Log P as a tool in intramolecular hydrogen bond considerations

Log P as a tool in intramolecular hydrogen bond considerations

Giulia Caron, Maura Vallaro and Giuseppe Ermondi

Molecular Biotechnology and Health Sciences Dept., Università degli Studi di Torino, via Quarello 15, 10135 Torino, Italy.

E-mail: giulia.caron@unito.it, telephone: +39 011 6708337

Abstract

Intramolecular hydrogen bonding (IMHB) considerations are gaining relevance in drug discovery and a molecular descriptor which can predict very early the capacity of a compound to form IMHB is needed to speed up the optimization process of drug candidates.

Although log P_{oct} is largely used for optimization purposes, in this paper we firstly use the Block Relevance (BR) analysis to theoretically show how log P_{oct} is not a convenient choice to assess IMHB properties of candidates. Then we discuss the limits of log P_{oct} and introduce Δlog $P_{oct\text{-}tol}$, i.e. the difference between log P_{oct} and log P_{tol} (the logarithm of the partition coefficient in the toluene/water system). Finally, we provided some examples also including bRo5 protease inhibitors, to clarify how to interpret Δlog $P_{oct\text{-}tol}$ values.

Preface

Lipophilicity represents the affinity of a molecule or a moiety for a lipophilic environment. It is commonly measured by the partition coefficient P defined as the ratio between the concentration of the solute in the partition solvent (immiscible with water) and its concentration in water [1]. Octanol is considered the partition solvent by most researchers in drug discovery thus log P_{oct} is often simply written as log P. Here the partition solvent will be explicitly indicated in subscript since more solvents and thus more log Ps will be discussed in the paper. The second descriptor of lipophilicity is the distribution coefficient, expressed as D^{pH}, and its logarithm (log D^{pH}). It is a pH dependent descriptor for ionizable solutes and results from the weighted contributions of all ionised forms present at the indicated pH [1]. Since in this study we refer to neutral compounds, log P_{oct} = log $D^{pH}{}_{oct}$ all along the manuscript.

log P_{oct}: an established tool in drug discovery

The role of lipophilicity descriptors in determining the overall quality of drug candidates is of paramount importance [2] and in fact log P is implemented in the Lipinski's rule of five (Ro5) [3]. The Ro5 states that an orally active drug has no more than one violation of the following criteria: no more than 5 hydrogen bond donors (HBD, the total number of nitrogen–hydrogen and oxygen–hydrogen bonds), no more than 10 hydrogen bond acceptors (HBA, all nitrogen or oxygen atoms), a molecular mass less than 500 Daltons and a log P_{oct} not greater than 5. To identify opportunities for oral drug discovery beyond the Ro5 (bRo5), it has been suggested that oral druggable space bRo5 may extend up to a log P_{oct} of 6 [4].

In addition to implementation in the Ro5, log P_{oct} is also applied as indicators of project progresses through the lipophilicity efficiency (LipE) equation [5]. The concept of LipE allows medicinal chemists to normalize the observed potency with changes in lipophilicity and thus LipE is used in the evaluation of specific structural modifications during the progression of a chemical series (e.g. the development of homologues).

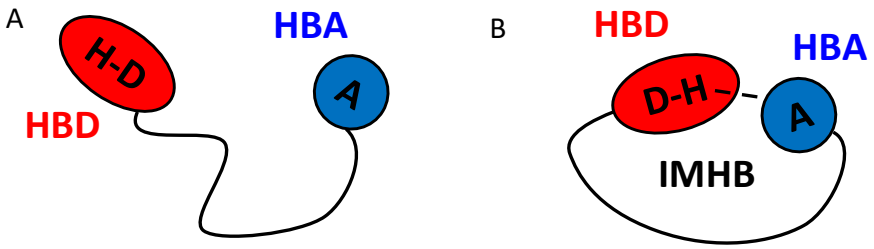
Finally, increasing lipophilicity is one of the potential approaches used by many medicinal chemists to improve passive permeability (an example is discussed in [6]), whereas the hERG-mediated cardiovascular liability is traditionally avoided by lowering lipophilicity and eliminating basic amine functionality (for instance as discussed by Dow [7]). Lipophilicity modulation is also applied to limit metabolism issues.

