Critical comparison of shake-flask, potentiometric and chromatographic methods for lipophilicity evaluation (log $P_{o/w}$) of neutral, acidic, basic, amphoteric, and zwitterionic drugs

Adriana Port[a], Magda Bordas[a], Raquel Enrech[a], Rosalia Pascual[a], Martí Rosés[b], Clara Ràfols[b], Xavier Subirats[b,*], Elisabeth Bosch[b]

[a]*ESTEVE, Drug Discovery and Preclinical Development, Parc Científic de Barcelona, Baldiri Reixac, 4-8, 08028 Barcelona*

[b]*Institute of Biomedicine (IBUB) and Department of Chemical Engineering and Analytical Chemistry, Universitat de Barcelona, Martí i Franquès 1-11, 08028 Barcelona*

*Corresponding author

Dr. Xavier Subirats

xavier.subirats@ub.edu

Abstract

In the present study three different procedures have been compared for the determination of the lipophilicity of the unionized species (log $P_{o/w}$) of neutral, acidic, basic, amphoteric, and zwitterionic drugs. Shake-flask, potentiometric and chromatographic approaches have been assayed in a set of 66 representative compounds in different phases of advanced development. An excellent equivalence has been found between log $P_{o/w}$ values obtained by shake-flask and potentiometry, while the chromatographic approach is less accurate but very convenient for screening purposes when a high-throughput is required. In the case of zwitterionic and amphoteric compounds, either for shake-flask and chromatographic methods, the pH has to be accurately selected in order to ensure the compound to be in its neutral form.

Keywords: Drug lipophilicity; log $P_{o/w}$; log $P_{o/w}$ by liquid chromatography; log $P_{o/w}$ by potentiometry; log $P_{o/w}$ by shake-flask

1. Introduction

The design of compounds with suitable physicochemical properties is of paramount importance in drug discovery, since working in the optimal space may lead to improved pharmacokinetic and safety profiles(Arnott and Planey, 2012; Gleeson, 2008; Leeson and Springthorpe, 2007; Manallack et al., 2013; Ritchie et al., 2011; Wager et al., 2010; Walker, 2014). Among the relevant characteristics of a compound, lipophilicity is considered a very significant parameter and, commonly, it is expressed as the partition coefficient between *n*-octanol and aqueous phase (log $P_{o/w}$). Thus, lipophilicity determination is compulsory at early stages of the drug discovery process. Due to the wide variety of compound characteristics and the need of fast and reliable lipophilicity evaluation, different approaches for log $P_{o/w}$ estimation have been developed. Among them some stand out: the reference shake-flask method, accurate but excessively time consuming (Andrés et al., 2015; EPA, 1996; OECD, 1995), the automated chromatographic techniques, faster but suitable only for unionized compounds in working conditions (Donovan and Pescatore, 2002; Pallicer et al., 2010), or the potentiometric titration approaches, which are appropriated only for substances with acid-base properties, requiring in addition high purity samples (Avdeef, 1993, 1992; Ràfols et al., 2012). Therefore, a critical evaluation of the most used log $P_{o/w}$ determination methods through experimental measurements of a wide collection of drugs belonging to different chemical groups can provide a useful criterion to select the most appropriate approach for each instance.

In fact, a variety of lipophilicity values for a single drug are commonly reported in the literature and depend on the evaluation technique and the experimental conditions, mainly the working pH, used in the measurement procedure (Bio-Loom, 2017). In order to select the most appropriate measurement tool according to the drug chemical features, a critical study of the main causes of variability for the most used techniques becomes urgent. In this study, a diverse set of representative drugs of different chemical classes (neutral, acidic, basic, amphoteric, and zwitterionic compounds) have been independently examined and their log $P_{o/w}$ values evaluated by means of fully experimental procedures. The selected set of 66 acid-base APIs is representative of a previous selection of 2401 compounds of pharmaceutical interest, as explained later in the paper. Thus, shake-flask (with LC analysis using UV, MS or NMR detection), as well as a chromatographic approach which combines the chromatographic retention with the hydrogen bond donor molecular descriptor have been carefully examined. Biphasic potentiometric titrations involving the drug neutralization have been also tested for compounds with acid-base properties. The quality of the results obtained and the applicability, accuracy and precision of the mentioned approaches for each class of compounds have been critically evaluated. It should be pointed out here that “zwitterionic class of compounds” refers to compounds showing a zwitterion as the predominant neutral form, whereas the “amphoteric class

of compounds" presents a non-charged species as the main form at pH values in between the acidic and basic pK_a values.

2. Experimental

2.1 Materials

2.1.1 Apparatus

pH measurements were taken with a combined Crison (Hach Lange Spain, L'Hospitalet de Llobregat, Spain) 5202 electrode in a Crison 2001 pH meter. The electrode system was calibrated with ordinary aqueous buffers of pH 4.01, 7.00 and 9.21 (25 ºC).

Centrifugation was carried out with an Eppendorf (Hamburg, Germany) 5810R Centrifuge operating at 3000 rpm (radius of 16.8 cm) and 25 ºC for 15 minutes.

Chromatographic measurements were performed with a Waters (Milford, MA, USA) Alliance HPLC system and a Shimadzu (Kyoto, Japan) Nexera UHPLC instrument, both with diode array detector, or with an Agilent (Santa Clara, CA, USA) 1200 HPLC equipped with an Agilent 6540 UHD Accurate-Mass QTOF system which was operating in electron spray ionization (ESI) in positive mode. Instrument control and processing were performed by Empower1, LabSolutions and Masshunter 4.0, respectively. Retention data and peak areas were obtained from several columns: Waters XBridge BEH C_{18} (50 x 4.6 mm, 2.5 µm and 50 x 2.1 mm, 2.5 µm), Acquity BEH C_{18} (50 x 2.1 mm, 1.7 µm), and Agilent Zorbax Eclipse XDB-C18 (50 x 4.6 mm, 1.8 µm).

^{1}H NMR experiments were recorded on an Agilent Mercury 400 MHz (spectrometer fitted with a 5 mm ID/PFG probe) with ^{2}H lock in deuterated solvents. Acquisition and processing were done by VNMRj 4.2 and MestreNova, respectively.

A T3 titrator (from Sirius Analytical Instruments Ltd., East Sussex, UK) was used for the potentiometric determination of pK_a and log $P_{o/w}$. Instrument control and data processing were done by Sirius T3 v1.1 software.

2.1.2 Chemicals

Acetonitrile HPLC supragradient grade and acetonitrile LiChrosolv were purchased from Scharlab (Sentmenat, Spain) and Merck (Billerica, MA, USA), respectively. Solutions and solvent mixtures were made up of water purified by the Milli-Q® plus system from Millipore (Bedford, MA, USA) with a resistivity of 18.2 mΩ cm. Readymade 0.5M solutions of potassium hydroxide and hydrochloric acid were obtained from Merck and Sigma, respectively. 1-octanol ACS reagent and potassium bromide were purchased from Sigma (Saint Louis, Missouri, USA). Deuterated chloroform-d and methanol-d4 were obtained from Sigma Aldrich (Steinheim, Germany).

The chemicals used for buffer preparation were obtained from Merck (Darmstadt, Germany), Fluka (Steinheim, Germany), Sigma–Aldrich (Steinheim, Germany), Baker (Deventer, Netherlands), Riedel de Haën (Seelze, Germany) and Carlo Erba (Milano, Italy). Some of the drugs were purchased from Sigma-Aldrich (Steinheim, Germany), Alfa-Aesar (Ward Hill, USA), Prestwick Chemical (Illkirch, France) in high purity grade. Other drugs were synthesized in ESTEVE (Barcelona, Spain).

2.2 Methods

2.2.1 Shake-flask method: chromatographic quantification

Shake-flask measurements were performed following standard procedures described in OECD Guidelines (OECD, 1995) or from DMSO solutions as stated by Andrés et al., 2015. The *n*-octanol and aqueous phase were mutually saturated for 24 h. The compounds were grouped according to their pK_a values in order to select the appropriate buffers to ensure the neutral form of the solutes. 0.1N HCl, 0.1M NaOH, phosphoric acid, *N*-cyclohexyl-3-aminopropanesulfonic acid (CAPS) or acetic acid solutions were used to prepare aqueous buffer systems for the pH range 2-12. In the case of the amphoteric or zwitterionic molecules, the log $P_{o/w}$ values were measured at several pH values, including the isoelectric point.

