. Am. Chem. Soc.* **1997**, *119*, 10509–10524. https://doi.org/10.1021/ja9718937.
103. Wang, T.; Yuan, X.; Wu, M.-B.; Lin, J.-P.; Yang, L.-R. The Advancement of Multidimensional QSAR for Novel Drug Discovery - Where Are We Headed? *Expert Opinion on Drug Discovery* **2017**, 1–16. https://doi.org/10.1080/17460441.2017.1336157.
104. Vedani, A.; Dobler, M. 5D-QSAR: The Key for Simulating Induced Fit? *J Med Chem* **2002**, *45*, 2139–2149. https://doi.org/10.1021/jm011005p.
105. Suay-Garcia, B.; Bueso-Bordils, J.I.; Falcó, A.; Pérez-Gracia, M.T.; Antón-Fos, G.; Alemán-López, P. Quantitative Structure–Activity Relationship Methods in the Discovery and Development of Antibacterials. *WIREs Comput Mol Sci* **2020**, *10*, e1472. https://doi.org/10.1002/wcms.1472.
106. Vedani, A.; Dobler, M.; Lill, M.A. Combining Protein Modeling and 6D-QSAR. Simulating the Binding of Structurally Diverse Ligands to the Estrogen Receptor. *J. Med. Chem.* **2005**, *48*, 3700–3703. https://doi.org/10.1021/jm050185q.
107. Leo, A.J.; Hansch, C. Role of Hydrophobic Effects in Mechanistic QSAR. *Perspectives in Drug Discovery and Design* **1999**, *17*, 1–25.
108. Benigni, R.; Bossa, C.; Netzeva, T.; Rodomonte, A.; Tsakovska, I. Mechanistic QSAR of Aromatic Amines: New Models for Discriminating between Homocyclic Mutagens and Nonmutagens, and Validation of Models for Carcinogens. *Environ and Mol Mutagen* **2007**, *48*, 754–771. https://doi.org/10.1002/em.20355.
109. Chaudhry, Q.; Piclin, N.; Cotterill, J.; Pintore, M.; Price, N.R.; Chrétien, J.R.; Roncaglioni, A. Global QSAR Models of Skin Sensitisers for Regulatory Purposes. *Chemistry Central Journal* **2010**, *4*, S5. https://doi.org/10.1186/1752-153X-4-S1-S5.
110. Doucet, J.P.; Doucet-Panaye, A.; Papa, E. Topological QSAR Modelling of Carboxamides Repellent Activity to *Aedes Aegypti*. *Molecular Informatics* **2019**, *38*, 1900029. https://doi.org/10.1002/minf.201900029.
111. Hall, L.H.; Hall, L.M. QSAR Modeling Based on Structure-Information for Properties of Interest in Human Health. *SAR and QSAR in Environmental Research* **2005**, *16*, 13–41. https://doi.org/10.1080/10629360412331319853.
112. Holder, A.J.; Ye, L.; Eick, J.D.; Chappelow, C.C. A Quantum-Mechanical QSAR Model to Predict the Refractive Index of Polymer Matrices. *QSAR Comb. Sci.* **2006**, *25*, 905–911. https://doi.org/10.1002/qsar.200510203.
113. Trossini, G.; Maltarollo, V.; Schmidt, T. Hologram QSAR Studies of Antiprotozoal Activities of Sesquiterpene Lactones. *Molecules* **2014**, *19*, 10546–10562. https://doi.org/10.3390/molecules190710546.
114. Su, B.-H.; Shen, M.; Esposito, E.X.; Hopfinger, A.J.; Tseng, Y.J. In silico Binary Classification QSAR Models Based on 4D-Fingerprints and MOE Descriptors for Prediction of hERG Blockage. *J. Chem. Inf. Model.* **2010**, *50*, 1304–1318. https://doi.org/10.1021/ci100081j.
115. Sahlin, U.; Filipsson, M.; Öberg, T. A Risk Assessment Perspective of Current Practice in Characterizing Uncertainties in QSAR Regression Predictions. *Molecular Informatics* **2011**, *30*, 551–564. https://doi.org/10.1002/minf.201000177.

116. Price, E.; Kalvass, J.C.; DeGoey, D.; Hosmane, B.; Doktor, S.; Desino, K. Global Analysis of Models for Predicting Human Absorption: QSAR, In Vitro , and Preclinical Models. *J. Med. Chem.* **2021**, *64*, 9389–9403. https://doi.org/10.1021/acs.jmedchem.1c00669.
117. Linnankoski, J.; Ranta, V.-P.; Yliperttula, M.; Urtti, A. Passive Oral Drug Absorption Can Be Predicted More Reliably by Experimental than Computational Models--Fact or Myth. *Eur J Pharm Sci* **2008**, *34*, 129–139. https://doi.org/10.1016/j.ejps.2008.03.001.
118. Kansy, M.; Senner, F.; Gubernator, K. Physicochemical High Throughput Screening: Parallel Artificial Membrane Permeation Assay in the Description of Passive Absorption Processes. *J Med Chem* **1998**, *41*, 1007–1010. https://doi.org/10.1021/jm970530e.
119. Akamatsu, M.; Fujikawa, M.; Nakao, K.; Shimizu, R. In silico Prediction of Human Oral Absorption Based on QSAR Analyses of PAMPA Permeability. *Chemistry & Biodiversity* **2009**, *6*, 1845–1866. https://doi.org/10.1002/cbdv.200900112.
120. Diukendjieva, A.; Alov, P.; Tsakovska, I.; Pencheva, T.; Richarz, A.; Kren, V.; Cronin, M.T.D.; Pajeva, I. In Vitro and in silico Studies of the Membrane Permeability of Natural Flavonoids from Silybum Marianum (L.) Gaertn. and Their Derivatives. *Phytomedicine* **2019**, *53*, 79–85. https://doi.org/10.1016/j.phymed.2018.09.001.
121. Boik, J.C.; Newman, R.A. Structure-Activity Models of Oral Clearance, Cytotoxicity, and LD50: A Screen for Promising Anticancer Compounds. *BMC Pharmacol* **2008**, *8*, 12. https://doi.org/10.1186/1471-2210-8-12.
