Research Article

Experimental lipophilicity for beyond Rule of 5 compounds

Giuseppe Ermondi[1], Maura Vallaro[1], Gilles Goetz[2], Marina Shalaeva[3] & Giulia Caron*[,1]

[1]Molecular Biotechnology & Health Sciences Department, Università degli Studi di Torino, via Quarello 15, Torino 10135, Italy
[2]Hit Discovery & Optimization, Discovery Sciences, WWRD, Pfizer Inc, Eastern Point Road, Groton, CT 06340, USA
[3]Independent Scientist, Dana Point, CA 92629, USA
*Author for correspondence: Tel.: +39 011 670 8337; giulia.caron@unito.it

Aim: To set up a chromatographic strategy for the determination of log P for beyond Rule of 5 (bRo5) drugs. **Materials & methods:** Capacity factors measured by reverse phase-HPLC. Balance of intermolecular interactions governing systems assessed by partial least squares regression (PLSR) coupled with block relevance anaysis (PLSR-BR) and multiblock PLSR (MBPLSR). Determination of virtual log P obtained through conformational sampling. **Results:** log k'60 is highly correlated with log P for a dataset of 36 Ro5 compliant compounds ($R^2 = 0.93$, $Q^2 = 0.90$). We refer to the value generated via this method as BRlogP. The balance of intermolecular forces controlling BRlogP and log P are very similar. The ElogPs measured for the bRo5 dataset are significantly higher than corresponding BRlogP. **Conclusion:** The combination of BRlogP and ElogP provides an experimental lipophilicity range for bRo5 compounds.

The major role played by lipophilicity in drug discovery is widely known [1,2] and recently reviewed by Goetz and Shalaeva [3], who illustrated the potential of lipophilicity in balancing potency and absorption, distribution, metabolism, and excretion (ADME) properties from the point of view of the pharmaceutical industry. Lipophilicity is commonly described as log D, where the distribution coefficient D is quantified by the concentration of all species (unionized and ionized) of a compound at a given pH in two immiscible phases (commonly 1-octanol and water/buffer) at equilibrium. Distribution coefficient (D) is replaced with partition coefficient (P) at any given pH if only one species (typically neutral) is present.

HPLC is a widely used analytical technique that characterizes the physicochemical properties of compounds [4]. Chemicals are retained proportionally to their partition coefficient in the octanol/water system under the reverse phase-HPLC (RP-HPLC) conditions on analytical columns packed with a solid phase made of long hydrocarbon chains (e.g., C8 and C18) chemically bound onto silica. The RP-HPLC methods are therefore a reliable alternative to other experimental log D/log P measurements (e.g., shake-flask) [5].

A number of research organizations have developed their own RP-HPLC methods for lipophilicity determinations [3] and ElogP (ElogD) is the tool implemented, along with shake-flask, for Pfizer research programs [6,7]. Shortly, this method uses an octanol saturated mobile phase with methanol as organic modifier and an LC-ABZ column as a stationary phase. The obtained log K_w data were calibrated with known log P (log D) values of a training set and the obtained log P (log D) data are reported as ElogP (ElogD). The solvation analysis was performed to verify the closeness in terms of balance of intermolecular interactions between the classical shake flask log P (log D) values and ElogP (ElogD) [6,7].

The first aim of the present study is to test if a fast HPLC method based on a single log k' value would yield reliable log P values for neutral drug-like molecules. To do this, we tested two well-known columns (Supelcosil LC-ABZ and XBridge™ Shield RP18, see methods) and mobile phases using acetonitrile as organic modifier. The method was applied to a dataset of 36 known neutral Ro5 compliant compounds. It was validated for the balance of intermolecular forces with two computational tools; partial least squares regression (PLSR) coupled with block

relevance (BR) analysis and multiblock PLSR (MBPLSR). We refer to the value generated via this method as BRlogP.

Recent investigations have highlighted that compounds residing just outside of traditional small molecule drug space, in other words in the beyond rule of 5 (bRo5) space, allow modulation of difficult classes of targets [8]. The bRo5 drugs have a huge relevance in immunosuppression, treatment of infectious and viral diseases and in oncology [9]. Reliable determination of physicochemical properties is required [10] for a successful design of bRo5 drugs, lipophilicity measurement being one of the most important parameters [3,9].

Lipophilicity accounts for molecular size and polarity of compounds in solution as well as for 3D conformational changes [11] adopted by flexible molecules, which in turn depend on the polarity of the environment [12,13]. Therefore, it has to be demonstrated that methods optimized to measure log P for Ro5 molecules can be transferred and successfully applied to bRo5 compounds since different systems (either biphasic or RP-HPLC), characterized by different environments, can affect resulting lipophilicity values. Therefore, the second and main aim of the study is to check the performances of BRlogP to assess the lipophilicity of bRo5 compounds. Since ElogP is one of the most known lipophilicity indices implemented in drug discovery, it was selected as the method to compare against. We selected nine bRo5 drugs and measured BRlogP and ElogP for in depth analyses. To investigate the role played by flexibility, we also submitted the bRo5 drugs to a standard conformational sampling protocol and calculated the lipophilicity of the conformers (i.e., virtual log P [14]). Experimental and computational lipophilicity were finally compared.

Materials & methods

BRlogP

Deionized water and HPLC grade acetonitrile were used throughout. The mobile phases consisted of 20 mM ammonium/acetate at pH 7.0 and acetonitrile from 60 to 80% v/v. Samples were dissolved in a 1:1 mixture of buffer/acetonitrile in the concentration range of 50–100 ug/ml. The flow rate was 1 ml/min. The dead time, t_0, was determined by monitoring the baseline disturbance. The retention time (t_R) was measured on a Supelcosil LC-ABZ column (Supelco, 5 cm × 4.6 mm, 5 μm particle size, 120Å pore size) and XBridge™ Shield RP18 (Waters, 5 cm × 4.6 mm, 5 μm particle size, 130Å pore size). Isocratic log k′ (capacity factor k′ = [t_R - t_0]/t_0) values were measured.