Summing up, the role of lipophilicity in determining the overall quality of drug candidates is well established and log P_{oct} and log D_{oct} are routinely used in drug discovery programs for optimization purposes.

log P_{oct} and IMHB

An intramolecular hydrogen bond (IMHB) occurs when a hydrogen bond (HB) is formed between a hydrogen bond donor (HBD) and a hydrogen bond acceptor (HBA) that belong to the same molecule (Figure 1).

Figure 1. A molecule with a hydrogen donor group HBD (D-H) and a hydrogen bond acceptor group HBA (A) can exist in two forms A) open (extended), in which the HBA and HBD moieties are exposed and B) closed (folded), in which an IMHB is formed and masks the HB properties of the HBA and HBD moieties.



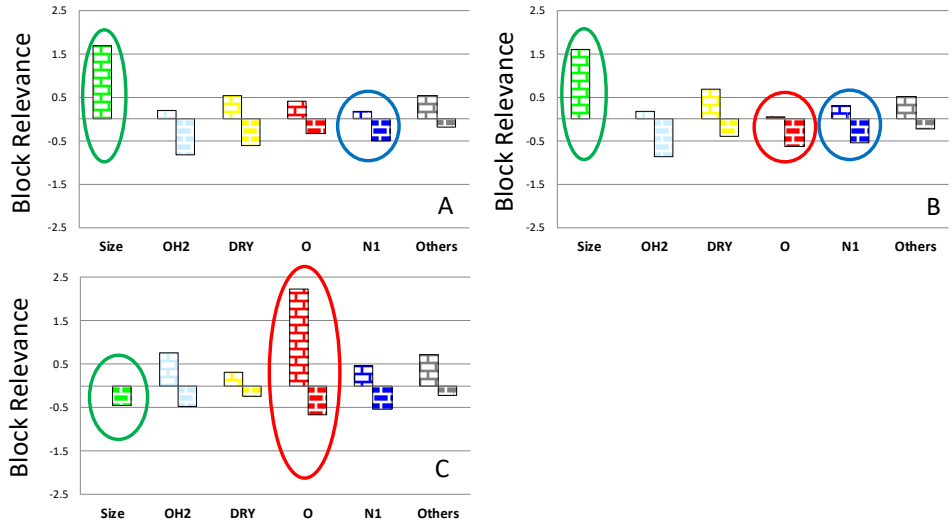
Giordanetto and other researchers recently figured out the formation or disruption IMHB as one valuable strategy to transform molecules into drugs [8] and suggested to include IMHB considerations in drug discovery programs [9]. In fact, IMHB could improve cell permeability and oral bioavailability without necessarily decreasing solubility and binding affinity [10] [11] [12]. For practical purposes it is therefore important to find a descriptor (i.e. a number) which quantifies how easily a molecule can form IMHBs.

In principle, lipophilicity in octanol/water could be used to describe the propensity of compounds to form IMHBs since the formation of IMHBs is expected to mask the polarity of the HBA and HBD moieties (Figure 1) and thus increase log P_{oct}.

However, log P_{oct} is not a convenient choice to assess IMHB properties. This is revealed by a Quantitative Structure-Property Relationship (QSPR) where log P_{oct} was related to a set of Volsurf+ descriptors (https://www.moldiscovery.com/) by a Partial Least Square (PLS) algorithm and the final model was analysed using the BR analysis [13]. Shortly, BR analysis shows in one picture (see Fig. 2A for the meaning of the graphs at a glance) the molecular features (e.g. solutes dimensions and HB properties) governing log P_{oct} : blocks with positive weighting (e.g. the green block) increase log P_{oct} (the higher, the more), whereas those with negative weighting indicate how much the property decreases log P_{oct} (the lower, the less). Moreover, blocks with comparable positive and negative contributions are poorly relevant in the description of log P_{oct}. From BR analysis one can appreciate the dominant impact of the green Size/Shape block (from here on called Size) which is the major limit in the use of log P_{oct} as a tool for IMHB considerations. In fact, log P_{oct} mostly depends on solutes' dimensions rather than HB properties, the red and the blue blocks in Fig. 2A are due to HB properties and are small. In practice, to use log P_{oct} to explore IMHB properties, you need to work in a pairwise fashion: the couple of compounds should be formed by structurally related molecules with similar size, one of them can form IMHBs (the sample), the other not (the control). Some examples are provided below in the *Case studies* section.