122. Wessel, M.D.; Jurs, P.C.; Tolan, J.W.; Muskal, S.M. Prediction of Human Intestinal Absorption of Drug Compounds from Molecular Structure. *J Chem Inf Comput Sci* **1998**, *38*, 726–735. https://doi.org/10.1021/ci980029a.
123. Moda, T.L.; Andricopulo, A.D. Consensus Hologram QSAR Modeling for the Prediction of Human Intestinal Absorption. *Bioorganic & Medicinal Chemistry Letters* **2012**, *22*, 2889–2893. https://doi.org/10.1016/j.bmcl.2012.02.061.
124. Yoshida, F.; Topliss, J.G. QSAR Model for Drug Human Oral Bioavailability. *J. Med. Chem.* **2000**, *43*, 2575–2585. https://doi.org/10.1021/jm0000564.
125. Suenderhauf, C.; Hammann, F.; Maunz, A.; Helma, C.; Huwyler, J. Combinatorial QSAR Modeling of Human Intestinal Absorption. *Mol. Pharmaceutics* **2011**, *8*, 213–224. https://doi.org/10.1021/mp100279d.
126. Zhao, Y.H.; Le, J.; Abraham, M.H.; Hersey, A.; Eddershaw, P.J.; Luscombe, C.N.; Boutina, D.; Beck, G.; Sherborne, B.; Cooper, I.; et al. Evaluation of Human Intestinal Absorption Data and Subsequent Derivation of a Quantitative Structure–Activity Relationship (QSAR) with the Abraham Descriptors. *Journal of Pharmaceutical Sciences* **2001**, *90*, 749–784. https://doi.org/10.1002/jps.1031.
127. Niwa, T. Using General Regression and Probabilistic Neural Networks To Predict Human Intestinal Absorption with Topological Descriptors Derived from Two-Dimensional Chemical Structures. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 113–119. https://doi.org/10.1021/ci020013r.
128. Bai, J.P.F.; Utis, A.; Crippen, G.; He, H.-D.; Fischer, V.; Tullman, R.; Yin, H.-Q.; Hsu, C.-P.; Jiang, L.; Hwang, K.-K. Use of Classification Regression Tree in Predicting Oral Absorption in Humans. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 2061–2069. https://doi.org/10.1021/ci040023n.
129. Liu, H.X.; Hu, R.J.; Zhang, R.S.; Yao, X.J.; Liu, M.C.; Hu, Z.D.; Fan, B.T. The Prediction of Human Oral Absorption for Diffusion Rate-Limited Drugs Based on Heuristic Method and Support Vector Machine. *J Comput Aided Mol Des* **2005**, *19*, 33–46. https://doi.org/10.1007/s10822-005-0095-8.
130. Jones, R.; Connolly, P.C.; Klamt, A.; Diedenhofen, M. Use of Surface Charges from DFT Calculations To Predict Intestinal Absorption. *J. Chem. Inf. Model.* **2005**, *45*, 1337–1342. https://doi.org/10.1021/ci049653f.
131. Deconinck, E.; Hancock, T.; Coomans, D.; Massart, D.L.; Heyden, Y.V. Classification of Drugs in Absorption Classes Using the Classification and Regression Trees (CART) Methodology. *Journal of Pharmaceutical and Biomedical Analysis* **2005**, *39*, 91–103. https://doi.org/10.1016/j.jpba.2005.03.008.
132. Iyer, S.S.; Cheng, G. Role of Interleukin 10 Transcriptional Regulation in Inflammation and Autoimmune Disease. *Crit. Rev. Immunol.* **2012**, *32*, 23–63.
133. Hou, T.; Wang, J.; Zhang, W.; Xu, X. ADME Evaluation in Drug Discovery. 7. Prediction of Oral Absorption by Correlation and Classification. *J. Chem. Inf. Model.* **2007**, *47*, 208–218. https://doi.org/10.1021/ci600343x.

 doi:10.20944/preprints202501.2308.v1

134. Yan, A.; Wang, Z.; Cai, Z. Prediction of Human Intestinal Absorption by GA Feature Selection and Support Vector Machine Regression. *Int J Mol Sci* **2008**, *9*, 1961–1976. https://doi.org/10.3390/ijms9101961.
135. Reynolds, D.P.; Lanevskij, K.; Japertas, P.; Didziapetris, R.; Petrauskas, A. Ionization-Specific Analysis of Human Intestinal Absorption. *Journal of Pharmaceutical Sciences* **2009**, *98*, 4039–4054. https://doi.org/10.1002/jps.21730.
136. Obrezanova, O.; Segall, M.D. Gaussian Processes for Classification: QSAR Modeling of ADMET and Target Activity. *J. Chem. Inf. Model.* **2010**, *50*, 1053–1061. https://doi.org/10.1021/ci900406x.
137. Shen, J.; Cheng, F.; Xu, Y.; Li, W.; Tang, Y. Estimation of ADME Properties with Substructure Pattern Recognition. *J. Chem. Inf. Model.* **2010**, *50*, 1034–1041. https://doi.org/10.1021/ci100104j.
138. Guerra, A.; Campillo, N.E.; Páez, J.A. Neural Computational Prediction of Oral Drug Absorption Based on CODES 2D Descriptors. *European Journal of Medicinal Chemistry* **2010**, *45*, 930–940. https://doi.org/10.1016/j.ejmech.2009.11.034.
139. Basant, N.; Gupta, S.; Singh, K.P. Predicting Human Intestinal Absorption of Diverse Chemicals Using Ensemble Learning Based QSAR Modeling Approaches. *Computational Biology and Chemistry* **2016**, *61*, 178–196. https://doi.org/10.1016/j.compbiolchem.2016.01.005.