The compounds were initially dissolved in DMSO at a usual concentration of 10 mM, followed by a dilution (generally 1:100) in aqueous buffer in the pH range 2-12. The resulting solutions were equilibrated with *n*-octanol for 1 h at 25 ºC. The phase ratio (V_w/V_o) varied from 1:1 to 100:1 depending on the expected log $P_{o/w}$ value of the given compound (Andrés et al., 2015). Generally, the procedure 1 proposed in the cited reference was followed, with the exception of olmesartan (1b) and oxybutynin (3). The two phases were separated by centrifugation and the solute concentration in the aqueous phase was determined by liquid chromatography with UV or MS detection.

2.2.2 Shake-flask method: NMR procedure

The NMR procedure was intended for the most hydrophilic compounds. A sample of 3-4 mg of the compound was dissolved in different partitions of non-deuterated aqueous phases and *n*-octanol, both saturated with each other, and the usual shake-flask equilibration was performed. Since the unionized form of the drug was needed, different aqueous phases were used depending on the sample: HCl 0.1 M (pH 1-2) was used for acidic compounds, NaOH 0.1 M (pH 12-13) for basic compounds, and MilliQ water in neutral compounds.

Just after the separation of the aqueous and octanol phases, the NMR samples were prepared with 400 µL of aqueous or octanol phase and 300 µL of DMSO-d6. The NMR experiments were conducted on an Agilent Mercury 400 MHz spectrometer at 25 ºC . The ^{1}H-NMR spectra were acquired using a 15º pulse with a relation delay of 15 s and 128 scans. For aqueous samples, the H_2O

signal was suppressed using WET sequence, with 64 scans and 15 s of relaxation delay. All spectra were carefully baseline corrected before integration.

log $P_{o/w}$ was calculated from the molar ratio of the drug in the two phases, using the integral of DMSO-d6 signal as a reference (Mo et al., 2010).

2.2.3 Potentiometric methods

Drug pK_a values were potentiometrically determined. A solution of the compound was neutralized over a wide pH range (from 2.0 to 12.0), and pK_a values were fitted from titration curves (mL of titrant vs. pH) by applying equations based on mass and charge balances. For bases and ampholytes, 1-2 mg of the samples were dissolved in 0.15 M KCl or in methanol/0.15 M KCl mixtures, pre-acidified to pH 2.0 with 0.5 M HCl, and then titrated with 0.5 M KOH solution. In the case of acids, the titration was performed in the opposite direction. In case of poorly soluble drugs, pK_a values were measured at several methanol/water compositions and aqueous pK_a was obtained from the Yasuda-Shedlovsky model (Avdeef et al, 1993).

log $P_{o/w}$ values were obtained from the difference between the aqueous pK_a of the species and the *apparent* pK_a determined from dual-phase titrations (*n*-octanol/KCl 0.15 M) (Avdeef, 1993, 1992). Typically, 1-2 mg of the samples was titrated as in aqueous pK_a, in presence of various amounts of the partitioning solvent, water-saturated *n*-octanol. The phase ratio applied was varied depending on the expected log $P_{o/w}$ value of the compound. log $P_{o/w}$ values were estimated and refined by a weighted non-linear least-squares, where the aqueous pK_a values were used as unrefined contributions. log $P_{o/w}$ values determined from different phase volume ratios were averaged and the ion-pair partitioning of charged species was also characterized.

2.2.4 Chromatographic methods

Retention data were obtained working with mobile phases at different pH values containing aqueous buffers with usually 40 or 50% (v/v) of acetonitrile. The compounds were grouped according to their pK_a values in order to select the appropriate buffers to ensure the neutral form of the solutes (generally, pH≈ pK_a -2 for acids and pH≈ pK_a +2 for bases). Trifluoroacetic, acetic, phosphoric, or citric acids or ammonium bicarbonate buffer (each at 0.010 or 0.1M, treated with various amounts of 0.1M NaOH or ammonium) were used as the aqueous buffer phase for the pH range 2-11. For pH 12 measurements, 10 mM pyrrolidine served as the aqueous phase. In case of amphoteric or zwitterionic molecules, log $P_{o/w}$ was measured at pH values corresponding to the isoelectric point. When pK_a values were too close to achieve the isoelectric point, a retention vs. pH profile was obtained in the pH range between 3 and 12 using 10 mM solutions of citric acid or ammonium hydrogencarbonate.

The drugs were dissolved in methanol, potassium bromide was used as dead time marker and a UV detector was employed.

In order to properly determine lipophilicity from chromatographic data it is needed to complement the measured retention with the hydrogen bond donor ability of the specific compound. In the present study the *A* molecular descriptor proposed by Abraham (Abraham, 1993) was selected as solute hydrogen bond acidity (Pallicer et al., 2013, 2010). *A* values for the selected compounds were calculated from the Absolv prediction module (ACD/Labs, 2017).

2.3 Computational software

A list of 122 2D-descriptors were calculated for the whole 2401 compounds with hydrogens added to the structure, using either Discovery Studio 2016 (Dassault Systèmes, 2017) or MOE (Molecular Operating Environment, 2017) software. Different classes of descriptors were used: topological such as the Balaban, Wiener, Zagreb, Kappa Shape and connectivity indices; adjacency and distance matrix descriptors such as the diameter or largest value in the distance matrix; graph-theoretical information content descriptors; molecular property, atom and bound counts, such as the number of hydrogen bond donors and acceptors and the vertex adjacency information; partial charge descriptors and finally physical properties such as atomic polarizabilities and surface area and volume related descriptors (listed in appendix 1). Only 2D-descriptors were selected in order to avoid any conformational ambiguity, and the finally chosen ones are well known for their contribution in determining quantitative structure activity relationships (Todeschini and Consonni, 2000).

Principal component analysis (PCA) and the calculation of the diversity metrics were done with Discovery Studio 2016.

3. Results and discussion

3.1 Representativeness assessment of the studied set of drugs

Firstly, a working set of 66 Active Pharmaceutical Ingredients (APIs) was *a priori* selected taking into account, mainly, the structural diversity. In order to assess the adequacy of the selected set of APIs, in relation to its representativeness versus the whole set of advanced compounds (phase III, launched or withdrawn), a search in the Thomson Reuters Integrity database (Clarivate Analytics, 2017) was performed (dietary supplements, polymers, mixtures and drug delivery systems, as well as inorganic complexes were excluded). The resulting 2679 unique structures were then further filtered by their molecular weight, requiring a value equal or lower than 600 Da, further by their number of rotable single bonds, accepting 15 or less, and by the fraction of rotatable single bonds divided by the number of bonds between heavy atoms, having to be equal or less than 0.5. The purpose of this

filtering stage was to concentrate on typical small molecule drugs, ending up with a set of 2401 unique structures. The 66 APIs selected *a priori* as working set were among these large group of drugs. To characterize the whole set of 2401 compounds a PCA was performed on the 122 calculated descriptors described above, with the first three principal components accounting for the 67.1% of the total variance (Fig. 1). Then the representativeness of our set of 66 APIs in the total space covered by the 2401 compounds was evaluated. The diversity metrics are shown in Table 1, which compiles minimum and average distances among the 66 APIs and the whole set of 2401 compounds using two alternatives for distance calculations: the properly scaled 122 descriptors and the ECFP_6 fingerprints (Rogers and Hahn, 2010). Table 1 also shows the cell coverage of both groups of compounds, where cells have been defined by the first three principal components, using two different bin sizes. The representativeness of the set of the 66 APIs can be assessed by looking at the minimum distance among its compounds, which increases by ten in comparison to the minimum distance that can be found among the compounds in the whole set of 2401 substances, and this either measuring the distance using real-value descriptors or ECFP_6 fingerprints. Additionally, the average cell density of the 66APIs decreases to almost one when partitioning the coordinate space defined by the first three PCAs using a bin size of two. This indicates low redundancies in terms of the chemical properties covered by the descriptors. For bigger bin sizes, as shown in the table for a bin size of five, the cell density for the selected compounds is still low and the density reduction when compared to the whole set is of one order of magnitude. In summary, it can be concluded that the selected set of 66APIs preserves the chemical diversity of the whole set but reducing redundancies.