ElogP

Deionized water and HPLC grade methanol were used throughout. The mobile phases consisted of 20 mM MOPS buffer with 0.15% n-decylamine in octanol-1 saturated water at pH 7.4 and 0.25% octanol-1 in methanol at 60, 65 and 70% v/v. Samples were dissolved in methanol/water in the concentration range of 50–200 ug/ml. The flow rate was 2 ml/min. Injections of pure methanol were used to determine the dead time t_0. The retention time (t_R) was measured on a Supelcosil LC-ABZ column. Isocratic log k′ (capacity factor k′ = [t_R - t_0]/t_0) values were measured. The ELogP was determined by the published equation $logP_{oct} = 1.1267* log k'w + 0.2075$, where log k′w is the log k′ extrapolated to 0% methanol, using a linear procedure [7].

PLSR & BR analysis

The PLSR models were built according to the following strategy. First, compounds were split in a training and a test set and then the simplified molecular-input line-entry system codes of the compounds were submitted to VolSurf+ (VS+, version 1.0.7, www.moldiscovery.com) using default settings and four probes (OH2, DRY, N1 and O probes that mimic, respectively, water, hydrophobic, hydrogen bond acceptor [HBA] and hydrogen bond donor [HBD] properties of the environment). PCA and PLSR tools implemented in VS+ were used to obtain and validate models.

The BR analysis was performed as described elsewhere [15–17]. Briefly, BR analysis organizes the VS+ descriptors into six blocks (Size, Water, DRY, N1, O and Others) of straightforward significance detailed in Figure 1, panel B. In a very simplistic way, BR analysis gives the relevance of each parameter (block) to the model.

Multiblock partial least squares regression

The multiblock PLSR as described by Westerhuis [18] was carried out using MVAPACK, implemented in Octave. Briefly, MBPLSR generates a model of the response Y (log P and log k′60) as a linear function of the six blocks

$$y = 3.31x + 2.79$$
$$R^2 = 0.9281$$
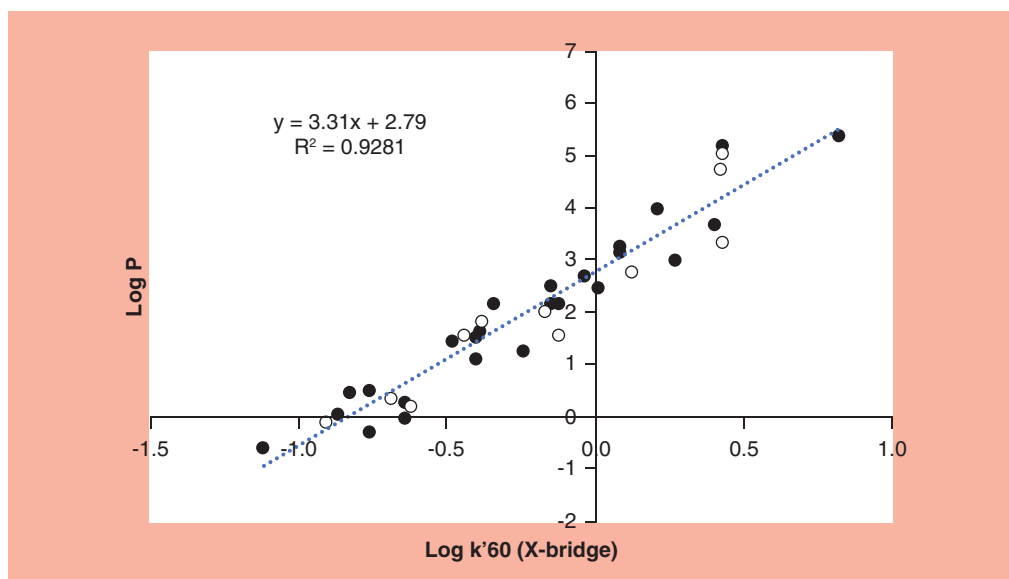
Figure 1.   Relationship (Eq. 1) between log P and log k'60 measured with the X-Bridge column (training set: full dots; test set: empty dots).

based on the VS+ descriptors (see above for blocks definition). To do that, descriptors (X) are first structured as reported in the literature and then submitted to the modified PLSR algorithm [18].

The model with a number of latent variables that maximizes the leave-many-out cross validation coefficient (Q^2_{LMO}) is chosen as the best model. Finally, the relative importance of each block is evaluated using the super weights *ws* vector [19].

Molecular descriptors
All molecular descriptors were calculated with Dragon (software for molecular descriptor calculation) version 7.0.10, 2017, Kode srl, https://chm.kode-solutions.net, kindly provided by Todeschini.

Conformational sampling & virtual log P calculation
The simplified molecular-input line-entry system (SMILES) codes of the ten bRo5 drugs were obtained from the DrugBank database (www.drugbank.ca/) and where used directly as input for OMEGA (OMEGA 3.0.1.2: OpenEye Scientific Software, NM, USA. www.eyesopen.com, 2018). Default settings were applied. Virtual log P, in other words the log P of any conformer, was calculated by projecting PLSR models obtained with VS+ (see above) on the conformers generated by Omega.