140. Wang, Z. *Bst: Gradient Boosting*; 2016;
141. Czub, N.; Szlęk, J.; Pacławski, A.; Klimończyk, K.; Puccetti, M.; Mendyk, A. Artificial Intelligence-Based Quantitative Structure-Property Relationship Model for Predicting Human Intestinal Absorption of Compounds with Serotonergic Activity. *Mol Pharm* **2023**, *20*, 2545–2555. https://doi.org/10.1021/acs.molpharmaceut.2c01117.
142. Talevi, A.; Goodarzi, M.; Ortiz, E.V.; Duchowicz, P.R.; Bellera, C.L.; Pesce, G.; Castro, E.A.; Bruno-Blanch, L.E. Prediction of Drug Intestinal Absorption by New Linear and Non-Linear QSPR. *European Journal of Medicinal Chemistry* **2011**, *46*, 218–228. https://doi.org/10.1016/j.ejmech.2010.11.005.
143. Agatonovic-Kustrin, S.; Beresford, R.; Yusof, A.P.M. Theoretically-Derived Molecular Descriptors Important in Human Intestinal Absorption. *Journal of Pharmaceutical and Biomedical Analysis* **2001**, *25*, 227–237. https://doi.org/10.1016/S0731-7085(00)00492-1.
144. Klopman, G.; Stefan, L.R.; Saiakhov, R.D. ADME Evaluation. *European Journal of Pharmaceutical Sciences* **2002**, *17*, 253–263. https://doi.org/10.1016/S0928-0987(02)00219-1.
145. Toma, C.; Gadaleta, D.; Roncaglioni, A.; Toropov, A.; Toropova, A.; Marzo, M.; Benfenati, E. QSAR Development for Plasma Protein Binding: Influence of the Ionization State. *Pharm Res* **2018**, *36*, 28. https://doi.org/10.1007/s11095-018-2561-8.
146. Khaouane, A.; Khaouane, L.; Ferhat, S.; Hanini, S. Deep Learning for Drug Development: Using CNNs in MIA-QSAR to Predict Plasma Protein Binding of Drugs. *AAPS PharmSciTech* **2023**, *24*, 232. https://doi.org/10.1208/s12249-023-02686-6.
147. Khaouane, A.; Ferhat, S.; Hanini, S. A Quantitative Structure-Activity Relationship for Human Plasma Protein Binding: Prediction, Validation and Applicability Domain. *Adv Pharm Bull* **2023**, *13*, 784–791. https://doi.org/10.34172/apb.2023.078.
148. Zhivkova, Z.D. Quantitative Structure - Pharmacokinetics Relationships for Plasma Protein Binding of Basic Drugs. *J Pharm Pharm Sci* **2017**, *20*, 349–359. https://doi.org/10.18433/J33633.
149. Riedl, M.; Mukherjee, S.; Gauthier, M. Descriptor-Free Deep Learning QSAR Model for the Fraction Unbound in Human Plasma. *Mol Pharm* **2023**, *20*, 4984–4993. https://doi.org/10.1021/acs.molpharmaceut.3c00129.
150. Mathew, S.; Tess, D.; Burchett, W.; Chang, G.; Woody, N.; Keefer, C.; Orozco, C.; Lin, J.; Jordan, S.; Yamazaki, S.; et al. Evaluation of Prediction Accuracy for Volume of Distribution in Rat and Human Using In Vitro, In vivo, PBPK and QSAR Methods. *J Pharm Sci* **2021**, *110*, 1799–1823. https://doi.org/10.1016/j.xphs.2020.12.005.
151. Luque Ruiz, I.; Gómez-Nieto, M.Á. Robust QSAR Prediction Models for Volume of Distribution at Steady State in Humans Using Relative Distance Measurements. *SAR QSAR Environ Res* **2018**, *29*, 529–550. https://doi.org/10.1080/1062936X.2018.1494038.

152. Simeon, S.; Montanari, D.; Gleeson, M.P. Investigation of Factors Affecting the Performance of in silico Volume Distribution QSAR Models for Human, Rat, Mouse, Dog & Monkey. *Mol Inform* **2019**, *38*, e1900059. https://doi.org/10.1002/minf.201900059.
153. Faramarzi, S.; Kim, M.T.; Volpe, D.A.; Cross, K.P.; Chakravarti, S.; Stavitskaya, L. Development of QSAR Models to Predict Blood-Brain Barrier Permeability. *Front Pharmacol* **2022**, *13*, 1040838. https://doi.org/10.3389/fphar.2022.1040838.
154. Janicka, M.; Sztanke, M.; Sztanke, K. Modeling the Blood-Brain Barrier Permeability of Potential Heterocyclic Drugs via Biomimetic IAM Chromatography Technique Combined with QSAR Methodology. *Molecules* **2024**, *29*, 0. https://doi.org/10.3390/molecules29020287.
155. Toropov, A.A.; Toropova, A.P.; Beeg, M.; Gobbi, M.; Salmona, M. QSAR Model for Blood-Brain Barrier Permeation. *J Pharmacol Toxicol Methods* **2017**, *88*, 7–18. https://doi.org/10.1016/j.vascn.2017.04.014.
156. Yu, R.; Wang, J.; Wang, R.; Lin, Y.; Hu, Y.; Wang, Y.; Shu, M.; Lin, Z. Combined Pharmacophore Modeling, 3D-QSAR, Homology Modeling and Docking Studies on CYP11B1 Inhibitors. *Molecules* **2015**, *20*, 1014–1030. https://doi.org/10.3390/molecules20011014.
157. Yu, T.; Huang, T.; Yu, L.; Nantasenamat, C.; Anuwongcharoen, N.; Piacham, T.; Ren, R.; Chiang, Y.-C. Exploring the Chemical Space of CYP17A1 Inhibitors Using Cheminformatics and Machine Learning. *Molecules* **2023**, *28*, 1679. https://doi.org/10.3390/molecules28041679.
158. Li, Y.; Ning, J.; Wang, Y.; Wang, C.; Sun, C.; Huo, X.; Yu, Z.; Feng, L.; Zhang, B.; Tian, X.; et al. Drug Interaction Study of Flavonoids toward CYP3A4 and Their Quantitative Structure Activity Relationship (QSAR) Analysis for Predicting Potential Effects. *Toxicol Lett* **2018**, *294*, 27–36. https://doi.org/10.1016/j.toxlet.2018.05.008.