3.2 Prior remarks: techniques, log $P_{o/w}$ ranges, and ionization of compounds

Recommending a unique and very reliable procedure for log $P_{o/w}$ determination of any drug or drug candidate is, at the moment, not realistic. This is because of the wide variety of chemicals involved in therapeutic treatments and also the specific limitations of each experimental approach. However, some remarks should be considered before discussion of results. Thus,

a) The reference method for lipophilicity estimation is the shake-flask, since it deals with the direct measurement of *n*-octanol/aqueous buffer partition. However, according to general guidelines, the experimental log $P_{o/w}$ values to be determined should be in principle between -2 and 4 (occasionally up to 5) (OECD, 1995).
b) To get reliable log $P_{o/w}$ values for acidic, basic, amphoteric, and zwitterionic compounds, when using either shake flask or chromatographic procedures, it is compulsory to make the measurements when the compound of interest is entirely in its neutral form, avoiding partially ionized species. In the shake-flask method, as long as pK_a values are accurately determined in water, it is easy to select an aqueous buffer with a suitable pH. However, some cautions should be

taken when the chromatographic approach is selected, because the hydroorganic mobile phase contains an important fraction of organic modifier that might significantly change the compound pK_a value, modifying in this way the molar fraction of the neutral species. Thus, acids increase their pK_a values with the content of organic modifier, whereas bases show the reversed trend (Subirats et al., 2007). This variation on pK_a values not only affects analytes, but also buffering species, leading to an increase or decrease of mobile phase pH values depending on the nature of the buffers and the content of organic modifier. This might be especially relevant in the case of amphoteric compounds. For instance, in aqueous solution when the acidic and the basic pK_a are close, the molar fraction of the neutral form might be far from 1. Nevertheless, the organic modifier content of the mobile phase is responsible for significant pH and pK_a variations, which can change the drug species present in working solution. As a general rule, it might be pointed out that increasing the acetonitrile content in the mixed solvent has a clear advantage for amphoteric compounds since the mole fraction of neutral species is generally raised. On the contrary, for zwitterions the mole fraction of neutral species decreases with the content of the organic solvent. Figure 2 shows the cases of tapentadol, enalapril, and telmisartan as representative examples of the above mentioned. The first drug is an amphoteric compound with a mole fraction of neutral species of about 60% at pH 10 in water, but with the addition of acetonitrile this mole fraction increases above 95% in a mobile phase containing a 50% of organic modifier, because of both the pK_a decrease of the basic functional group (from 9.44 to 8.77) and the pK_a increase of the acidic one (from 10.47 to 12.28), broadening the difference between the pK_a values of basic and acidic moieties from 1.0 in water to 3.5 in the hydroorganic mixture. Contrarily, enalapril is a zwitterionic compound which in the presence of organic solvent displays a lower separation between acidic and basic pK_a values (from 3.03 to 4.04 and from 5.35 to 4.43, respectively), shortening the pK_a separation from 2.3 in water to 0.4 in the mixed solvent. Finally, telmisartan is a nice and complex example of a zwitterionic drug in aqueous solution that becomes, mainly, an amphoteric compound with the addition of acetonitrile. In water the pK_a of the acidic moiety (4.39) is lower than that of the basic one (6.02), but this is no longer the case at 50% of acetonitrile (5.78 and 5.00). However, either in water or in the hydroorganic solvent, the molar fraction is significantly lower than 1.

c) The chromatographic method used in this work is a simplified approach (Pallicer et al., 2013) of the more complex one previously proposed (Pallicer et al., 2010). Both procedures take into account the experimental chromatographic retention and some molecular descriptors that were shown to be necessary for the proper calculation of log $P_{o/w}$. The solely retention measurement is unable to correctly describe the drug lipophilicity, since log $P_{o/w}$ quantity is independent of the hydrogen bond acidity whereas the chromatographic retention on C18 columns strongly depends on it (Pallicer et al., 2013; Valko et al., 2001). The used simplified approach involves only one

descriptor, the Abraham's *A* parameter, which can be easily estimated from the compound structure using either a commercial software package (ACD/Labs) or the open access UFZ-LSER Database (Ulrich et al., 2017). It was proved that this method is able to properly estimate log $P_{o/w}$ values from -1 to 7 (Pallicer et al., 2012), broadening in two orders of magnitude the lipophilicity range proposed in previous guidelines (OECD, 2004).

d) Two potentiometric titrations of the drug, with and without *n*-octanol, allow the log $P_{o/w}$ determination of compounds with acid-base properties. In these instances, the solubility of both acidic and basic species is the limiting feature to get reliable results. Regarding to the limits of applicability, a very good agreement was reported between log $P_{o/w}$ values obtained by shake-flask and potentiometry in the range between -1.8 and 5.8 for 23 structurally diverse compounds (Takács-Novák and Avdeef, 1996).

3.3 Comparison between shake-flask, potentiometric and chromatographic methods

The first challenge to solve was the adoption of an equivalence criterion between log $P_{o/w}$ values obtained from different techniques. There is a well-known experimental variability in log $P_{o/w}$ measurements, even following the same procedure, which led the Association of Official Analytical Chemists (AOAC) to admit differences of 0.3 units between log $P_{o/w}$ values measured from replicates using the shaking flask reference method (EPA, 1996; OECD, 1995). A tolerance limit of 0.6 was proposed for the chromatographic method used in this work (Pallicer et al., 2010), bearing in mind that different columns and mobile phases can be employed. This equivalence criterion is consistent with the variety of values shown for most compounds in the literature (Leo et al., 1995) and in the Bio-Loom database (Bio-Loom, 2017). Therefore, in the present work all the results included in an interval of 0.6 units between the higher and the lower obtained log $P_{o/w}$ values are considered as equivalent.

Table 2 shows the measured lipophilicity values for the representative set of 66 drugs belonging to five chemical classes, which were measured by the three methods previously mentioned: shake flask, chromatography, and potentiometry. Potentiometrically determined acidity constants, together with experimental literature (when available) and calculated log $P_{o/w}$ values (ClogP), are also listed in this table. ClogP (Bio-Loom, 2017) was the selected software because of its superior performance in the case of highly diverse compounds (Pallicer et al., 2014).

For some of the studied compounds not all the three experimental procedures could be successfully applied. The potentiometric procedure failed for those compounds with log $P_{o/w}$ values higher than 5.5 (clofazimine and rimonabant), and in the case of clopamide due to solubility issues. In case of very poorly soluble compounds the presence of a precipitate might not be visually evident for very low octanol-water volume ratios, since the stirred solutions are turbid. Diltiazem suffers from

degradation in alcohol-water mixtures (Andrisano et al., 2001), leading to non-reliable lipophilicity values in a shake-flask procedure involving long times in solution. A similar behavior was observed for clopidogrel. In relation to the chromatographic approach, topiramate and amantadine were not detected by UV, and isoproterenol was not sufficiently retained in the column.

As presented in Figure 3 and considering shake-flask as the reference method for log $P_{o/w}$ determination, excellent correlations were observed with potentiometry independently of the acid-base properties of the compounds. Only four compounds showed differences higher than 0.6 log $P_{o/w}$ units, the basic oxybutynin and prenylamine, and the zwitterionic levodopa and telmisartan. This figure also shows a good correlation between shake-flask and chromatography, but with a lower degree of equivalency between log $P_{o/w}$ values.

Table 3 shows the percentage of compounds with differences between the lowest and the highest log $P_{o/w}$ values obtained by different techniques not higher than 0.6 units. Equivalent results for the three approaches were obtained for 60% of the substances, being especially remarkable the high equivalency in the case of acids. As expected, the results obtained by shake-flask excellently matched those retrieved by potentiometry (Figure 3A), and a lower degree of equivalency was found when compared with chromatography (Figure 3B). However, it should be pointed out that several compounds, which were considered as non-equivalent, were in fact only slightly beyond the 0.6 limit.