Results & discussion
BRlogP
To set up a method to measure log P by RP-HPLC, the same dataset of 36 compounds used by Lombardo and coworkers [6] was selected. It includes neutral, Ro5 compliant compounds, covering a log P range from -0.55 to 5.4. Three major points were evaluated: the stationary and mobile phases to be used, the chromatographic index to be determined and the tools to be applied to validate the method, in other words how to compare the balance of intermolecular forces governing the retention mechanism in the RP-HPLC system and the partitioning mechanism in the reference system 1-octanol/water.

Among stationary phases, LC-ABZ provided excellent results and high correlations were found between log P and chromatographic indexes for a wide range of structurally diverse compounds [6,20]. In this stationary phase, the alkyl chains contain amide groups that electrostatically shield silanols from highly polar analytes. However, the application of this type of stationary phases is still limited to the pH below eight because of instability in alkaline media. A more recent RP-HPLC column (XBridge™ Shield RP18, from here on named XBridge) affords a wider pH resistance (1–12) and bears the same alkyl chains as those in the LC-ABZ [21]. Overall both LC-ABZ and X-Bridge deserve being considered as valuable stationary phases for the determination of a lipophilicity index.

Acetonitrile was chosen since it was suggested to be the best organic modifier [19] and it is already applied in other chromatographic systems recently developed in our laboratories (i.e., log k′80 PLRP-S) [22].

Two chromatographic indexes can be determined in principle: k′X% represents the capacity factor of the solute at an X% organic solvent concentration, and K_w is the capacity factor extrapolated to a 0% concentration of the organic solvent. Many studies tried to establish which descriptor gives the best correlation with lipophilicity, but their review is beyond the scope of the paper. Here we decided to focus on log k′ since it is faster to obtain than log K_w. Some years ago, we showed that for Lombardo's dataset (n = 36) the balance of intermolecular forces governing log k′60 determined on a LC-ABZ column using methanol as cosolvent was closer to log P than that expressed by log K_w [16]. Therefore, in this study, we first focused on log k′60. However, since log k′80 is faster to obtain (shorter retention times) and already implemented to determine lipophilicity in a chromatographic apolar system [22], for the sake of comparison we also considered log k′80. Summing up, in this study we determined four chromatographic indexes: log k′60 LC-ABZ, log k′80 LC-ABZ, log k′60 X-Bridge and log k′80 X-Bridge (Table 1).

Data analysis first revealed that the performances produced by the two columns are largely superimposable (Supplementary Figure 1). Moreover, in agreement with a previous study [16], we verified that the relationship between log P and log k′60 is stronger than the corresponding with log k′80 (Supplementary Figure 2). Therefore, to set up a new lipophilicity index, we focused on log k′60 measured with the X-Bridge column which endowed better pH stability than LC-ABZ.

The dataset was randomly split in training (24) and test (12) sets defined in Table 1. The relationship between log P and log k′60 for the training set (the dataset) is reported in Figure 1 and the corresponding equation is below (Eq. 1)

$$\log P = 3.31 * \log k'60 + 2.79$$
$$R^2 = 0.93; Q^2_{LMO} = 0.91; R^2_{Yscr} = 0.042$$

(Eq. 1)

where R^2 is coefficient of determination, Q^2_{LMO} is leave-many-out cross validation coefficient and R^2_{Yscr} is the coefficient of determination of the corresponding scrambled models

In practice, once measured log k′60 with the X-Bridge column, we can calculate log P via Eq. 1. We termed this log P as BRlogP since we used BR analysis for its mechanistic interpretation (see below).

Some of us recently published a study focusing on methods enabling the deconvolution of intermolecular forces governing physicochemical descriptors [18]. The BR analysis is one of them and it was already successfully applied to model chromatographic indexes [15,16,23]. Shortly, BR analysis provides a mechanistic 'at a glance' of QSPR PLSR-based models using a 'post processing' supervised grouping of VS+ descriptors in blocks with a clear physicochemical meaning. To avoid a postproduction process of the PLSR results, a modified PLSR algorithm implementing blocks (MBPLSR) was also recently applied to interpret the logarithm of the retention coefficients determined in an SDS MLC system (log K_w^{SDS}) [18]. Both computational approaches were applied in this study as described below.

Log k′60 values were imported into VS+ as response variables (Y) and a relation between Y and the 82 VS+ molecular descriptors (X) was sought using the PLSR algorithm implemented in the software (see methods). Statistically stable PLSR models were found (Table 2) and shown by the determination coefficients of the test set (R^2[test] in Table 2).

Graphical outputs of BR analysis are reported in Figure 2 along with the meaning and significance of the blocks. Panel A and B (Figure 2) show the balance of intermolecular forces governing log P and log k′60 (Table 1). At first, glance and without entering in the very details of BR plots interpretation, Figure 2A and B are almost superimposable; the only slight difference being the size block (green) slightly smaller for the chromatographic descriptor. This finding demonstrated that log P and BRlogP (calculated from log k′60) encode the same information.

The same data matrix were also submitted to MBPLSR and panel C and D in Figure 2 show the graphical results. The same trend observed for PLSR-BR analysis is found for MBPLSR.

Lipophilicity of bRo5 drugs

As mentioned above the determination of lipophilicity for bRo5 compounds is a major need in modern drug discovery. Since ElogP is a widely applied chromatographic descriptor of lipophilicity of industrial relevance, we measured BRlogP and ElogP for nine neutral bRo5 drugs (Table 3, chemical structures in Supplementary Figure 3).