159. Gonzalez, E.; Jain, S.; Shah, P.; Torimoto-Katori, N.; Zakharov, A.; Nguyễn, Đ.-T.; Sakamuru, S.; Huang, R.; Xia, M.; Obach, R.S.; et al. Development of Robust Quantitative Structure-Activity Relationship Models for CYP2C9, CYP2D6, and CYP3A4 Catalysis and Inhibition. *Drug Metab Dispos* **2021**, *49*, 822–832. https://doi.org/10.1124/dmd.120.000320.
160. Wongrattanakamon, P.; Lee, V.S.; Nimmanpipug, P.; Jiranusornkul, S. 3D-QSAR Modelling Dataset of Bioflavonoids for Predicting the Potential Modulatory Effect on P-Glycoprotein Activity. *Data Brief* **2016**, *9*, 35–42. https://doi.org/10.1016/j.dib.2016.08.004.
161. Prachayasittikul, V.; Worachartcheewan, A.; Shoombuatong, W.; Prachayasittikul, V.; Nantasenamat, C. Classification of P-Glycoprotein-Interacting Compounds Using Machine Learning Methods. *EXCLI J* **2015**, *14*, 958–970. https://doi.org/10.17179/excli2015-374.
162. Ghaemian, P.; Shayanfar, A. Image-Based QSAR Model for the Prediction of P-Gp Inhibitory Activity of Epigallocatechin and Gallocatechin Derivatives. *Curr Comput Aided Drug Des* **2019**, *15*, 212–224. https://doi.org/10.2174/1573409914666181003152042.
163. Xiang, Y.; Liu, S.; Yang, J.; Wang, Z.; Zhang, H.; Gui, C. Investigation of the Interactions between Flavonoids and Human Organic Anion Transporting Polypeptide 1B1 Using Fluorescent Substrate and 3D-QSAR Analysis. *Biochim Biophys Acta Biomembr* **2020**, *1862*, 183210. https://doi.org/10.1016/j.bbamem.2020.183210.
164. Gui, C.; Li, Y.; Peng, T. Development of Predictive QSAR Models for the Substrates/Inhibitors of OATP1B1 by Deep Neural Networks. *Toxicol Lett* **2023**, *376*, 20–25. https://doi.org/10.1016/j.toxlet.2023.01.006.
165. Ni, Y.; Duan, Z.; Zhou, D.; Liu, S.; Wan, H.; Gui, C.; Zhang, H. Identification of Structural Features for the Inhibition of OAT3-Mediated Uptake of Enalaprilat by Selected Drugs and Flavonoids. *Front Pharmacol* **2020**, *11*, 802. https://doi.org/10.3389/fphar.2020.00802.
166. Bi, Y.; Wang, X.; Li, H.; Tian, Y.; Han, L.; Gui, C.; Zhang, Y. 3D-QSAR Analysis of the Interactions of Flavonoids with Human Organic Cation Transporter 2. *Toxicol Lett* **2022**, *368*, 1–8. https://doi.org/10.1016/j.toxlet.2022.07.811.
167. Sheng, M.; Ma, L.; Li, Z.; Peng, X.; Cen, S.; Feng, M.; Tian, Y.; Dai, X.; Shi, X. A Hybrid Evaluation of the Intestinal Absorption Performance of Compounds from Molecular Structure. *Chem Biol Drug Des* **2024**, *104*, e14576. https://doi.org/10.1111/cbdd.14576.
168. Hildebrand, P.W.; Rose, A.S.; Tiemann, J.K.S. Bringing Molecular Dynamics Simulation Data into View. *Trends in Biochemical Sciences* **2019**, *44*, 902–913. https://doi.org/10.1016/j.tibs.2019.06.004.
169. Hollingsworth, S.A.; Dror, R.O. Molecular Dynamics Simulation for All. *Neuron* **2018**, *99*, 1129–1143.

170. McCammon, J.A.; Gelin, B.R.; Karplus, M. Dynamics of Folded Proteins. *Nature* **1977**, *267*, 585–590. https://doi.org/10.1038/267585a0.
171. Liu, X.; Shi, D.; Zhou, S.; Liu, H.; Liu, H.; Yao, X. Molecular Dynamics Simulations and Novel Drug Discovery. *Expert Opinion on Drug Discovery* **2018**, *13*, 23–37. https://doi.org/10.1080/17460441.2018.1403419.
172. Venable, R.M.; Krämer, A.; Pastor, R.W. Molecular Dynamics Simulations of Membrane Permeability. *Chem. Rev.* **2019**, *119*, 5954–5997. https://doi.org/10.1021/acs.chemrev.8b00486.
173. Wan, G.; Dai, X.; Yin, Q.; Shi, X.; Qiao, Y. Interaction of Menthol with Mixed-Lipid Bilayer of Stratum Corneum: A Coarse-Grained Simulation Study. *J Mol Graph Model* **2015**, *60*, 98–107. https://doi.org/10.1016/j.jmgm.2015.06.005.
174. Wang, H.; Meng, F. The Permeability Enhancing Mechanism of Menthol on Skin Lipids: A Molecular Dynamics Simulation Study. *J Mol Model* **2017**, *23*, 279. https://doi.org/10.1007/s00894-017-3457-y.
175. Dai, X.; Yin, Q.; Wan, G.; Wang, R.; Shi, X.; Qiao, Y. Effects of Concentrations on the Transdermal Permeation Enhancing Mechanisms of Borneol: A Coarse-Grained Molecular Dynamics Simulation on Mixed-Bilayer Membranes. *IJMS* **2016**, *17*, 1349. https://doi.org/10.3390/ijms17081349.
176. Wimley, W.C.; Hristova, K. Antimicrobial Peptides: Successes, Challenges and Unanswered Questions. *J Membr Biol* **2011**, *239*, 27–34. https://doi.org/10.1007/s00232-011-9343-0.