3.4 Comparison with literature values

With the exception of levodopa, there is a good agreement between literature log $P_{o/w}$ values and those measured in this work by shake-flask, potentiometry, and chromatography. Only experimental lipophilicity values tagged with the "highest quality" in Bio-Loom database (obtained from different techniques, such as shake-flask, potentiometry, liquid chromatography, micellar or microemulsion electrokinetic chromatography, thin layer chromatography…) were used in this study for comparison. It must be pointed out that in some cases literature log $P_{o/w}$ values were determined for ionized species and then these values were somehow corrected for ionization. As shown in Table 4, lipophilicity data measured from shake-flask and potentiometric methods nicely agree with literature values, presenting relatively high determination coefficients (R^2), and intercepts and slopes close to 0 and 1, respectively. Correlation is slightly worse for the chromatographic method.

Figure 4 is a comprehensive plot of the correlations above mentioned, showing the different nature of the compounds studied (acids, bases, neutral, amphoteric, and zwitterionic) and the experimental procedures followed (shake-flask, pontentiometry, and chromatography). The main outlier is levodopa, for which both the shake-flask and chromatographic methods gave the same value (1.58 and 1.57, respectively), while potentiometry provided one log $P_{o/w}$ unit lower (0.50). These values were much higher than those selected by Bio-Loom database and calculated by the prediction

software ClogP (-2.74 and -2.84, respectively). In fact, the database reports several lipophilicity values for levodopa considered of lower quality in the range between -1.72 and -4.70. Labetalol, also a zwitterionic compound, shows very similar log $P_{o/w}$ values measured from the three different procedures, but about 1.5 units lower than the literature highest quality value (although other reported values in the same Bio-Loom database are comprised between 0.66 and 1.24). However, Avdeef reports a value of 1.33 for this compound (Avdeef, 2012), which would be equivalent to those obtained in this work by shake-flask, potentiometry, and chromatography. A similar feature is observed in the case of the basic compound ranitidine, with additional literature values of 0.27 (Bio-Loom, 2017) or 0.45 (Avdeef, 2012), much closer to those obtained by the three different techniques compared in this study.

Concerning the highly lipophilic substances rimonabant and clofazimine, there is again a good match between chromatographic and shake flask procedures (such a high lipophilicity is beyond the reach of actual potentiometric techniques), but they are lower than the reported literature log $P_{o/w}$. The value of 7.48 for clofazimine is very similar to the calculated ClogP value, but we should bear in mind that experimental determination of highly lipophilic compounds may lead to significant levels of uncertainty. In the case of rimonabant, a log $P_{o/w}$ value of 5.8 from chromatographic measurements is also reported in the Bio-Loom database, showing in this case a better agreement for both the chromatographic (6.00) and the shake-flask (5.57) methods assayed in this work.

3.5 Chromatographic considerations for amphoteric and zwitterionic compounds

Regarding the chromatographic approach for amphoteric and zwitterionic compounds, it must be pointed out that direct measurement of retention at a particular pH in the vicinity of isoelectric point might not be sufficient for the determination of a proper log $P_{o/w}$ value, especially in the case of compounds with close acidic and basic pK_a values. As previously commented (Figure 2), for zwitterionic compounds the pK_a difference between acidic and basic groups shortens with the addition of acetonitrile and thus the maximum recorded retention might not correspond to that of the fully unionized species (molar fraction lower than 1). Thus, it is convenient to measure retention in relative wide range of pH values (measured in the mobile phase) and fit the data to the following equation in order to obtain the "true" retention time of the unionized species (k_{HX}) (Rosés and Bosch, 2002):

$$k = \frac{k_{H_2X^+} + k_{HX}10^{(pH-pK_{a1})} + k_{X^-}10^{(2pH-pK_{a1}-pK_{a2})}}{1+10^{(pH-pK_{a1})}+10^{(2pH-pK_{a1}-pK_{a2})}} \quad (1)$$

where k is the retention factor measured at any pH, $k_{H_2X^+}$ and k_{X^-} are the fitted retention factors of the positively and negatively charged species, respectively, and pK_{a1} and pK_{a2} are the fitted acidity constants. Examples of application of this approach are presented in Figure 5 for tapentadol and

rosiglitazone. In mobile phases containing 50% of acetonitrile, the pK_a values of tapentadol are separated enough to have the compound at about pH 10.5 in its nearly fully neutral form. Thus, the fitted k_{HX} value (3.40) is very close to that measured at pH 10.49 (3.32). In contrast, the fitted k_{HX} of rosiglitazone (3.52) is significantly higher than the maximum retention directly achieved from a single chromatogram (2.97 and 2.96 at pH 6.22 and 6.59, respectively). Clearly, the elaboration of a retention vs. pH profile could be excessively time consuming for a chromatographic approach for lipophilicity determination, mainly intended to be time saving.

4. Conclusions

The present study confirms that shake-flask with LC-UV, LC-MS or NMR detection is the most universal method for determination of partition coefficients, since it is relatively simple and adequate for both neutral and ionizable compounds. For the set of 66 compound studied in this work it gives an excellent correlation with literature data, but its major drawbacks are that it is time consuming (phase equilibration + quantification procedure) and in the case of highly lipophilic or sparingly soluble compounds highly sensitive quantification systems are required. Potentiometric titrations are fully automatizable and significantly reduce the time required for measurement, but are only suitable for acidic or basic compounds, require samples of a high purity and might not be adequate for highly lipophilic substances due to solubility issues. Chromatographic measurements can be used for both neutral and ionizable compounds, are fully automatizable and time saving, but mobile phase compositions must be carefully selected in order to ensure the compound in its unionized species (mobile phase pH and analyte pK_a depend on the content of organic modifier).

The equivalency between log $P_{o/w}$ values obtained by shake-flask and potentiometry is excellent in the whole studied range (between -1.1 and 5.6). These results, obtained with a broader set of representative compounds, are consistent with a previous study (Takács-Novák and Avdeef, 1996), suggesting that the upper limit of log $P_{o/w}$ 4 recommended in the shake-flask guidelines (OECD, 1995) can be shifted up provided that and adequate ratio octanol/water is used and the quantification system is sensitive enough.

Generally, good correlations are obtained when comparing the shake-flask and the chromatographic methods. This time saving approach, although it might be less accurate, is particularly convenient for those highly lipophilic compounds beyond the limits of shake-flask and potentiometry and for solutes presenting stability issues in time consuming determinations.

Among all the compound classes studied, the amphoteric and zwitterionic compounds are the more complex. The determination of their lipophilicity is not straightforward, mainly because the pH at which the neutral species can be found might be difficult to identify. This is especially the case

when using chromatographic methods, where the variations on analyte pK_a and mobile phase pH with the addition of organic solvent should be taken into consideration, and consequently each case must be particularly examined.

ACKNOWLEDGEMENTS

This work was supported by the Ministry of Science, Innovation and Universities of Spain (project CTQ2017-88179-P AEI/FEDER, UE). Dr. Carmen Almansa is acknowledged for her helpful comments on the manuscript.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

TABLES

Table 1. Diversity metrics from the PCA study using the calculated 122 descriptors.

