Section III

Cheminformatics - Advanced: *The role of pharmacokinetics/toxicity and the concept of drug-likeness*

Andrei Leitão, Dan C. Fara and Tudor I. Oprea

Up to now many pharmacodynamic properties were taken into account, basically considering the ligand-target interactions. Nonetheless, a good drug candidate does not rely only on a good fit and the best match of hydrophobic and lipophilic interactions with the target macromolecule. In fact, there must be a compromise with many different properties that compose the pharmacokinetic profile of a compound. While pharmacodynamics deals with the compound-target interaction, the pharmacokinetics takes into consideration all the steps that the compound should go to reach the site of interaction – see Figure 1[1]. A compound must be bioavailable in a good extend in order to reach the site of action. For oral drugs, it means that it should be soluble and permeate through the enterocyte, being metabolized into the liver and distributed into the bloodstream.

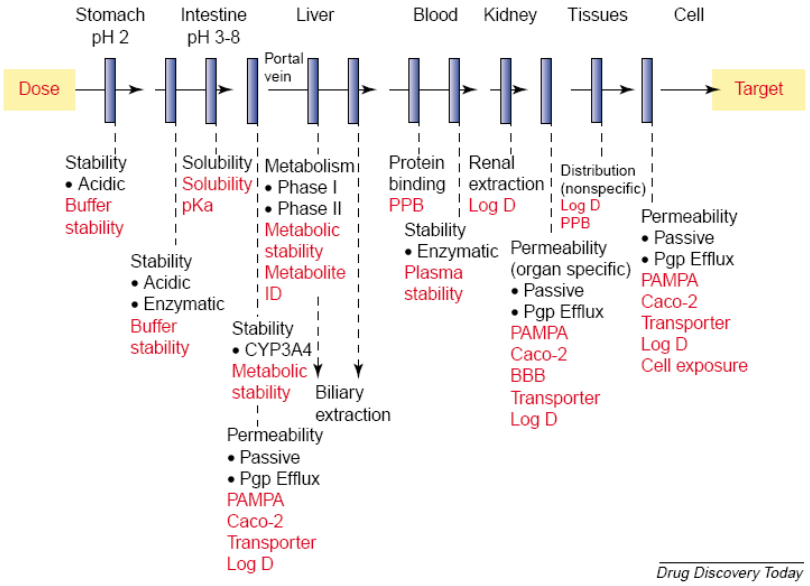


Figure 1. Representation of the pharmacokinetics process and different models in use to study the pharmaceutical profile (taken from reference 1).

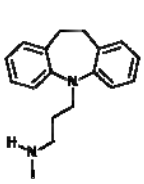
Basically, the pharmacokinetics is composed by the absorption, distribution, metabolism and excretion. Together with the toxicity studies, they are the so called ADME-Tox (also found as ADMET). This plethora of events that takes place before the pharmacodynamics represents nowadays the main drawback in the drug development phase, increasing the attrition rate to the process (i.e. rate of failure to develop a new compound).[2]

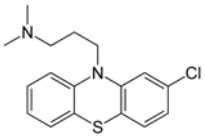
The high rate of failure impairs the development of many new compounds that are not considered to be attractive drug candidates after the pharmaceutical companies have spent years of work and millions of dollars. Therefore, it is expected that the integration of the ADME-Tox knowledge and other global properties (drug-likeness, among others) can streamline the drug discovery & development pipeline.[3]

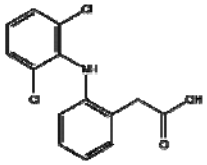
Early ADME studies in the end of the nineteen century showed that there was a correlation between the narcotic property of some compounds and the solubility in oil, but the first quantitative evidence came with the Hansch studies in the 60's, which showed the correlation between the logarithmic of the octanol/water partition coefficient (log P or log $P_{o/w}$) and the biological activity. The octanol-water system is considered the first useful model to the partition in the membrane of a cell. The method is straightforward, where a compound is added to a separation funnel with two liquids (water and octanol) and, after shacking it, the concentration of the molecule in each liquid is measured. Hydrophilic compounds tend to stay in the aqueous solution, while more hydrophobic ones are buried in the octanol moiety. The result of this balance is described by *P* (the partition parameter) meaning that the lower it goes, the more compound accumulates into the hydrophilic part (consequently it can be considered as a hydrophilic molecule).

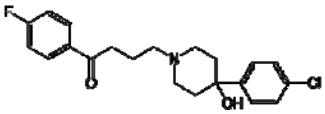
Many studies throughout the last and this century reinforce the importance of taking into account the log P of a compound as a good metric of the compound permeability profile. Too hydrophilic compounds should not be considered good drug candidates because they are not able to cross the membrane using the passive diffusion. On the other hand, too hydrophobic compounds tend to be insoluble, forming aggregates, which are font of the many false-positive results got in the high-throughput screening (HTS). Studies with different classes of compounds showed that CNS drugs (i.e. molecules that act at the Central Nervous System) use to have higher log P value than other classes of drugs, once they should cross the Blood-Brain Barrier (BBB).

Methods for calculating the log P are now available, based on different sets of rules and concepts. The most popular one is known as Clog P, which is based on the Hansch and Leo studies on thousands of molecules and comprise the biggest measured log P dataset available so far – see Figure 2.