177. Sager, J.E.; Yu, J.; Ragueneau-Majlessi, I.; Isoherranen, N. Physiologically Based Pharmacokinetic (PBPK) Modeling and Simulation Approaches: A Systematic Review of Published Models, Applications, and Model Verification. *Drug Metab Dispos* **2015**, *43*, 1823–1837. https://doi.org/10.1124/dmd.115.065920.
178. Kuepfer, L.; Niederalt, C.; Wendl, T.; Schlender, J.-F.; Willmann, S.; Lippert, J.; Block, M.; Eissing, T.; Teutonico, D. Applied Concepts in PBPK Modeling: How to Build a PBPK/PD Model. *CPT Pharmacometrics Syst Pharmacol* **2016**, *5*, 516–531. https://doi.org/10.1002/psp4.12134.
179. Upton, R.N.; Foster, D.J.R.; Abuhelwa, A.Y. An Introduction to Physiologically-Based Pharmacokinetic Models. *Paediatr Anaesth* **2016**, *26*, 1036–1046. https://doi.org/10.1111/pan.12995.
180. Lin, L.; Wong, H. Predicting Oral Drug Absorption: Mini Review on Physiologically-Based Pharmacokinetic Models. *Pharmaceutics* **2017**, *9*, 41. https://doi.org/10.3390/pharmaceutics9040041.
181. Lin, W.; Chen, Y.; Unadkat, J.D.; Zhang, X.; Wu, D.; Heimbach, T. Applications, Challenges, and Outlook for PBPK Modeling and Simulation: A Regulatory, Industrial and Academic Perspective. *Pharm Res* **2022**, *39*, 1701–1731. https://doi.org/10.1007/s11095-022-03274-2.
182. Maharaj, A.R.; Edginton, A.N. Physiologically Based Pharmacokinetic Modeling and Simulation in Pediatric Drug Development. *CPT Pharmacometrics Syst Pharmacol* **2014**, *3*, e150. https://doi.org/10.1038/psp.2014.45.
183. Jones, H.; Rowland-Yeo, K. Basic Concepts in Physiologically Based Pharmacokinetic Modeling in Drug Discovery and Development. *CPT Pharmacometrics Syst Pharmacol* **2013**, *2*, e63. https://doi.org/10.1038/psp.2013.41.
184. Pumkathin, S.; Hanlumyuang, Y.; Wattanathana, W.; Laomettachit, T.; Liangruksa, M. Investigating Pharmacokinetic Profiles of *Centella Asiatica* Using Machine Learning and PBPK Modelling. *Journal of Biopharmaceutical Statistics* **2024**, 1–16. https://doi.org/10.1080/10543406.2024.2358797.
185. Jin, J.; Xu, X.; Li, F.; Weng, F.; Zou, B.; Li, Y.; Zhao, J.; Zhang, S.; Yan, D.; Qiu, F. Physiologically Based Pharmacokinetic Modeling for Confirming the Role of CYP3A1/2 and P-Glycoprotein in Detoxification Mechanism between Glycyrrhizic Acid and Aconitine in Rats. *J Appl Toxicol* **2024**, *44*, 978–989. https://doi.org/10.1002/jat.4595.
186. Tanveer, A.; Hussain, K.; Tasneem, H.; Arif, I.; Rashid, M.; Abbas, N.; Shamim, R.; Shah, P.A.; Bukhari, N.I. Prediction of CYP-Mediated Silybin A-Losartan Pharmacokinetic Interactions Using Physiological Based Pharmacokinetic Modeling. *J Pharmacokinet Pharmacodyn* **2022**, *49*, 311–323. https://doi.org/10.1007/s10928-022-09804-0.
187. Chen, L.; Ji, N.; Zhang, M.; Chen, W. The Influence of Wuzhi Capsule on the Pharmacokinetics of Cyclophosphamide. *Recent Pat Anticancer Drug Discov* **2022**, *17*, 195–203. https://doi.org/10.2174/1574892816666211110152119.

188. Pilla Reddy, V.; Jo, H.; Neuhoff, S. Food Constituent- and Herb-Drug Interactions in Oncology: Influence of Quantitative Modelling on Drug Labelling. *Br J Clin Pharmacol* **2021**, *87*, 3988–4000. https://doi.org/10.1111/bcp.14822.
189. Lehmann, A.; Geburek, I.; Hessel-Pras, S.; Enge, A.-M.; Mielke, H.; Müller-Graf, C.; Kloft, C.; Hethey, C. PBTK Model-Based Analysis of CYP3A4 Induction and the Toxicokinetics of the Pyrrolizidine Alkaloid Retrorsine in Man. *Arch Toxicol* **2024**, *98*, 1757–1769. https://doi.org/10.1007/s00204-024-03698-2.
190. Rivero-Segura, N.A.; Gomez-Verjan, J.C. In silico Screening of Natural Products Isolated from Mexican Herbal Medicines against COVID-19. *Biomolecules* **2021**, *11*, 216. https://doi.org/10.3390/biom11020216.
191. Li, X.; Chen, C.; Ding, N.; Zhang, T.; Zheng, P.; Yang, M. Physiologically Based Pharmacokinetic Modelling and Simulation to Predict the Plasma Concentration Profile of Schaftoside after Oral Administration of Total Flavonoids of Desmodium Styracifolium. *Front Pharmacol* **2022**, *13*, 1073535. https://doi.org/10.3389/fphar.2022.1073535.
192. Daina, A.; Michielin, O.; Zoete, V. SwissADME: A Free Web Tool to Evaluate Pharmacokinetics, Drug-Likeness and Medicinal Chemistry Friendliness of Small Molecules. *Sci Rep* **2017**, *7*, 42717. https://doi.org/10.1038/srep42717.
193. Pires, D.E.V.; Blundell, T.L.; Ascher, D.B. pkCSM: Predicting Small-Molecule Pharmacokinetic and Toxicity Properties Using Graph-Based Signatures. *J. Med. Chem.* **2015**, *58*, 4066–4072. https://doi.org/10.1021/acs.jmedchem.5b00104.