	APIs	Average / Minimum distance	Occupied cells / Average cell density	
			Bin-size 2	Bin-size 5
122 descriptors	2401	1.313 / 0.036	766 / 3.13	130 / 18.47
	66	1.350 / 0.304	59 / 1.12	31 / 2.13
Fingerprint	2401	0.909 / 0.039		
	66	0.899 / 0.380		

Table 2. pK_a and log $P_{o/w}$ values of the studied drugs.

		log $P_{o/w}$					
Compound	**pK_a values**[a]	**Shake-flask**[b]	**Potent.**[a]	**Chrom.**	**Bio-Loom**[i]	**Lit.**[j]	**ClogP**
Acidic							
Acetaminophen	9.39(A)	0.40 (1.0)	0.49	0.57	0.51	0.34	0.49
Atorvastatin	4.04(A)[d]	4.00 (2.0)	4.08[d]	4.50[g]	4.18	-	4.46
Celecoxib	9.55(A)	3.90 (2.0)	3.91	4.20[g]	-	-	4.57
Flufenamic	4.16(A)[d]	4.64 (2.0)	5.19[d]	4.83[g]	5.25	5.56	5.38
Glimepiride	5.38(A)[d]	4.02 (2.0)	3.97[d]	4.30[g]	-	-	3.96
Hydrochlorothiazide	8.72(A), 9.96(A)	0.00 (2.0)	-0.04	0.74	-0.07	-0.03	-0.37
Indomethacin	3.98(A)[d]	3.89 (2.0)	4.10[d]	3.83[g]	4.27	3.51	4.18
Ketorolac	3.50(A)[d]	2.71 (2.0)	2.62[d]	2.60	1.68	1.88	1.62
Naproxen	4.28(A)[d]	3.12 (2.0)	3.24[d]	2.93[g]	3.34	3.24	2.82
R-Flurbiprofen	4.35(A)[d]	3.97 (2.0)	3.84[d]	3.73[g]	3.86	3.99	3.75
Rosuvastatin	4.44(A)[d]	2.46 (2.0)	2.52[d]	2.58[g]	-	-	1.90
Topiramate	8.55(A)	0.47 (1.0)	0.58	-	-	-	0.04
Valsartan	3.84(A)[d], 4.69(A)[d]	3.37 (2.0)	3.52[d]	3.20[g]	3.9	3.90	4.86
Warfarin	5.01(A)[c]	3.19 (2.0)	3.28	3.41[g]	2.70	3.54	2.90
Zonisamide	9.49(A)	0.50 (2.0)	0.77	1.01	-	-	-0.36
Basic							
Amantadine	10.62(B)	2.32 (12.5)	2.52	-	2.44	-	2.00
Atenolol	9.40(B)[d]	0.13 (12.0)	0.06[d]	0.22[g]	0.16	0.22	-0.11
Chlorpromazine	9.25(B)[d]	5.40 (12.0)	5.27[d]	5.44[g]	5.35	5.40	5.50
Clofazimine	8.38(B)	6.30 (12.0)	-	5.93[g]	7.48	-	7.55
Clopidogrel	4.99(B)[d]	-	4.52[d]	4.84[g]	-	-	4.21
Diltiazem	7.79(B)[d]	-	2.84[d]	3.02[g]	2.8	2.89	3.65
Duloxetine	9.81(B)[d]	4.07 (12.0)	4.54[d]	4.04[g]	-	3.76[k]	4.26
Famotidine	6.67(B)	-0.75 (11.5)	-0.36	0.65	-0.8	-0.81	-0.86
Fluoxetine	9.89(B)[d]	4.21 (12.0)	4.42[d]	4.26[g]	4.5	4.50	4.57
Loratadine	4.86(B)[d]	4.45 (12.0)	4.88[d]	4.30[g]	4.4	-	5.05
Miconazole	5.99(B)[d]	5.58 (12.0)	5.38[d]	5.68[g]	5.34	4.89	5.81
Milnacipran	9.55(B)[d]	1.37 (12.0)	1.72[d]	2.57[g]	2.03	-	1.91
Mirtazapine	3.77(B), 7.65(B)	3.06 (12.0)	3.28	2.78	3.48	-	3.07
Oxybutynin	7.72(B)	5.29 (10.5)	4.59	5.34	-	-	4.69

Prenylamine	9.31(B)	3.82 (12.0)	5.07	5.94	-	-	5.80
Quetiapine	3.57(B)[o], 6.97(B)[o]	2.91 (12.0)	3.13[d]	2.50[g]	2.84	-	2.99
Ranitidine	2.18(B)[d], 8.38(B)[d]	-0.24 (11.5)	0.26[d]	0.06[g]	1.03	0.45	0.85
Rimonabant	2.80 (B)[d]	5.57 (12.0)	-	6.00[g]	6.7	-	6.12
Sertraline	9.31(B)[d]	4.73 (12.0)	5.17[d]	5.38[g]	-	4.9[l]	5.35
Terfenadine	9.27(B)	4.96 (12.0)	4.47	6.08	5.69	5.52	6.07
Tramadol	9.50(B)[d]	2.64 (11.5)	2.70[d]	3.28[g]	2.63	2.31	3.10
Trimipramine	9.21(B)	4.55 (12.0)	4.77	5.87	-	-	5.44
Venlafaxine	9.59(B)[d]	2.81 (12.0)	3.05[d]	3.74[g]	3.0	-	3.27
Verapamil	8.81(B)	3.63 (11.5)	4.07	4.43	3.97	4.33	4.27
Vildagliptin	7.52(B)[d]	-0.57 (11.5)	-0.16[d]	-0.08[g]	-	-	0.97
Neutral							
Carbamazepine	-	1.40 (7.0)	-	1.90[g]	2.19	2.45	2.24
Lacosamide	-	0.21 (7.0)	-	-0.14	-	-	0.39
Levetiracetam	-	-0.14 (2.0)	-	0.48[g]	-	-	-0.34
Oxcarbazepine	-	1.17 (7.0)	-	1.08	-	-	1.21
Sulfinpyrazone	-	1.35 (7.0)	-	2.26	2.30	3.93	2.43
Taranabant	-	4.94 (2.0)	-	5.69	-	5.2[m]	5.79
Amphoteric							
Clopamide	2.72(B), 8.95(A)	1.00 (5.2)	-	1.33	-	-	2.37
Folic acid	2.30(B), 3.79(A), 4.67(A), 7.97(B)	-	0.10	-	-	-	-1.70
Haloperidol	8.54(B)[f], 10.98(A)[f]	3.52[f]	3.61[f]	3.66[g]	3.82	3.67	3.85
Isoniazid	3.53(B)[e], 11.14(A)[e]	-0.65[e]	-0.85[e]	-0.95[h]	-0.70	-	-0.67
Isoproterenol	8.66(B), 9.95(A)	-	-0.62	-	-	-	0.15
Mebendazole	3.53(B)[f], 9.88(A)[f]	3.09[f]	2.92[f]	1.82	2.83	3.28	3.08
Nalidixic acid	6.00(A)[f]	1.36[f]	1.48[f]	1.98	1.59	1.41	1.02
Omeprazole	4.25(B)[d], 8.64(A)[d]	2.23	2.14[d]	1.40[g]	2.23	5.42	2.57
Pantoprazole	3.84(B), 8.22(A)[f]	2.07[f]	1.84[f]	1.34	-	-	2.11
Pioglitazone	5.56(B), 6.52(A)	-	4.03	3.22	-	3.31[n]	3.53
Rosiglitazone	6.26(B)[d], 6.67(A)[d]	-	3.10[d]	3.29	-	2.78[n]	3.02
Sulfamethoxazole	1.67(B), 5.65(A)[f]	0.86[f]	0.90[f]	1.44[g]	0.89	-	0.56
Tapentadol	9.44(B), 10.47(A)	-	2.88	3.37	-	-	3.15
Zwitterionic							
Benazepril	3.35(A)[f], 5.43(B)[f]	1.24[f]	1.38[f]	2.05	-	-	1.82

Ciprofloxacin	6.20(A)[f], 8.56(B)[f]	-1.13[f]	-1.15[f]	-1.20	-1.08	-1.08	-0.47
Enalapril	3.03(A), 5.35(B)	-0.04 (4.2)	-0.09	0.14	-0.07	0.16	0.67
Labetalol	7.41(A)[f], 9.37(B)[f]	1.45[f]	1.37[f]	1.74	3.09	1.33	2.50
Levodopa	2.77(A), 8.49(B), 10.29(A)	1.58 (5.6)	0.50	1.57	-2.74	-	-2.82
Ramipril	3.53(A), 5.79(B)	1.06 (4.7)	0.72	0.77	1.04	-	1.54
Telmisartan	3.01(B)[f], 4.39(A)[f], 6.02(B)[f]	4.18[f]	3.54[f]	4.03	-	7.46	7.29

[a] Potentiometrically determined at a ionic strength of 0.15 M
[b] pH of the aqueous phase in brackets; determined at a ionic strength of 0.10 M
[c] From ref. (Völgyi et al., 2007).
[d] From ref. (Pallicer et al., 2012).
[e] From ref. (Ràfols et al., 2012).
[f] From ref. (Ràfols et al., 2017).
[g] Calculated from ref. (Pallicer et al., 2013).
[h] Calculated from ref. (Pallicer et al., 2010).
[i] Compiled experimental log $P_{o/w}$ values with the "highest quality" tag (Bio-Loom, 2017), obtained from different techniques (shake-flask, potentiometry, chromatography…).
[j] Compiled experimental log $P_{o/w}$ values from ref. (Avdeef, 2012).
[k] From ref. (Martin et al., 2008).
[l] From ref. (Avdeef and Sun, 2011).
[m] From ref. (Liu et al., 2007).
[n] From ref. (Giaginis et al., 2007).
[o] From ref. (Garrido et al., 2006).