## Table 1. Chromatographic indexes discussed in the paper.

Compound	tr/ts	Log k'60 X-Bridge	Log k'80 X-Bridge	Log k'60 LC-ABZ	Log k'80 LC-ABZ	Log P
3,5-dichlorophenol	Training	0.4	-0.03	0.45	0.02	3.68
3-bromoquinoline	Training	0.27	-0.06	0.22	-0.12	3.03
3-chlorophenol	Training	0.01	-0.27	0.07	-0.29	2.50
Acetaminophen	Training	-0.76	-0.8	-0.68	-0.8	0.51
Acetophenone	Test	-0.12	-0.36	-0.17	-0.48	1.58
Allopurinol	Training	-1.12	-1.12	-0.95	-0.98	-0.55
Antipyrine	Test	-0.69	-0.72	-0.57	-0.64	0.38
Bifonazole	Test	0.42	-0.11	0.39	-0.16	4.77
Bromazepam	Test	-0.39	-0.52	-0.24	-0.51	1.65
Caffeine	Training	-0.91	-0.82	-0.72	-0.72	-0.07
Carbamazepine	Test	-0.34	-0.55	-0.27	-0.59	2.19
Chloramphenicol	Training	-0.4	-0.67	-0.41	-0.67	1.14
Clotrimazole	Training	0.43	-0.08	0.37	-0.13	5.20
Dexamethasone	Training	-0.38	-0.55	-0.28	-0.54	1.83
Diazepam	Test	0.12	-0.24	0.05	-0.31	2.79
Diethylstilbestrol	Test	0.43	-0.08	0.39	-0.14	5.07
Estradiol	Test	0.21	-0.13	0.33	-0.03	4.01
Fluconazole	Training	-0.83	-0.88	-0.66	-0.75	0.50
Griseofulvin	Training	-0.15	-0.43	-0.15	-0.54	2.18
Hydrocortisone	Training	-0.4	-0.55	-0.36	-0.54	1.55
Hydrocortisone-21-acetate	Training	-0.12	-0.42	-0.12	-0.48	2.19
Lorazepam	Training	-0.15	-0.46	-0.15	-0.44	2.51
Lormetazepam	Training	-0.04	-0.35	0.01	-0.42	2.72
Methylthioinosine	Training	-0.87	-0.95	-0.88	-0.91	0.09
Metronidazole	Training	-0.64	-0.76	-0.64	-0.7	-0.02
Naphtalene	Training	0.43	0.01	0.38	-0.1	3.37
Nifedipine	Test	0.08	-0.34	-0.07	-0.43	3.17
Nifuroxime	Training	-0.24	-0.52	-0.27	-0.59	1.28
Nitrofurazone	Training	-0.62	-0.8	-0.62	-0.8	0.23
Pentoxifylline	Test	-0.64	-0.67	-0.64	-0.79	0.29
Prednisolone	Training	-0.44	-0.56	-0.37	-0.57	1.60
Prednisone	Test	-0.48	-0.63	-0.45	-0.66	1.46
Quinoline	Training	-0.17	-0.36	-0.13	-0.36	2.03
Testosterone	Test	0.08	-0.22	0.12	-0.12	3.29
Thiamphenicol	Training	-0.76	-0.81	-0.71	-0.82	-0.27
Tolnaftate	Training	0.82	0.15	0.68	-0.02	5.40

Compounds belonging either to the training or to the test set are indicated. Log P data were taken from the paper by Lombardo *et al.* [6].

Table 2. Final models obtained using partial least squares regression and multiblock partial least squares regression.

Algorithm	Descriptor	N	N training	N test	R²	Q²LMO	LVs	Slope	Intercept	R²(Test)
PLSR	Log k'60	36	24	12	0.86	0.58	3	1.16	0.00	0.74
PLSR	Log P	36	24	12	0.93	0.57	3	1.06	-0.36	0.70
MBPLSR	Log k'60	36	24	12	0.86	0.51	3	1.10	0.06	0.50
MBPLSR	Log P	36	24	12	0.90	0.62	3	1.05	-0.11	0.56