194. Banerjee, P.; Eckert, A.O.; Schrey, A.K.; Preissner, R. ProTox-II: A Webserver for the Prediction of Toxicity of Chemicals. *Nucleic acids research* **2018**, *46*, W257–W263.
195. Banerjee, P.; Kemmler, E.; Dunkel, M.; Preissner, R. ProTox 3.0: A Webserver for the Prediction of Toxicity of Chemicals. *Nucleic Acids Research* **2024**, *52*, W513–W520. https://doi.org/10.1093/nar/gkae303.
196. Ntie-Kang, F. An in silico Evaluation of the ADMET Profile of the StreptomeDB Database. *SpringerPlus* **2013**, *2*, 353. https://doi.org/10.1186/2193-1801-2-353.
197. Gu, Y.; Yu, Z.; Wang, Y.; Chen, L.; Lou, C.; Yang, C.; Li, W.; Liu, G.; Tang, Y. admetSAR3.0: A Comprehensive Platform for Exploration, Prediction and Optimization of Chemical ADMET Properties. *Nucleic Acids Research* **2024**, *52*, W432–W438. https://doi.org/10.1093/nar/gkae298.
198. Fu, L.; Shi, S.; Yi, J.; Wang, N.; He, Y.; Wu, Z.; Peng, J.; Deng, Y.; Wang, W.; Wu, C.; et al. ADMETlab 3.0: An Updated Comprehensive Online ADMET Prediction Platform Enhanced with Broader Coverage, Improved Performance, API Functionality and Decision Support. *Nucleic Acids Research* **2024**, *52*, W422–W431. https://doi.org/10.1093/nar/gkae236.
199. Xiong, G.; Wu, Z.; Yi, J.; Fu, L.; Yang, Z.; Hsieh, C.; Yin, M.; Zeng, X.; Wu, C.; Lu, A.; et al. ADMETlab 2.0: An Integrated Online Platform for Accurate and Comprehensive Predictions of ADMET Properties. *Nucleic Acids Res* **2021**, *49*, W5–W14. https://doi.org/10.1093/nar/gkab255.
200. Sohlenius-Sternbeck, A.-K.; Terelius, Y. Evaluation of ADMET Predictor in Early Discovery Drug Metabolism and Pharmacokinetics Project Work. *Drug Metab Dispos* **2022**, *50*, 95–104. https://doi.org/10.1124/dmd.121.000552.
201. Lagorce, D.; Bouslama, L.; Becot, J.; Miteva, M.A.; Villoutreix, B.O. FAF-Drugs4: Free ADME-Tox Filtering Computations for Chemical Biology and Early Stages Drug Discovery. *Bioinformatics* **2017**, *33*, 3658–3660. https://doi.org/10.1093/bioinformatics/btx491.
202. Chen, F.; Yin, X.; Wang, Y.; Lv, Y.; Sheng, S.; Ouyang, S.; Zhong, Y. Pharmacokinetics, Tissue Distribution, and Druggability Prediction of the Natural Anticancer Active Compound Cytisine N-Isoflavones Combined with Computer Simulation. *Biol Pharm Bull* **2020**, *43*, 976–984. https://doi.org/10.1248/bpb.b20-00004.
203. Tian, H.; Ketkar, R.; Tao, P. ADMETboost: A Web Server for Accurate ADMET Prediction. *J Mol Model* **2022**, *28*, 408. https://doi.org/10.1007/s00894-022-05373-8.
204. Wei, Y.; Li, S.; Li, Z.; Wan, Z.; Lin, J. Interpretable-ADMET: A Web Service for ADMET Prediction and Optimization Based on Deep Neural Representation. *Bioinformatics* **2022**, *38*, 2863–2871. https://doi.org/10.1093/bioinformatics/btac192.
205. Zhai, J.; Man, V.H.; Ji, B.; Cai, L.; Wang, J. Comparison and Summary of in silico Prediction Tools for CYP450-Mediated Drug Metabolism. *Drug Discovery Today* **2023**, *28*, 103728. https://doi.org/10.1016/j.drudis.2023.103728.

206. Testa, B.; Crivori, P.; Reist, M.; Carrupt, P.-A. The Influence of Lipophilicity on the Pharmacokinetic Behavior of Drugs: Concepts and Examples. *Perspectives in Drug Discovery and Design* **2000**, *19*, 179–211.
207. Chen, L.; Lian, G.; Han, L. Prediction of Human Skin Permeability Using Artificial Neural Network (ANN) Modeling. *Acta Pharmacologica Sinica* **2007**, *28*, 591–600. https://doi.org/10.1111/j.1745-7254.2007.00528.x.
208. Chang, Y.-C.; Chen, C.-P.; Chen, C.-C. Predicting Skin Permeability of Chemical Substances Using a Quantitative Structure-Activity Relationship. *Procedia Engineering* **2012**, *45*, 875–879. https://doi.org/10.1016/j.proeng.2012.08.252.
209. Potts, R.O.; Guy, R.H. Predicting Skin Permeability. *Pharm Res* **1992**, *9*, 663–669. https://doi.org/10.1023/a:1015810312465.
210. An, N.; Machado, J.-H.; Tang, Y.; Kostal, J.; Voutchkova-Kostal, A. Predicting Skin Permeation Rate from Nuclear Magnetic Resonance Spectra. *Green Chem.* **2016**, *18*, 4468–4474. https://doi.org/10.1039/C6GC00945J.
211. Daina, A.; Zoete, V. A BOILED-Egg To Predict Gastrointestinal Absorption and Brain Penetration of Small Molecules. *ChemMedChem* **2016**, *11*, 1117–1121. https://doi.org/10.1002/cmdc.201600182.
212. Sharom, F.J. The P-Glycoprotein Multidrug Transporter. *Essays Biochem* **2011**, *50*, 161–178. https://doi.org/10.1042/bse0500161.
213. Swiss Institute of Bioinformatics. SwissADME. Frequently Asked Questions.
214. Walters, W.P. Going Further than Lipinski's Rule in Drug Design. *Expert Opinion on Drug Discovery* **2012**, *7*, 99–107. https://doi.org/10.1517/17460441.2012.648612.