Desipramine
4.5


Chlorpromazine
5.3


Acetaminofen
4.7


Haloperidol
3.8

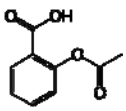

Aspirin (acetylsalicylic acid)
1.2

Figure 2. Clog P values of some drugs

But, is there an optimal range for log P? This question arose as expected in the scientific community and, together with many others, led to the research on this still very young field. Studies have been evolving from the description of these phenomena to the collection of data and standardization of methods in order to select and, if possible, predict better molecules based on computational models. The seminal work of Lipinski published in 1997 shed light to this complex problem by means of a truly easy way.[4] Analyzing the set of oral drugs in the market at that time using the following calculated parameters (log P, molecular weight = MW, number of hydrogen bond donors = HBD, and hydrogen bond acceptors = HBA) it came to the conclusion that:

1. there are some optimal ranges for oral drugs: MW ≤ 500, HBA ≤ 10, HBD ≤ 5, ClogP ≤ 5;
2. oral drugs usually do not violate more than one these ranges.

This simple set of descriptors could define the profile of oral drugs and, because of its multiplicity of five, it was so called 'rule of five' (and sometimes Lipinski's rule). The main objective of this work was to advise the

scientific community about the usefulness of such a set of descriptors as a guide to the drug discovery of new molecules, avoiding structures with are far outside these limits, because the higher probability of the attrition rate in the further steps of the process. By doing that, the new concept of drug-likeness (i.e. compounds that have properties similar to drugs) was born. Further studies in this area not only expanded the rule of five and better defined the drug-like term, but also provided similar concepts for leads[5] and chemical fragments.

Many other models are available to describe the permeability of a drug. For example, Caco-2 and MDCK cell lines and PAMPA (Parallel Artificial Membrane Permeation Assay) are used to check if a compound can permeate the membrane and also the mechanism. Nowadays, these cellular assays can be done under a HTS platform, where compound libraries are tested. The main drawback of these cellular systems is the time to grow the cell monolayer, once that the assay times are short.[6]

The compound metabolism can be studied using microsomes (cellular assay) and/or CYP450 (Cytocrome P450) enzymes in enzymatic assays. Nowadays, molecular modeling is also another tool used to understand the mechanism of catalysis provided by these enzymes, considering also the wide range of scaffolds that are substrates of those enzymes. Of particular concern today are the enzymes CYP405 2D6 and 3A4, which are responsible for the metabolism of many drug classes. Some models were also developed based on the current knowledge in the field in order to predict drug metabolism, but are still limited to some classes and enzymes, due to the limitation of the databases. This is the case of MetaSite, new software that was built based on the CYP2D6 and CYP3A4 three-dimensional structures together with the metabolic profile of some compound classes.[7] Certainly, these models still deserve improvement, but they are providing to be useful in a near future – see Figure 3.

Some models for distribution and toxicity are still in the beginning. It is especially hard to figure out the toxic profile of a compound. The excretion of compounds usually depends on the metabolism and there are no good models that can be used in a HTS assay.

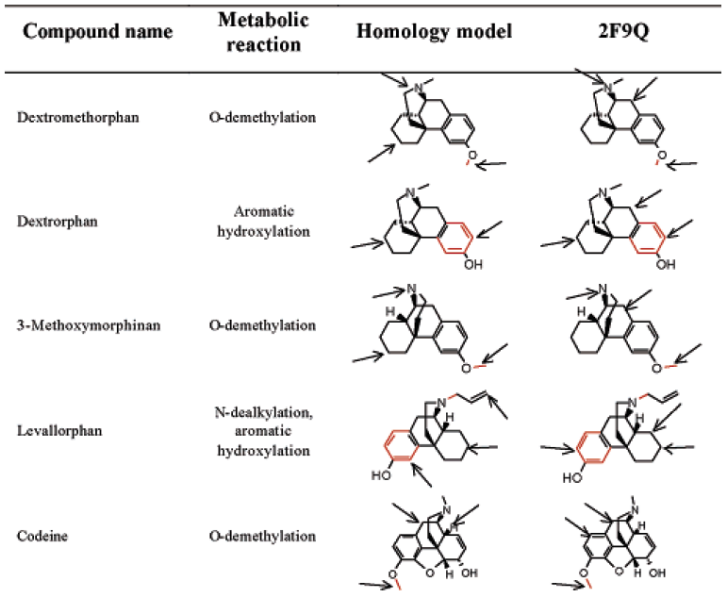


Figure 3. Prediction of metabolic sites for some compounds based on homology model and the three-dimensional structure of the CYP450 2D6

[1] Kerns, E. H.; Di. L.; Pharmaceutical Profiling in Drug discovery, *Drug Discovery Today* **2003**, *8*, 316-323

[2] Kubinyi, H.; Drug Research: Myths, Hype and Reality, *Nature Reviews Drug Discovery* **2003**, *2*, 665-668

[3] Oprea, T. I.; Virtual Screening in Lead Discovery: A Viewpoint, *Molecules* **2002**, *7*, 51-62

[4] Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J.; Experimental and Computational Approaches to Estimate Solubility and Permeability in Drug Discovery and Development Settings, *Advanced Drug Delivery Reviews* **1997**, *23*, 3-25

[5] Oprea, T. I.; Davis, A. M.; Teague, S. J.; Leeson, P. D.; Is There a Difference between Leads and Drugs? A Historical Perspective, *Journal of Chemical Information and Computer Sciences* **2001**, *41*, 1308-1315

[6] Balimane, P. V.; Chong, S.; Cell Culture-Based Models for Intestinal Permeability: A Critique, *Drug Discovery Today* **2005**, *10*, 335-343

[7] Kjellander, B.; Masimirembwa, C. M.; Zamora, I.; Exploration of Enzyme-Ligand Interactions in CYP2D6 & 3A4 Homology Models and Crystal Structures Using a Novel Computational Approach, *Journal of Chemical Information and Modeling* **2007**, *47*, 1234-1247