N = total number of compounds, N training = training set compounds, N test = test set compounds, R2 = coefficient of determination, Q2LMO = cross-validation coefficient leave-many-out, in which about 30% of compounds are randomly excluded for prediction for 2000 iterations, LVs = number of latent variables, slope = slope of the linear regression obtained by plotting experimental vs predicted data of the training set, intercept = intercept of the linear regression obtained by plotting experimental vs predicted data of the training set, R2 (test) = coefficient of determination for the test set.
LV: Latent variable; MBPLSR: Multiblock PLSR; PLSR: Partial least squares regression.
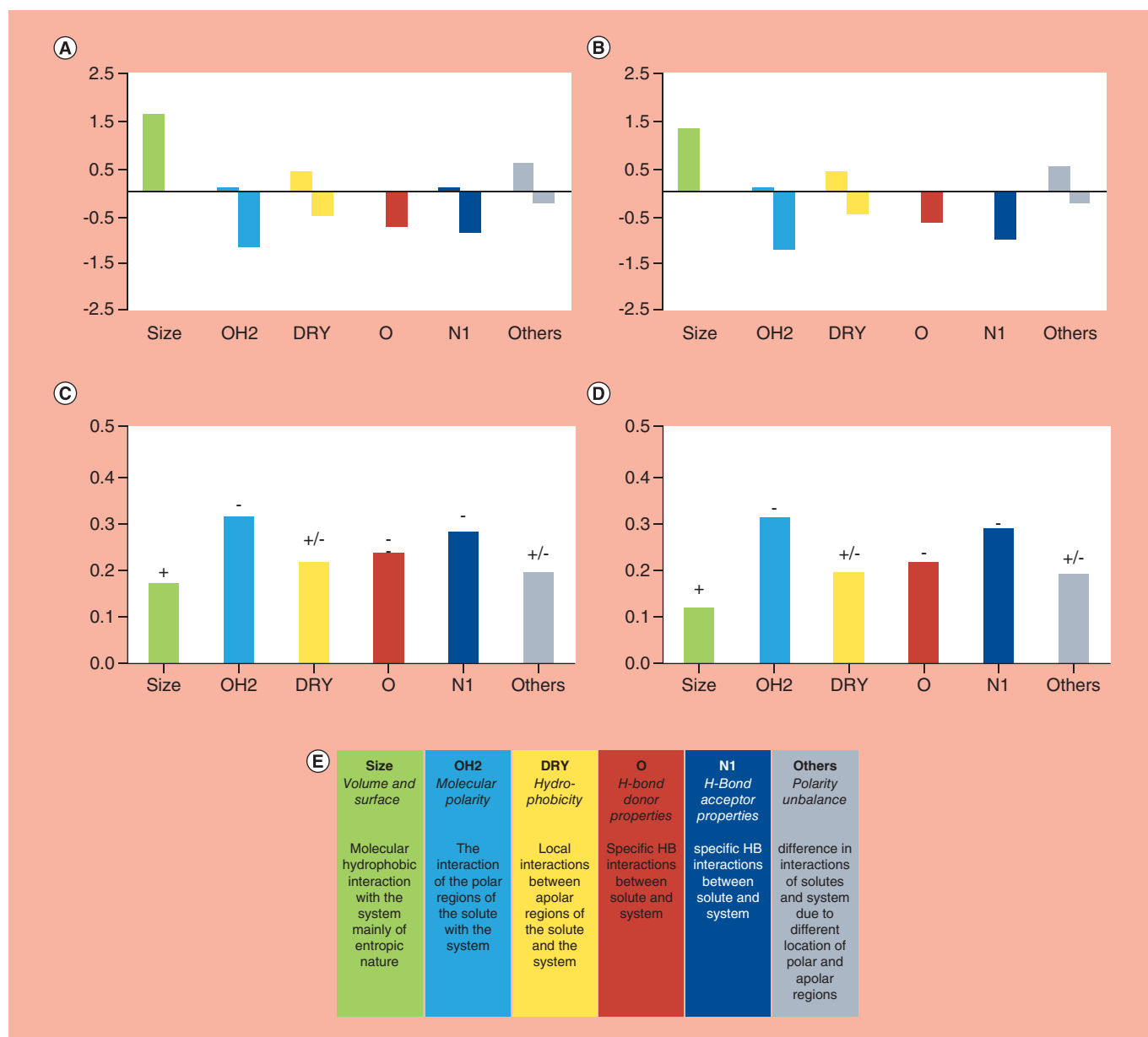
Figure 2.   Mechanistic interpretation of log P and log k'60. (A) log P (block relevance analysis), **(B)** log k'60 block relevance analysis), **(C)** log P (multiblock PLSR), **(D)** log k'60 (multiblock PLSR) and **(E)** meaning and significance of the blocks.

Moreover, in Table 3 we also report AlogP, a calculated value of lipophilicity, recently included in the determination of the lipophilicity permeability efficiency to replace experimental data [24]. Shortly, AlogP is a fragment-based 2D octanol/water partition coefficient calculator [25] which is expected to provide good accuracy for large molecules log P calculations. The relationship between BRlogP/ElogP and AlogP for the nine bRo5 drugs is poor (R^2 about 0.5, Supplementary Figure 4) and confirms that caution should be used in replacing an experimental lipophilicity value with a calculated one in the bRo5 chemical space.

Figure 3 shows the relationship between BRlogP and ElogP. For the traditional Ro5 compliant compounds, the two methods show an excellent correlation (R^2 = 0.93, slope about 1 and Y-intercept about 0). For bRo5 drugs, although BRlogP and ElogP are still highly correlated (R^2 = 0.96), the slope is higher than one (1.24) and the Y-intercept higher than 0 (0.67). This finding shows that the two chromatographic systems provide the same log P for small compounds but not for bRo5 drugs. This is particularly evident for lipophilic compounds.

Table 3. bRo5 compounds: lipophilicity data.				
	BRlogP	ElogP	AlogP	Exp log P range[†]
Atazanavir	3.0	4.7	5.3	1.68
Cyclosporine	5.8	8.5	5.6	2.69
Everolimus	4.9	6.7	6.0	1.84
Indinavir	1.6	3.8	3.2	2.24
Paclitaxel	3.2	4.4	3.1	1.24
Ritonavir	3.2	4.9	5.5	1.74
Tacrolimus	4.8	6.1	4.7	1.26
Telaprevir	3.5	4.4	3.4	0.94
Temsirolimus	4.8	6.9	5.8	2.06

[†]ElogP-BRlogP.
BRlogP and ElogP are experimental values whereas AlogP is computed using the Dragon software.
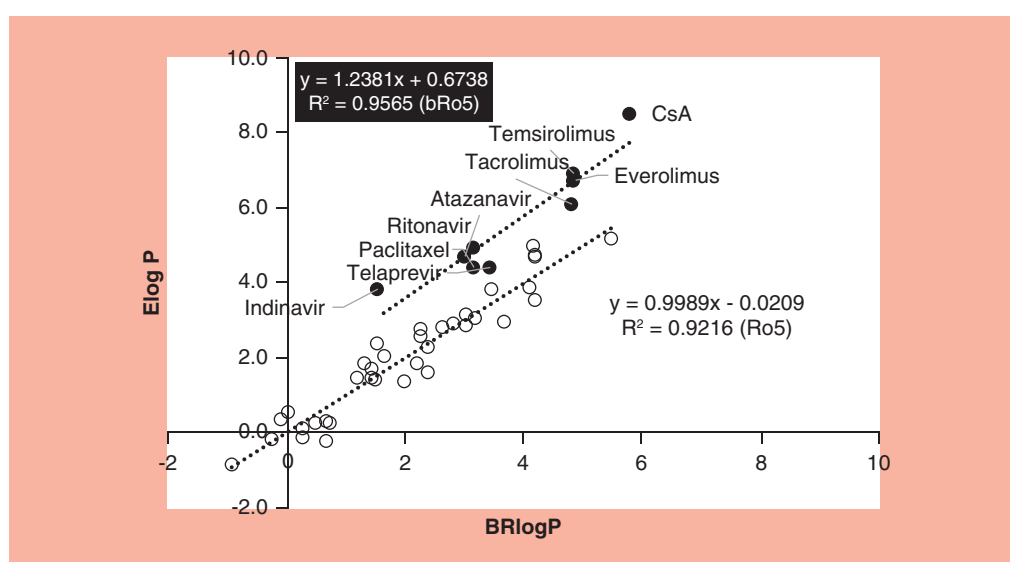
Figure 3.   Relationship between BRlogP and ElogP for the Ro5 compliant compounds (empty dots) and the bRo5 drugs (full dots).