Table 3. Percentage of compounds with equivalent log $P_{o/w}$ obtained by different techniques (total number of compounds in each category in brackets).

	SF - P - C	SF - P	SF - C	P - C
Overall	60% (48)	92% (50)	68% (57)	74% (53)
Acidic	93% (14)	100% (15)	93% (14)	93% (14)
Basic	45% (20)	90% (21)	59% (22)	64% (22)
Neutral	-	-	50% (6)	-
Amphoteric	43% (7)	100% (7)	50% (8)	70% (10)
Zwitterionic	57% (7)	71% (7)	86% (7)	71% (7)

SF: shake-flask; P: potentiometry; C: chromatography

Table 4. Correlations between log $P_{o/w}$ values obtained from literature and measured in the present work (excluding levodopa). Standard errors of the fitted parameters in brackets.

	Shake-flask	Potentiometry	Chromatography
Intercept	-0.02 (0.13)	0.01 (0.13)	0.26 (0.16)
Slope	0.92 (0.04)	0.96 (0.04)	0.89 (0.05)
R^2	0.940	0.943	0.904
N	39	36	39

FIGURES

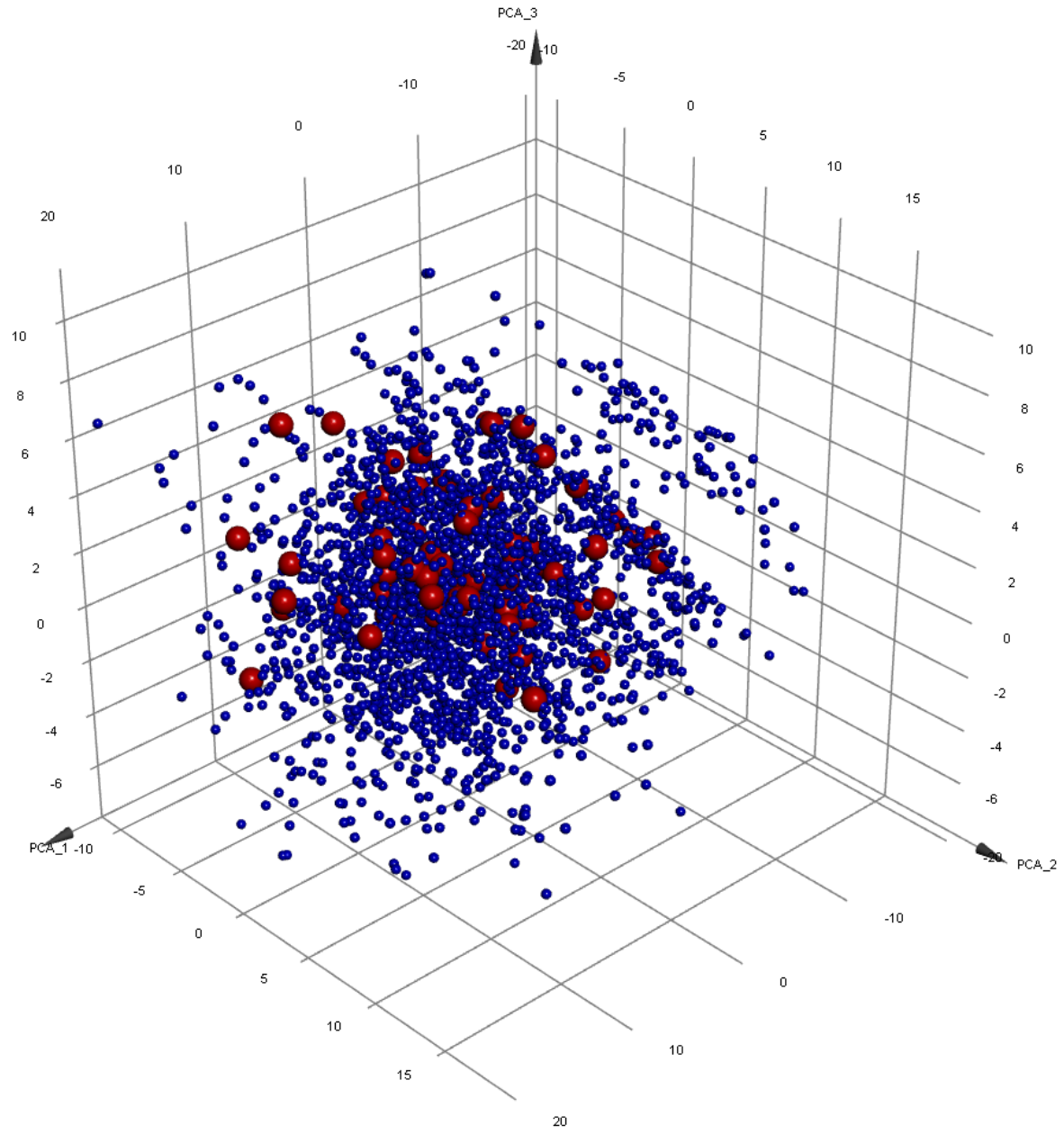

Figure 1. Representativeness assessment of the studied set of drugs by PCA. Large red symbols show the scores for the 66 selected APIs, and small blue ones those for the whole set of 2401 compounds.