215. Ghose, A.K.; Viswanadhan, V.N.; Wendoloski, J.J. A Knowledge-Based Approach in Designing Combinatorial or Medicinal Chemistry Libraries for Drug Discovery. 1. A Qualitative and Quantitative Characterization of Known Drug Databases. *Journal of combinatorial chemistry* **1999**, *1*, 55–68.
216. Ursu, O.; Rayan, A.; Goldblum, A.; Oprea, T.I. Understanding Drug-likeness. *WIREs Comput Mol Sci* **2011**, *1*, 760–781. https://doi.org/10.1002/wcms.52.
217. Egan, W.J.; Merz, K.M.; Baldwin, J.J. Prediction of Drug Absorption Using Multivariate Statistics. *Journal of medicinal chemistry* **2000**, *43*, 3867–3877.
218. Muegge, I.; Heald, S.L.; Brittelli, D. Simple Selection Criteria for Drug-like Chemical Matter. *Journal of medicinal chemistry* **2001**, *44*, 1841–1846.
219. Martin, Y.C. A Bioavailability Score. *Journal of medicinal chemistry* **2005**, *48*, 3164–3170.
220. Baell, J.B.; Holloway, G.A. New Substructure Filters for Removal of Pan Assay Interference Compounds (PAINS) from Screening Libraries and for Their Exclusion in Bioassays. *J. Med. Chem.* **2010**, *53*, 2719–2740. https://doi.org/10.1021/jm901137j.
221. Brenk, R.; Schipani, A.; James, D.; Krasowski, A.; Gilbert, I.H.; Frearson, J.; Wyatt, P.G. Lessons Learnt from Assembling Screening Libraries for Drug Discovery for Neglected Diseases. *ChemMedChem* **2008**, *3*, 435–444. https://doi.org/10.1002/cmdc.200700139.
222. Teague, S.J.; Davis, A.M.; Leeson, P.D.; Oprea, T. The Design of Leadlike Combinatorial Libraries. *Angew Chem Int Ed Engl* **1999**, *38*, 3743–3748. https://doi.org/10.1002/(SICI)1521-3773(19991216)38:24<3743::AID-ANIE3743>3.0.CO;2-U.
223. Duffy, E.M.; Jorgensen, W.L. Prediction of Properties from Simulations: Free Energies of Solvation in Hexadecane, Octanol, and Water. *J. Am. Chem. Soc.* **2000**, *122*, 2878–2888. https://doi.org/10.1021/ja993663t.
224. Jorgensen, W.L.; Duffy, E.M. Prediction of Drug Solubility from Structure. *Advanced Drug Delivery Reviews* **2002**, *54*, 355–366. https://doi.org/10.1016/S0169-409X(02)00008-X.
225. Jorgensen, W.L.; Duffy, E.M. Prediction of Drug Solubility from Monte Carlo Simulations. *Bioorganic & Medicinal Chemistry Letters* **2000**, *10*, 1155–1158. https://doi.org/10.1016/S0960-894X(00)00172-4.
226. Fioravanzo, E.; Cazzolla, N.; Durando, L.; Ferrari, C.; Mabilia, M.; Ombrato, R.; Parenti, M.D. General and Independent Approaches to Predict HERG Affinity Values. *Internet Electron J Mol Des* **2005**, *4*, 625–646.
227. Wang, J.; Du, H.; Yao, X.; Hu, Z. Using Classification Structure Pharmacokinetic Relationship (SCPR) Method to Predict Drug Bioavailability Based on Grid-Search Support Vector Machine. *Analytica Chimica Acta* **2007**, *601*, 156–163. https://doi.org/10.1016/j.aca.2007.08.040.

228. Oja, M.; Maran, U. Quantitative Structure–Permeability Relationships at Various pH Values for Neutral and Amphoteric Drugs and Drug-like Compounds. *SAR and QSAR in Environmental Research* **2016**, *27*, 813–832. https://doi.org/10.1080/1062936X.2016.1238408.
229. Fatima, S.; Gupta, P.; Sharma, S.; Sharma, A.; Agarwal, S.M. ADMET Profiling of Geographically Diverse Phytochemical Using Chemoinformatic Tools. *Future Med. Chem.* **2020**, *12*, 69–87. https://doi.org/10.4155/fmc-2019-0206.
230. Athar, M.; Sona, A.N.; Bekono, B.D.; Ntie-Kang, F. Fundamental Physical and Chemical Concepts behind "Drug-Likeness" and "Natural Product-Likeness." *Physical Sciences Reviews* **2019**, *4*. https://doi.org/10.1515/psr-2018-0101.
231. Wang, J.; Hou, T. Recent Advances on in silico ADME Modeling. *Annual reports in computational chemistry* **2009**, *5*, 101–127.
232. Bhat, J.; Chatterjee, S. Skeleton Selectivity in Complexation of Chelerythrine and Chelerythrine-like Natural Plant Alkaloids with the G-Quadruplex Formed at the Promoter of c-MYC Oncogene: In silico Exploration. *RSC Adv.* **2016**, *6*, 36667–36680. https://doi.org/10.1039/C6RA04671A.
233. Anna Malinowska, M.; Billet, K.; Drouet, S.; Munsch, T.; Unlubayir, M.; Tungmunnithum, D.; Giglioli-Guivarc'h, N.; Hano, C.; Lanoue, A. Grape Cane Extracts as Multifunctional Rejuvenating Cosmetic Ingredient: Evaluation of Sirtuin Activity, Tyrosinase Inhibition and Bioavailability Potential. *Molecules* **2020**, *25*. https://doi.org/10.3390/molecules25092203.
234. Skolnik, S.; Lin, X.; Wang, J.; Chen, X.-H.; He, T.; Zhang, B. Towards Prediction of In vivo Intestinal Absorption Using a 96-Well Caco-2 Assay. *Journal of Pharmaceutical Sciences* **2010**, *99*, 3246–3265. https://doi.org/10.1002/jps.22080.