The BRlogP and ElogP experimental conditions differ in the type of column (LC-ABZ vs X-Bridge), pH (7.0 vs 7.4), mobile phase composition (acetonitrile vs methanol and octanol) and nature of the chromatographic index (log k′ vs log K_w). In the first part of the study, we verified that LC-ABZ and X-Bridge provided very similar retention values and consequently the column type is not expected to impact retention. Since we are examining only neutral compounds here, the differences in pH and buffer composition are expected to be not relevant. Therefore, the major differences between the two methods are the mobile phase composition and the extrapolation process required to obtain ElogP. We can therefore conclude that BRlogP and ElogP provide two diverse environments with different polarity. This could impact the conformers' population and thus log P value of bRo5 compounds. Data suggest that ElogP seems to provide an environment in which compounds are forced to assume a more folded and lipophilic conformation than under BRlogP conditions.

To investigate this finding further, we performed conformational sampling on the bRo5 compounds and calculated log P for any conformer (i.e., virtual log P) using the second PLSR model (log P) reported in Table 1 (more details about the calculation of virtual log P and its relevance in drug discovery can be retrieved in the literature [26,27]). The computed distribution of lipophilicity among conformers is displayed as a box and whisker plot in Figure 4. Moreover, two conformers with low and high log P (data in Supplementary Table 1) are shown as examples for indinavir and CsA (Figure 4). Experimental BRlogP and ElogP are also indicated in red and black, respectively. The computed AlogP values are in green in Figure 4.
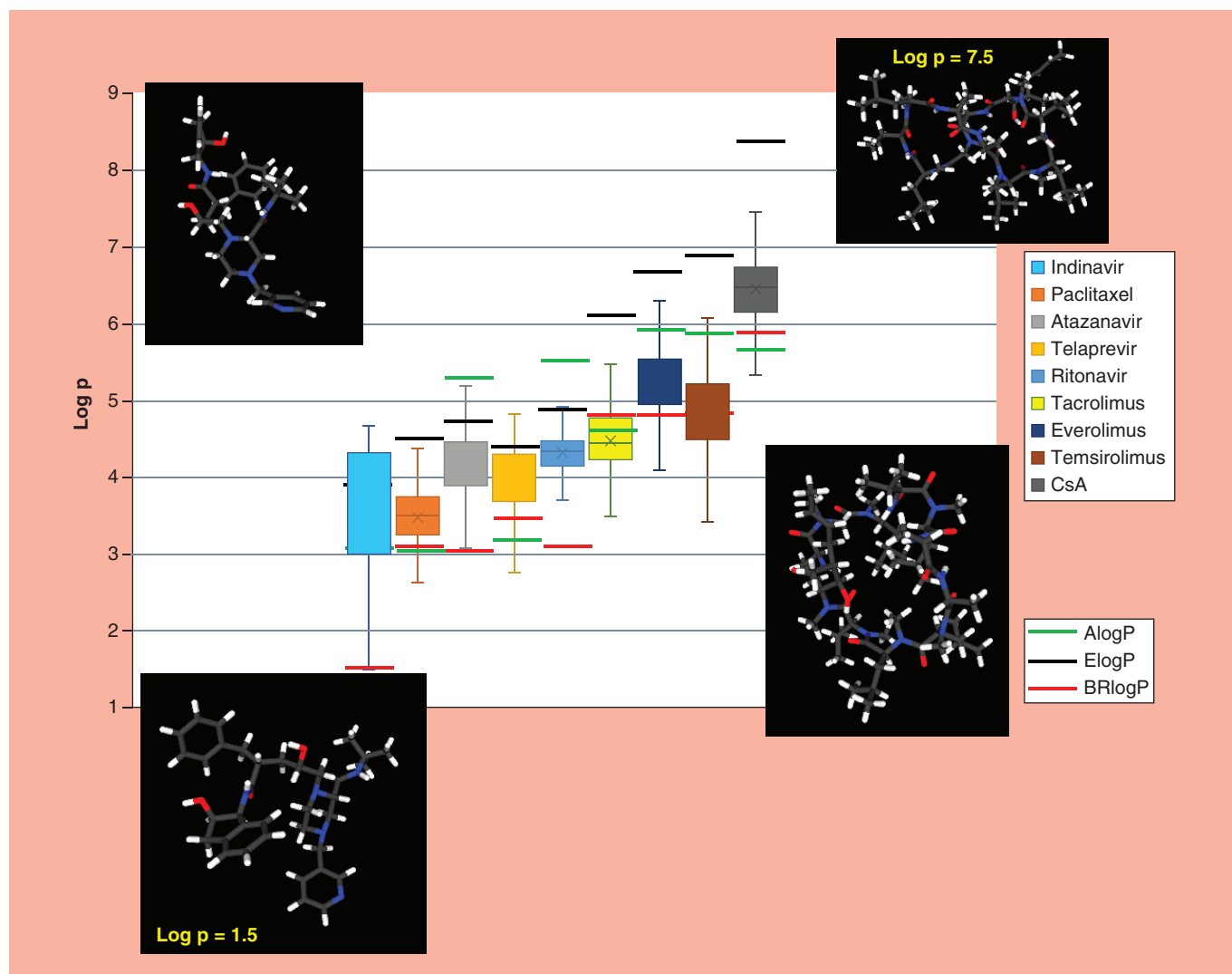
Figure 4.   The box and whiskers plot shows the minimum and maximum virtual log P values as whiskers; the boxes span the 25–75th percentile range. AlogP (green), BRlogP (red) and ElogP (black) values are also shown. Two conformers with low and high log P are shown as examples for indinavir and CsA.