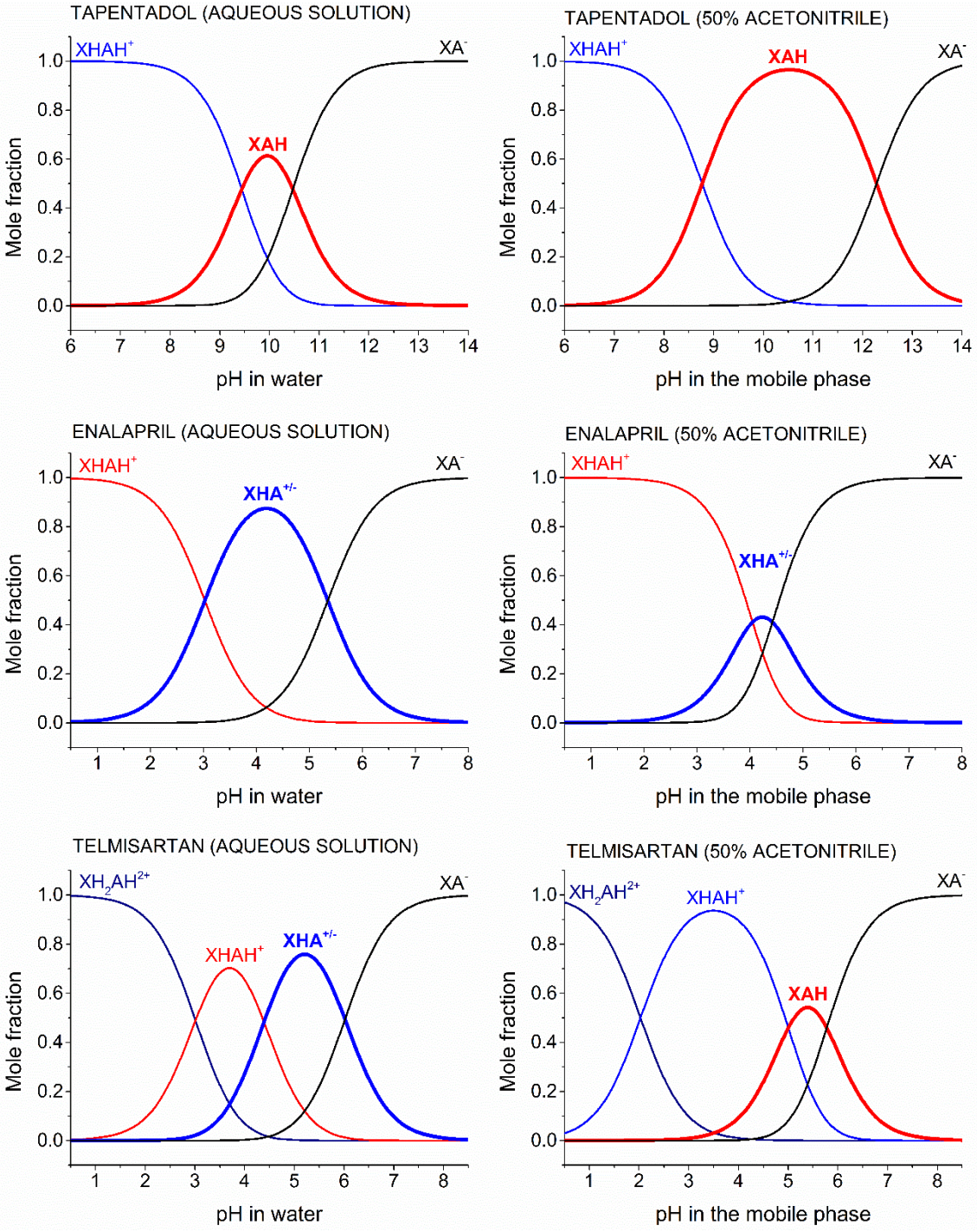

Figure 2. Mole fraction of the different species of tapentadol, enalapril and telmisartan in aqueous solution and in a mobile phase containing 50% in volume of acetonitrile. pK_a values in hydroorganic solvent were calculated (Subirats et al., 2006).

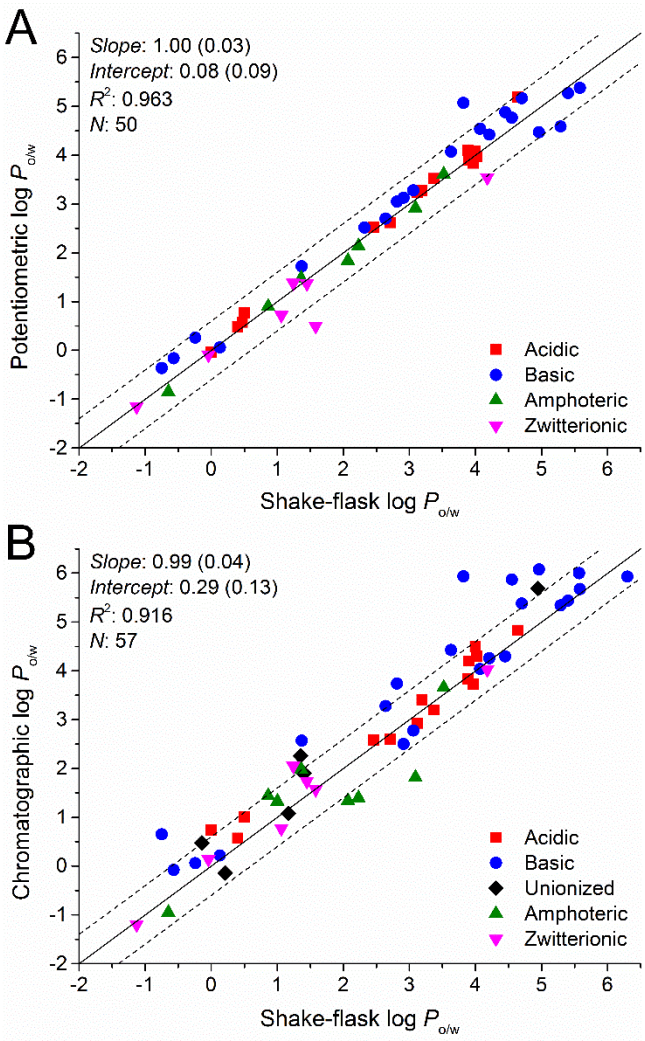

Figure 3. Correlations between potentiometric (A) and chromatographic (B) vs. shake flask log $P_{o/w}$ values. The statistics of linear regressions (slope, intercept, determination coefficient and number of observations) are presented in the figure, together with a solid straight line of null intercept and unitary slope, representing a total match between sets of log $P_{o/w}$ values, and dashed lines in order to show the ± 0.6 log $P_{o/w}$ range.

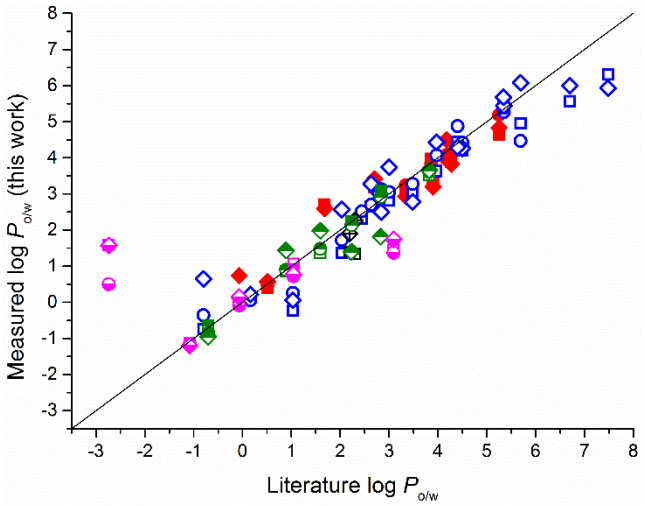

Figure 4. Correlations between log $P_{o/w}$ values measured in this work and found in the literature (Bio-Loom), depending on the technique employed and the acid-base properties of the drugs. Legend: (squares) shake-flask, (circles) potentiometry, (diamonds) chromatography; (filled symbols) acids, (empty) bases, (crossed) neutral, (upper-half filled) amphoteric, (lower-half filled) zwitterionic. A straight line of unitary slope and null intercept is also shown in the figure.

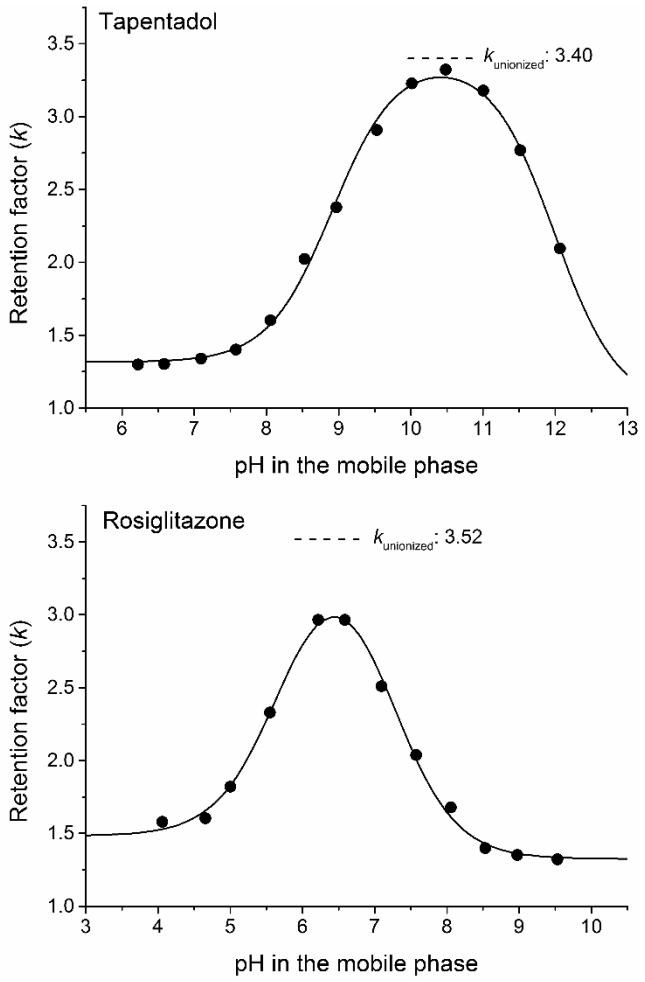

Figure 5. Retention vs. pH profile of the amphoteric tapentadol and rosiglitazone.