BR analysis also highlights a second limit of log P_{oct} as a tool for IMHB considerations. Fig. 2A shows that the red block (HBD properties) is split in two components with opposite sign. This means that H-bond donor groups do not force compounds to stay in the aqueous phase but instead they support an equal partition into octanol and water. Therefore, the formation of an IMHB which mostly reduces HDB properties exposure does not have a strong impact on log P_{oct}.

Figure 2. Some examples of BR analysis output : A) log P_{oct} [13], B) log P_{tol} [13] and C) Δlog$P_{oct\text{-}tol}$ [14]. Block name and color codes: size/shape (Size, green), molecular polarity (OH2, cyan), hydrophobicity (DRY, yellow), solutes' hydrogen bond donor properties (O, red), solutes' hydrogen bond acceptor properties (N1, blue), polarity unbalance (Others, grey). Panel C supports that Δlog $P_{oct\text{-}tol}$ is mostly influenced by solutes' HBD properties (red block, positive sign) and it is poorly influenced by steric descriptors (green block, negative sign). More details about methods and datasets could be foubd in the original literature. Shortly in any paper, a QSPR model was obtained by relating log P_{oct} (log P_{tol} and Δlog$P_{oct\text{-}tol}$) to a set of Volsuf+ descriptors by a PLS algorithm and interpreted using the BR analysis. 3D structures were generated by the builder implemented in VolSurf+ and the presence of IMHBs was neglected.



Going beyond octanol/water and combining log Ps

In the previous section we highlighted that log P_{oct} is not the best tool for IMHB considerations. From BR analysis one could reason to replace octanol/water with a biphasic system in which the Size contribution is less important whereas HBD and HBA blocks are more important. Such a system could be more sensible to characterize the propensity of compounds to form IMHBs.

Log P_{tol}, the logarithm of the partition coefficient P in the toluene/water system was recently introduced in drug discovery by Shalaeva and coworkers [9]. BR analysis (Fig. 2B) revealed that, if compared to log P_{oct}, log P_{tol} shows

a major impact of the HB-related blocks (both the red and the blue blocks are larger) but not a significative decrease of the Size (green block) which remains the most important block.

To limit the impact of Size, a difference between log Ps is expected to provide good results and in fact $\Delta\log P_{oct\text{-}tol}$ (i.e. the difference between log P_{oct} and log P_{tol}) is a clean descriptor of exposed HBD properties [9]. BR analysis supports this expectancy (Fig. 2C) showing that $\Delta\log P_{oct\text{-}tol}$ is mostly influenced by HBD solutes' property (red block, positive sign) and it is poorly influenced by steric descriptors (green block, negative sign) [14]. The presence of IMHBs thus produces low $\Delta\log P_{oct\text{-}tol}$ values.