Figure 4 shows that BRlogP (red bar) is in the 0–25th percentile of virtual log P except for ritonavir, tacrolimus and temsirolimus. The ElogP (black bar) is often much higher than any virtual log P values. The AlogP (green bar) varies with the considered compound and no clear trend can be found. Sometimes it is found within the 25–75th percentiles (c.f. indinavir and tacrolimus), but often it is located either in the 0–25th or in 75–100th percentile. The combination of BRlogP and ElogP provides an experimental range of log P (= ElogP - BRlogP in Table 3) which spans from about one for telaprevir to about 2.7 for CsA. Notably, the experimental log P range is larger for macrocyclic compounds (CsA, everolimus, paclitaxel, tacrolimus and temsirolimus) than for nonmacrocyclic drugs (atazanavir, indinavir, ritonavir and telaprevir). Finally, the experimental log P range is not superimposable with the virtual log P range. A lack of superposition between experimental and theoretical ranges is also found when using a different method to calculate virtual log P (i.e., log P from MLP calculated with Vega software v. 3.1.1.42, data in Supplementary Table 1). Work along these lines is in due course in our laboratories.

Overall Figure 4 supports that a single log P value often poorly represents the lipophilicity of conformationally flexible bRo5 compounds, which is better described by an experimental lipophilicity range. In addition, the lipophilicity range could be used to quantify the chameleonic properties of compounds in the bRo5 chemical space,

where a chameleon is a molecule for which the closed form is preferred in apolar environments and the open form in aqueous media [28]. This is a topic of huge relevance in modern drug discovery [29].

Conclusion

We have shown that BRlogP is a valuable and fast tool to experimentally access lipophilicity of neutral Ro5 compliant compounds. Its combination with ElogP provides an experimental log P range that could be implemented in bRo5 drug discovery programs. The possibility of using different HPLC methods to measure 'conformer-dependent' experimental log P values is highlighted. Work is in progress to find a similar method to determine log D at physiological pHs where most drugs are partially of fully ionized.

Future perspective

As recently reviewed in the literature [9,30], there is an increasing focus in the bRo5 chemical space. Notably in the past 3 years, 21% of new oral drugs (12) do not comply with Lipinski's Ro5. However, the understanding of the rules that govern a successful design of oral bRo5 drugs is far from being achieved. A major issue concerns drug metabolism and pharmacokinetic properties, since most bRo5 candidates show poor permeability, low solubility and high metabolic clearance. Lipophilicity is a molecular property largely impacting the ADME profile of compounds, and therefore there is the need to obtain lipophilicity descriptors for large and flexible structures. Since molecular properties such as log P of bRo5 compounds strongly depend on their chameleonic properties (a chameleon is a molecule for which the closed form is preferred in apolar environments and the open form in aqueous media), there is no guarantee that experimental and computational protocols developed on Ro5 compliant structures properly work on bRo5 compounds. In fact, the characteristics of the environment were generally poorly considered during the development of methods in the Ro5 era. Therefore, we expect that drug discovery campaigns in the bRo5 chemical space in the next years will implement specific methodological solutions for the determination of *ad hoc* physicochemical descriptors. The experimental lipophilicity range determined with ElogP and BRlogP is in line with this scenario.

Summary points

- A fast chromatographic index to assess log P (BRlogP) based on the determination of a single reverse phase high-performance liquid chromatography retention time is described and validated for Ro5 compliant compounds.
- The applicability domain of BRlogP to bRo5 compounds is verified in relation with ElogP.
- The limits of simple (AlogP) and more complex (virtual log P) computed lipophilicity descriptors are highlighted.
- The combination of BRlogP and ElogP provides an experimental log P range, which could quantify the chameleonic properties of bRo5 drugs.

References

Papers of special note have been highlighted as: ● of interest; ●● of considerable interest

1.   Valko K. *Physicochemical and Biomimetic Properties in Drug Discovery*. John Wiley & Sons, NJ, USA (2014).

2.   Arnott JA, Planey SL. The influence of lipophilicity in drug discovery and design. *Expert Opin. Drug Discov.* 7(10), 863–875 (2012).

3.   Goetz GH, Shalaeva M. Leveraging chromatography based physicochemical properties for efficient drug design. *ADMET and DMPK* 6(2), 85–104 (2018).

4.   Valkó K. Application of high-performance liquid chromatography based measurements of lipophilicity to model biological distribution. *J. Chromatogr. A.* 1037(1–2), 299–310 (2004).

5.   OECD. Test No. 117: Partition Coefficient (n-octanol/water), HPLC Method. *OECD Guidelines for the Testing of Chemicals, Section 1, OECD Publishing, Paris* (2004). https://dx.doi.org/10.1787/9789264069824-en.

6.   Lombardo F, Shalaeva MY, Tupper KA, Gao F, Abraham MH. ElogPoct: a tool for lipophilicity determination in drug discovery. *J. Med. Chem.* 43(15), 2922–2928 (2000).