235. Ntie-Kang, F.; Lifongo, L.L.; Judson, P.N.; Sippl, W.; Efange, S.M.N. How "Drug-like" Are Naturally Occurring Anti-Cancer Compounds? *J Mol Model* **2014**, *20*, 2069. https://doi.org/10.1007/s00894-014-2069-z.
236. Veber, D.F.; Johnson, S.R.; Cheng, H.-Y.; Smith, B.R.; Ward, K.W.; Kopple, K.D. Molecular Properties That Influence the Oral Bioavailability of Drug Candidates. *J. Med. Chem.* **2002**, *45*, 2615–2623. https://doi.org/10.1021/jm020017n.
237. Avdeef, A.; Tam, K.Y. How Well Can the Caco-2/Madin–Darby Canine Kidney Models Predict Effective Human Jejunal Permeability? *J. Med. Chem.* **2010**, *53*, 3566–3584. https://doi.org/10.1021/jm901846t.
238. Ibezim, A.; Debnath, B.; Ntie-Kang, F.; Mbah, C.J.; Nwodo, N.J. Binding of Anti-Trypanosoma Natural Products from African Flora against Selected Drug Targets: A Docking Study. *Med Chem Res* **2017**, *26*, 562–579. https://doi.org/10.1007/s00044-016-1764-y.
239. Bolelli, K.; Ertan-Bolelli, T.; Unsalan, O.; Altunayar-Unsalan, C. Fenoterol and Dobutamine as SARS-CoV-2 Main Protease Inhibitors: A Virtual Screening Study. *Journal of Molecular Structure* **2021**, *1228*, 129449. https://doi.org/10.1016/j.molstruc.2020.129449.
240. Ntie-Kang, F.; Zofou, D.; Babiaka, S.B.; Meudom, R.; Scharfe, M.; Lifongo, L.L.; Mbah, J.A.; Mbaze, L.M.; Sippl, W.; Efange, S.M.N. AfroDb: A Select Highly Potent and Diverse Natural Product Library from African Medicinal Plants. *PLoS ONE* **2013**, *8*, e78085. https://doi.org/10.1371/journal.pone.0078085.
241. Malik, R.; Bunkar, D.; Choudhary, B.S.; Srivastava, S.; Mehta, P.; Sharma, M. High Throughput Virtual Screening and in silico ADMET Analysis for Rapid and Efficient Identification of Potential PAP248-286 Aggregation Inhibitors as Anti-HIV Agents. *Journal of Molecular Structure* **2016**, *1122*, 239–246. https://doi.org/10.1016/j.molstruc.2016.05.086.
242. Shelke, S.M.; Bhosale, S.H. Synthesis, Antidepressant Evaluation and QSAR Studies of Novel 2-(5H-[1,2,4]Triazino [5,6-b]Indol-3-Ylthio)-N-(Substituted Phenyl)Acetamides. *Bioorganic & Medicinal Chemistry Letters* **2010**, *20*, 4661–4664. https://doi.org/10.1016/j.bmcl.2010.05.100.
243. West, C.; Zhang, Y.; Morin-Allory, L. Insights into Chiral Recognition Mechanisms in Supercritical Fluid Chromatography. I. Non-Enantiospecific Interactions Contributing to the Retention on Tris-(3,5-Dimethylphenylcarbamate) Amylose and Cellulose Stationary Phases. *Journal of Chromatography A* **2011**, *1218*, 2019–2032. https://doi.org/10.1016/j.chroma.2010.11.084.
244. Wang, J.; Xie, X.-Q.; Hou, T.; Xu, X. Fast Approaches for Molecular Polarizability Calculations. *J. Phys. Chem. A* **2007**, *111*, 4443–4448. https://doi.org/10.1021/jp068423w.

245. Janicka, M.; Mycka, A.; Sztanke, M.; Sztanke, K. Predicting Pharmacokinetic Properties of Potential Anticancer Agents via Their Chromatographic Behavior on Different Reversed Phase Materials. *IJMS* **2021**, *22*, 4257. https://doi.org/10.3390/ijms22084257.
246. Tandon, H.; Ranjan, P.; Chakraborty, T.; Suhag, V. Polarizability: A Promising Descriptor to Study Chemical–Biological Interactions. *Mol Divers* **2021**, *25*, 249–262. https://doi.org/10.1007/s11030-020-10062-w.
247. Pitsillou, E.; Liang, J.J.; Beh, R.C.; Prestedge, J.; Catak, S.; Hung, A.; Karagiannis, T.C. Identification of Novel Bioactive Compounds from Olea Europaea by Evaluation of Chemical Compounds in the OliveNet™ Library: In silico Bioactivity and Molecular Modelling, and in Vitro Validation of hERG Activity. *Computers in Biology and Medicine* **2022**, *142*, 105247. https://doi.org/10.1016/j.compbiomed.2022.105247.
248. Onguéné, P.A.; Ntie-Kang, F.; Mbah, J.A.; Lifongo, L.L.; Ndom, J.C.; Sippl, W.; Mbaze, L.M. The Potential of Anti-Malarial Compounds Derived from African Medicinal Plants, Part III: An in silico Evaluation of Drug Metabolism and Pharmacokinetics Profiling. *Org Med Chem Lett* **2014**, *4*, 6. https://doi.org/10.1186/s13588-014-0006-x.
249. Banerjee, P.; Dehnbostel, F.O.; Preissner, R. Prediction Is a Balancing Act: Importance of Sampling Methods to Balance Sensitivity and Specificity of Predictive Models Based on Imbalanced Chemical Data Sets. *Frontiers in Chemistry* **2018**, *6*. https://doi.org/10.3389/fchem.2018.00362.
250. Lombardo, F.; Desai, P.V.; Arimoto, R.; Desino, K.E.; Fischer, H.; Keefer, C.E.; Petersson, C.; Winiwarter, S.; Broccatelli, F. 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: Miniperspective. *J. Med. Chem.* **2017**, *60*, 9097–9113. https://doi.org/10.1021/acs.jmedchem.7b00487.