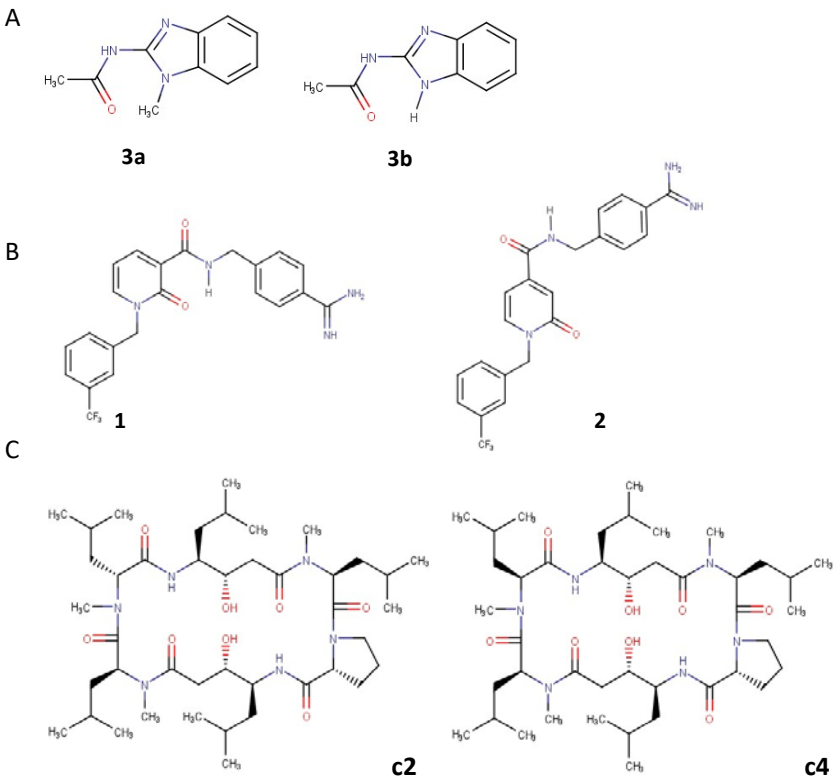
Case studies

As discussed above, the major impact of Size prevents log P_{oct} values from providing clean information about the propensity of compounds to form IMHB and a trick to overpass this issue is pair analysis. This consists in comparing a sample with a substructure prone to IMHB formation along with a control compound with similar structure (and thus Size) but incapable of forming that bond. Two examples are discussed below.

In the first example, we pay our attention on the couple of compounds **3a** and **3b** (Figure 3A) reported in the literature by Kuhn and coworkers [15]. **3b** shows a substructure prone to IMHB formation whereas **3a** (the control) is a compound incapable of forming that bond since a hydrogen atom is substituted by a more lipophilic methyl group. Crystallographic data supported the presence of an IMHB in **3b**. Despite the presence of the N-methyl substituent, log P_{oct} of **3a** (0.68) is lower than log P_{oct} of **3b** (1.39). In **3b** the presence of an IMHB reduces the polarity of the amide and the imidazole N-H and this is revealed by the increase in experimental log P_{oct}. In this case log P_{oct} is in line with NMR data and supports the presence of the IMHB.

In the second example, we focus on compounds **1** (the sample) and **2** (the control). Their chemical structures are in Fig. 3B [16]. EPSA, a supercritical fluid chromatography method specifically developed for the detection of IMHB [17], showed that **1** forms an IMHB. The lipophilicity in octanol/water was described by ElogD, a widely known and validated chromatographic surrogate of log D_{oct} [18]. ElogD is 0.9 for **1** and 0.8 for **2**. These data suggest that two regioisomers **1** and **2** show the same lipophilicity in the octanol/water system and thus this evidence does not support the presence/absence of the IMHB experimentally revealed by another experimental approach.

Figure 3. Compounds taken from the literature to discuss the limits of using log P_{oct} to describe the propensity of compounds to form IMHBs.



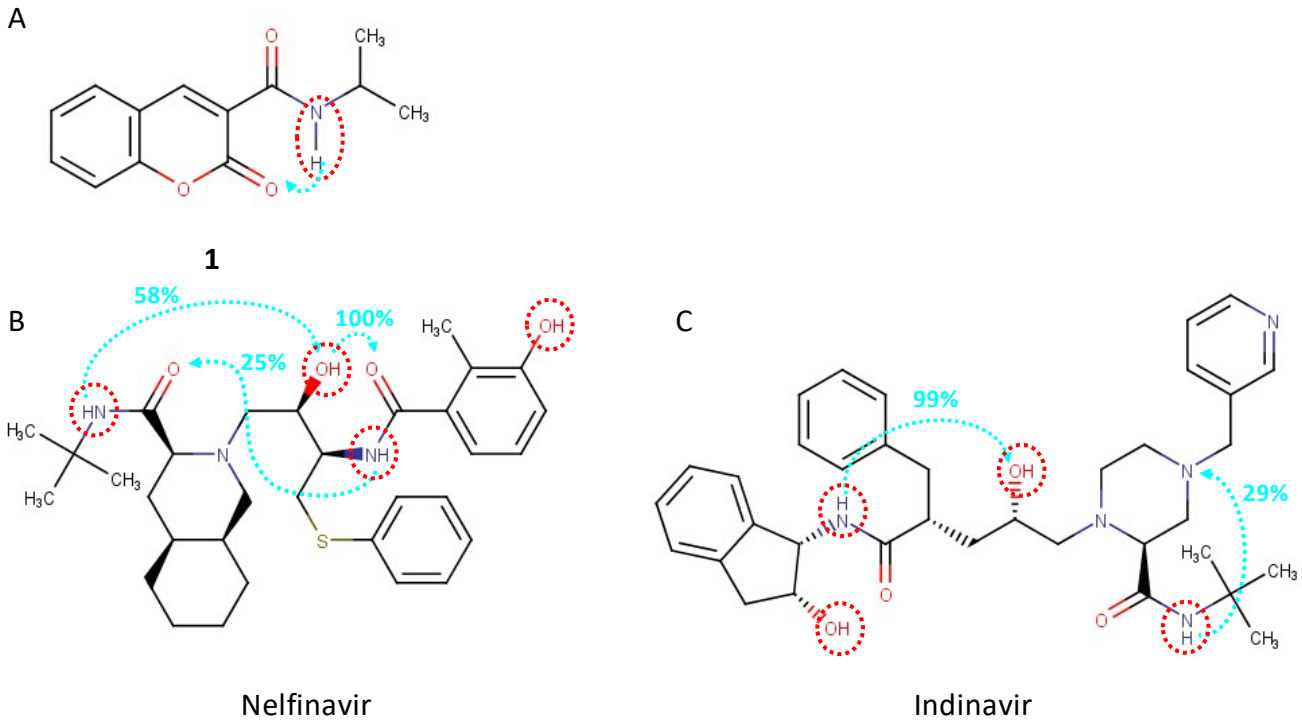
Due to the laboriousness of the experimental measurements, log P_{oct} values are often assessed through calculators. Furthermore, this option is mandatory in very early drug discovery when dealing with virtual structures. A plethora of both commercial and free tools can do that [19]. However, one should be aware that most log P_{oct} calculators are 2D tools and thus cannot be used for investigating IMHB formation which is conformation-dependent. An example is provided by Bockus and coworkers who synthetized a series of cyclic hexapeptide diastereomers containing γ-amino acids and determined lipophilicity and permeability properties [20]. Two compounds **c2** and **c4** are in Fig. 3C and showed identical calculated log P value, 4.60 and 2.48 respectively when MoKa (https://www.moldiscovery.com/) and Marvin Suite (https://chemaxon.com/products/marvin) are used, but their experimental ElogD differs by more than one logarithmic unity (5.7 and 6.9). This difference could be ascribed to different IMHB networks as highlighted by NMR studies.