●   **Includes the full description of the ElogP method and its validation.**

7.   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.* 44(15), 2490–2497 (2001).

8.   Doak BC, Kihlberg J. Drug discovery beyond the rule of 5 – opportunities and challenges. *Expert Opin. Drug Discov.* 12(2), 1–5 (2016).

9.   Degoey DA, Chen HJ, Cox PB, Wendt MD. Beyond the rule of 5: lessons learned from AbbVie's drugs and compound collection. *J. Med. Chem.* 61(7), 2636–2651 (2018).

10. Caron G, Ermondi G. Updating molecular properties during early drug discovery. *Drug Discov. Today* 22(6), 835–840 (2017).

●●   **The need of *ad hoc* descriptors for bRo5 chemical space is first described here.**

11. Guimarães CRW, Mathiowetz AM, Shalaeva M, Goetz G, Liras S. Use of 3D properties to characterize beyond rule-of-5 property space for passive permeation. *J. Chem. Inf. Model.* 52(4), 882–889 (2012).

12. Rossi Sebastiano M, Doak BC, Backlund M *et al.* Impact of dynamically exposed polarity on permeability and solubility of chameleonic drugs beyond the rule of 5. *J. Med. Chem.* 61(9), 4189–4202 (2018).

13. Poongavanam V, Danelius E, Peintner S *et al.* Conformational sampling of macrocyclic drugs in different environments – Can we find the relevant conformations? *ASC Omega* 3(9), 11742–11757 (2018).

14. Carrupt P-A, Testa B, Gaillard P. Computational approaches to lipophilicity: methods and applications. In: *Reviews in Computational Chemistry (Volume 11)*. Lipkowitz KB, Boyd DB (Eds), Wiley, NY, USA, (1997).

15. Ermondi G, Caron G. Molecular interaction fields based descriptors to interpret and compare chromatographic indexes. *J. Chromatogr. A.* 1252, 84–89 (2012).

16. Caron G, Vallaro M, Ermondi G. The Block Relevance (BR) analysis to aid medicinal chemists to determine and interpret lipophilicity. *Med. Chem. Commun.* 4(10), 1376–1381 (2013).

17. Caron G, Ermondi G. Block relevance (BR) analysis and polarity descriptors in property-based drug design. *ADMET and DMPK* 6, 215–224 (2018).

18. Ermondi G, Caron G. MLR, PLSR-BR analysis and MBPLSR to interpret multivariate QSPR models. The case of a micellar liquid chromatography descriptor (log K_W^{SDS}). *Mol. Inf.* 38, 1800144 (2019)

19. Pallicer JM, Pous-torres S, Sales J, Rosés M, Ràfols C, Bosch E. Determination of the hydrophobicity of organic compounds measured as log P o/w through a new chromatographic method. 1217, 3026–3037 (2010).

20. Poole SK, Poole CF. Separation methods for estimating octanol–water partition coefficients. *J. Chromatogr. B.* 797(1–2), 3–19 (2003).

21. Liu X, Hefesha H, Tanaka H, Scriba G, Fahr A. Lipophilicity measurement of drugs by reversed phase HPLC over wide pH range using an alkaline-resistant silica-based stationary phase, XBridgeTM shield RP18. *Chem. Pharm. Bull.* 56(10), 1417–1422 (2008).

22. Caron G, Vallaro M, Ermondi G *et al.* A fast chromatographic method for estimating lipophilicity and ionization in nonpolar membrane-like environment. *Mol. Pharm.* 13(3), 1100–1110 (2016).

23. Ermondi G, Vallaro M, Caron G. Learning how to use IAM chromatography for predicting permeability. *Eur. J. Pharm. Sci.* 114, 385–390 (2018).

24. Naylor MR, Ly AM, Handford MJ *et al.* Lipophilic Permeability Efficiency (LPE) reconciles the opposing roles of lipophilicity in membrane permeability and aqueous solubility. *J. Med. Chem.* doi: 10.1021/acs.jmedchem.8b01259 (2018) (Epub ahead of print).

25. Ghose AK, Crippen GM. Atomic physicochemical parameters for 3- dimensional structure-directed quantitative structure-activity-relationships 1. Partitioncoefficients as a measure of hydrophobicity. *J. Comput. Chem.* 7(4), 565–577 (1986).

26. Caron G, Ermondi G. Calculating virtual log P in the alkane/water system (log P^N_{alk}) and its derived parameters Δlog $P^N_{oct-alk}$ and log D^{pH}_{alk}. *J. Med. Chem.* 48(9), 3269–3279 (2005).

27. Vistoli G, Pedretti A, Testa B. Partition coefficient and molecular flexibility: the concept of lipophilicity space. *Chem. Biodivers.* 6(8), 1152–1169 (2009).

28. Carrupt PA, Testa B, Bechalany A, El Tayar N, Descas P, Perrissoud D. Morphine 6-glucuronide and morphine 3-glucuronide as molecular chameleons with unexpected lipophilicity. *J. Med. Chem.* 34(4), 1272–1275 (1991).

29. Whitty A, Zhong M, Viarengo L, Beglov D, Hall DR, Vajda S. Quantifying the chameleonic properties of macrocycles and other high-molecular-weight drugs. *Drug Discov. Today* 21(5), 712–717 (2016).

•• **Highlights that chameleonic properties deserve being considered also by pharma companies.**

30. Shalaeva M, Goetz G, Philippe L. Physical chemical properties for drug design in beyond rule-of-five chemical space. In: *2018 Medicinal Chemistry Reviews (Volume 53, chapter 21).* Bronson JJ (Ed.). Bristol-Myers Squibe, CT, USA, 421–435 (2018).