Overall, examples in Fig. 3 support BR analysis results and show that log P_{oct} is not the most effective tool to predict the capacity of compounds to form IMHBs since log P_{oct} can reveal the presence of some but not all IMHBs. Therefore, its application in drug discovery should be supported by additional descriptors.

As discussed above, BR analysis designates $\Delta\log P_{oct\text{-}tol}$ as a pure descriptor of HBD properties. Therefore, if $\Delta\log P_{oct\text{-}tol}$ can be experimentally obtained (solubility issues can limit the determination of $\log P_{tol}$), it can be used to predict the propensity of compounds to form IMHBs. Some examples are discussed below.

When a compound has a single HBD group (and at least one HBA), the application of $\Delta\log P_{oct\text{-}tol}$ is straightforward. If $\Delta\log P_{oct\text{-}tol}$ is close to 0 then the compound has high propensity to form IMHBs since this means that the contribution of the red block (i.e. HBD properties) is neglected. This was recently shown by some of us using **1** (Figure 4A, $\Delta\log P_{oct\text{-}tol}$ = 0.10) as an example [9].

*Figure 4. Chemical structures of compounds discussed in the text to highlight how to interpret $\Delta\log P_{oct\text{-}tol}$ values A) **1** ($\Delta\log P_{oct\text{-}tol}$ = 0.10,* [9]*) 2) B) nelfinavir ($\Delta\log P_{oct\text{-}tol}$ = 0.53) and C) indinavir ($\Delta\log P_{oct\text{-}tol}$ = 1.99). Red dotted circles highlight HBD moieties. Cyan dotted lines show the most probable IMHBs in the molecule. In **1** only one IMHB is possible whereas for indinavir and nelfinavir the probability of formation of different IMHBs is reported (see text for details).*



The interpretation of $\Delta\log P_{oct\text{-}tol}$ is more complex when more HBD groups are present in the chemical structure. Nelfinavir (Fig. 4B) is a protease inhibitor which shows a molecular property profile that is generally considered to be disadvantageous for the membrane permeability and drug absorption [21]. It is therefore important to

experimentally verify whether nelfinavir have propensity to form IMHBs which could modulate its ADME behavior.

Using a SiriusT3 instrument (www.sirius-analytical.com) to perform standard potentiometric titrations, we obtained for nelfinavir a value of $\Delta\log P_{oct\text{-}tol}$ = 0.53 (pK_as= 5.96, 11.97, $\log P_{oct}$ = 4.02, $\log P_{tol}$ = 3.49) which suggests that nelfinavir has a high capacity of forming IMHBs (considering the experimental error in the measure of partition coefficients, a value of $\Delta\log P_{oct\text{-}tol}$ close to 0.5 could be regarded as null). Since nelfinavir has 4 HBD groups, conformational analysis is expected to provide further information about the effective involvement of the different HBD moieties in the formation of IMHBs. The conformational sampling was carried out using standard conformational sampling tools, generally available in most molecular modeling packages. A low dielectric constant (ε=2.02) was used to mimic the apolar environment provided by toluene. The percentage of IMHB formed by each HBD group was evaluated weighing the influence of each conformation assuming a Boltzmann-like distribution of the conformers population. Such a percentage could be considered a rough estimation of the effective participation of HBD groups to an IMHB. Fig. 4B shows that for nelfinavir a hydroxyl group is always involved in the formation of an IMHB both as donor (100%) and acceptor group. Moreover, both the amide moieties participate to IMHB formation. Overall, all HBD groups of nelfinavir are significantly involved in the formation of IMHBs and justify a $\Delta\log P_{oct\text{-}tol}$ value close to 0.

Generally speaking, in the presence of more HBD groups, the experimental $\Delta\log P_{oct\text{-}tol}$ is not often equal or near to 0 and it is not trivial to establish a net threshold that discriminates when IMHBs are present or not. For example, indinavir (Fig. 4C) is another antiviral bearing four HBD groups. For indinavir, we measured $\Delta\log P_{oct\text{-}tol}$ = 1.99 (pK_as = 3.87 and 5.68, $\log P_{oct}$ = 2.85, $\log P_{tol}$ = 0.86). The significative difference in $\Delta\log P_{oct\text{-}tol}$ between indinavir and nelfinavir (which bears the same number and types of HBD groups) suggests that nelfinavir has a major capacity of forming IMHBs than indinavir. To verify whether indinavir forms IMHBs or not, conformational analysis was performed and showed that in indinavir (Fig. 4C), only one of the four HBD groups present in the molecule is involved in the formation of an IMHB (the amide moiety) as a donor group, whereas a hydroxyl group is involved in IMHB as an acceptor group (HBA). This example reveals that a value of $\Delta\log P_{oct\text{-}tol}$ about 2 in the presence of more HBD groups does not permit to exclude the formation of IMHBs and other techniques, experimental (e.g. NMR) or theoretical (e.g. conformational sampling), are required to confirm the hypotheses suggested by $\Delta\log P_{oct\text{-}tol}$.

Overall, these examples support that $\Delta\log P_{oct\text{-}tol}$ is a valuable tool to experimentally assess information about IMHB formation. In many cases, from the numerical value researchers can direct evaluate the skill of compounds to form IMHBs. For complex structures, standard conformational analysis can help $\Delta\log P_{oct\text{-}tol}$ interpretation.

Conclusions

Despite its wide application in research programs, the potential of lipophilicity as an elucidator of structural properties is not fully exploited. In this paper we showed that for a more efficient use of lipophilicity in IMHB considerations, the determination of log P_{oct} is not sufficient. Indeed, we need to go beyond the traditional octanol/water system, set-up a second system with a more apolar organic phase (i.e. toluene/water) and calculate Δlog $P_{oct\text{-}tol}$. This latter can provide a significant amount of information about IMHB formation and be used in early drug discovery.

References

[1] Caron G, Ermondi G, Scherrer RA. Lipophilicity, polarity, and hydrophobicity. In: Testa B, van de Waterbeemd H, editors. Comprehensive medicinal chemistry II, Oxford: Elsevier; 2007, Ch. 5.18, p. 425-452.

[2] Waring MJ. Lipophilicity in drug discovery. Expert Opin Drug Discov 2010;5:235–48.

[3] Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv Drug Deliv Rev 1997;23:3–26.

[4] Pye CR, Hewitt WM, Schwochert J, Haddad TD, Townsend CE, Etienne L, et al. Nonclassical Size Dependence of Permeation Defines Bounds for Passive Adsorption of Large Drug Molecules. J Med Chem 2017;60:1665–72.

[5] Ryckmans T, Edwards MP, Horne VA, Correia AM, Owen DR, Thompson LR, et al. Rapid assessment of a novel series of selective CB2 agonists using parallel synthesis protocols: A Lipophilic Efficiency (LipE) analysis. Bioorganic Med Chem Lett 2009;19:4406–9.

[6] Ammirati M, Bagley SW, Bhattacharya SK, Buckbinder L, Carlo AA, Conrad R, et al. Discovery of an in Vivo Tool to Establish Proof-of-Concept for MAP4K4-Based Antidiabetic Treatment. ACS Med Chem Lett 2015;6:1128–33.

[7] Dow RL, Andrews MP, Li J, Gibbs EM, Guzman-perez A, Laperle JL, et al. Bioorganic & Medicinal Chemistry Defining the key pharmacophore elements of PF-04620110 : Discovery of a potent , orally-active , neutral DGAT-1 inhibitor. Bioorg Med Chem 2013;21:5081–97.

[8] Giordanetto F, Tyrchan C, Ulander J. Intramolecular Hydrogen Bond Expectations in Medicinal Chemistry. ACS Med Chem Lett 2017;8:139-42.

[9] Shalaeva M, Caron G, Abramov YA, Connell TNO, Plummer MS, Yalamanchi G, et al. Integrating Intramolecular Hydrogen Bonding (IMHB) Considerations in Drug Discovery Using Δ logP As a Tool. J. Med. Chem. 2013;56:4870−4879

[10] Rezai T, Bock JE, Zhou M V., Kalyanaraman C, Lokey RS, Jacobson MP. Conformational flexibility, internal hydrogen bonding, and passive membrane permeability: Successful in silico prediction of the relative permeabilities of cyclic peptides. J Am Chem Soc 2006;128:14073–80.

[11] Labby KJ, Xue F, Kraus JM, Ji H, Mataka J, Li H, et al. Intramolecular hydrogen bonding: a potential strategy for more bioavailable inhibitors of neuronal nitric oxide synthase. Bioorg Med Chem 2012;20:2435–43.

[12] Over B, Matsson P, Tyrchan C, Artursson P, Doak BC, Foley MA, et al. Structural and conformational determinants of macrocycle cell permeability. Nat Chem Biol 2016;12:1065–74.

[13] Caron G, Vallaro M, Ermondi G, Goetz GH, Abramov YA, Philippe L, et al. A Fast Chromatographic Method for Estimating Lipophilicity and Ionization in Nonpolar Membrane-Like Environment. Mol. Pharmaceutics, 2016;13:1100–10.

[14] Ermondi G, Visconti A, Esposito R, Caron G. The Block Relevance (BR) analysis supports the dominating effect of solutes hydrogen bond acidity on ΔlogP(oct-tol). Eur J Pharm Sci 2014;53:50–4.

[15] Kuhn B, Mohr P, Stahl M. Intramolecular hydrogen bonding in medicinal chemistry. J Med Chem 2010;53:2601–11.

[16] Shalaeva M, Caron G, Abramov YA, Connell TNO, Plummer MS, Yalamanchi G, et al. Integrating Intramolecular Hydrogen Bonding (IMHB) Considerations in Drug Discovery Using Δ logP As a Tool. J. Med. Chem. 2013;56:4870−9.

[17] Goetz GH, Farrell W, Shalaeva M, Sciabola S, Anderson D, Yan J, et al. High throughput method for the indirect detection of intramolecular hydrogen bonding. J Med Chem 2014;57:2920–9.

[18] Lombardo F, Shalaeva MY, Tupper KA, Gao F. ElogD oct : A Tool for Lipophilicity Determination in Drug Discovery. 2. Basic and Neutral Compounds J. Med. Chem. 2001;44:2490–7.

[19] Mannhold M, Poda GI, Ostermann C, Tetko I V. Calculation of molecular lipophilicity: State of the art and comparison of methods on more than 96000 compounds. JPharmSci 2009;98:861–93.

[20] Bockus AT, Lexa KW, Pye CR, Kalgutkar AS, Gardner JW, Hund KCR, et al. Probing the Physicochemical Boundaries of Cell Permeability and Oral Bioavailability in Lipophilic Macrocycles Inspired by Natural Products. J Med Chem 2015;58:4581–9.

[21] Bergström CA, Strafford M, Lazorova L, Avdeef A, Luthman K, Artursson P. Absorption Classification of Oral Drugs Based on Molecular Surface Properties. J Med Chem 2003;46:558–70.