REFERENCES

Abraham, M.H., 1993. Scales of solute hydrogen-bonding: their construction and application to physicochemical and biochemical processes. Chem. Soc. Rev. 22, 73–83. doi:10.1039/cs9932200073

ACD/Labs. Advanced Chemistry Development, Inc. Toronto, ON, Canada [WWW Document], 2017. URL www.acdlabs.com

Andrés, A., Rosés, M., Ràfols, C., Bosch, E., Espinosa, S., Segarra, V., Huerta, J.M., 2015. Setup and validation of shake-flask procedures for the determination of partition coefficients (log *D*) from low drug amounts. Eur. J. Pharm. Sci. 76, 181–191. doi:10.1016/j.ejps.2015.05.008

Andrisano, V., Hrelia, P., Gotti, R., Leoni, A., Cavrini, V., 2001. Photostability and phototoxicity studies on diltiazem. J. Pharm. Biomed. Anal. 25, 589–597. doi:10.1016/S0731-7085(00)00588-4

Arnott, J.A., Planey, S.L., 2012. The influence of lipophilicity in drug discovery and design. Expert Opin. Drug Discov. 7, 863–875. doi:10.1517/17460441.2012.714363

Avdeef, A., 2012. Absorption and drug development, Second. ed, Absorption and Drug Development: Solubility, Permeability, and Charge State. John Wiley & Sons, Inc., Hoboken, NJ, USA. doi:10.1002/9781118286067

Avdeef, A., 1993. pH-metric log *P*. Part II. Refinement of partition coefficients and ionization constants of multiprotic substances. J. Pharm. Sci. 82, 183–190. doi:10.1002/jps.2600820214

Avdeef, A., 1992. pH-metric log *P*. Part I. Difference plots for determining ion-pair octanol-water partition coefficients of multiprotic substances. Quant. Struct. Relationships 11, 510–517. doi:10.1002/qsar.2660110408

Avdeef, A., Comer, J.E.A., Thomson, S.J., 1993. pH-Metric log *P*. 3. Glass electrode calibration in methanol-water, applied to pK_a determination of water-insoluble substances. Anal. Chem. 65, 42–49. doi:10.1021/ac00049a010

Avdeef, A., Sun, N., 2011. A new in situ brain perfusion flow correction method for lipophilic drugs based on the pH-dependent crone-renkin equation. Pharm. Res. 28, 517–530. doi:10.1007/s11095-010-0298-0

Bio-Loom. BioByte Corp. Claremont, CA, USA [WWW Document], 2017. URL www.biobyte.com

Clarivate Analytics, Philadelphia, PA, USA [WWW Document], 2017. URL www.clarivate.com

Dassault Systèmes BIOVIA, San Diego, CA, USA [WWW Document], 2017. URL www.accelrys.com

Donovan, S.F., Pescatore, M.C., 2002. Method for measuring the logarithm of the octanol-water partition coefficient by using short octadecyl-poly(vinyl alcohol) high-performance liquid chromatography columns. J. Chromatogr. A 952, 47–61. doi:10.1016/S0021-9673(02)00064-X

EPA, 1996. Product Properties Test Guidelines. OPPTS 830.7550. Partition coefficient (*n*-octanol/water), shake-flask method. U.S. Environmental Protection Agency, Prevention, Pesticides and Toxic Substances, United States.

Garrido, G., Ràfols, C., Bosch, E., 2006. Acidity constants in methanol/water mixtures of polycarboxylic acids used in drug salt preparations: Potentiometric determination of aqueous pK_a values of quetiapine formulated as hemifumarate. Eur. J. Pharm. Sci. 28, 118–127. doi:10.1016/j.ejps.2006.01.005

Giaginis, C., Theocharis, S., Tsantili-Kakoulidou, A., 2007. Investigation of the lipophilic behaviour of some thiazolidinediones. J. Chromatogr. B 857, 181–187. doi:10.1016/j.jchromb.2007.07.013

Gleeson, M.P., 2008. Generation of a set of simple, interpretable ADMET rules of thumb. J. Med. Chem. 51, 817–834. doi:10.1021/jm701122q

Leeson, P.D., Springthorpe, B., 2007. The influence of drug-like concepts on decision-making in medicinal chemistry. Nat. Rev. Drug Discov. 6, 881–890. doi:10.1038/nrd2445

Leo, A., Hoekman, D., Hansch, C., 1995. Exploring QSAR, hydrophobic, electronic, and steric constants. American Chemical Society, Washington, DC, USA.

Liu, P., Lin, L.S., Hamill, T.G., Jewell, J.P., Lanza, T.J., Gibson, R.E., Krause, S.M., Ryan, C., Eng, W., Sanabria, S., Tong, X., Wang, J., Levorse, D.A., Owens, K.A., Fong, T.M., Shen, C.P., Lao, J., Kumar, S., Yin, W., Payack, J.F., Springfield, S.A., Hargreaves, R., Burns, H.D., Goulet, M.T., Hagmann, W.K., 2007. Discovery of N-{(1S,2S)-2-(3-cyanophenyl)-3-[4-(2-[18F] fluoroethoxy)phenyl]-1-methylpropyl}-2-methyl-2-[(5-methylpyridin-2-yl)oxy] propanamide, a cannabinoid-1 receptor positron emission tomography tracer suitable for clinical use. J. Med. Chem. 50, 3427–3430. doi:10.1021/jm070131b

Manallack, D.T., Prankerd, R.J., Yuriev, E., Oprea, T.I., Chalmers, D.K., 2013. The significance of acid/base properties in drug discovery. Chem. Soc. Rev. 42, 485–496. doi:10.1039/C2CS35348B

Martin, R.S., Henningsen, R.A., Suen, A., Apparsundaram, S., Leung, B., Jia, Z., Kondru, R.K., Milla, M.E., 2008. Kinetic and Thermodynamic Assessment of Binding of Serotonin Transporter Inhibitors. J. Pharmacol. Exp. Ther. 327, 991 LP-1000.

Mo, H., Balko, K.M., Colby, D.A., 2010. A practical deuterium-free NMR method for the rapid determination of 1-octanol/water partition coefficients of pharmaceutical agents. Bioorganic Med. Chem. Lett. 20, 6712–6715. doi:10.1016/j.bmcl.2010.08.145

Molecular Operating Environment (MOE), 2013.08. Chemical Computing Group, Montreal, Canada [WWW Document], 2017. URL www.chemcomp.com

OECD, 2004. Test No. 117: Partition Coefficient (n-octanol/water), HPLC Method, OECD Guidelines for the Testing of Chemicals, Section 1. OECD Publishing, Paris.

doi:10.1787/9789264069824-en

OECD, 1995. Test No. 107: Partition Coefficient (n-octanol/water): Shake Flask Method, OECD Guidelines for the Testing of Chemicals, Section 1. OECD Publishing, Paris. doi:10.1787/9789264069626-en

Pallicer, J.M., Calvet, C., Port, A., Rosés, M., Ràfols, C., Bosch, E., 2012. Extension of the liquid chromatography/quantitative structure-property relationship method to assess the lipophilicity of neutral, acidic, basic and amphotheric drugs. J. Chromatogr. A 1240, 113–122. doi:10.1016/j.chroma.2012.03.089

Pallicer, J.M., Pascual, R., Port, A., Rosés, M., Ràfols, C., Bosch, E., 2013. The contribution of the hydrogen bond acidity on the lipophilicity of drugs estimated from chromatographic measurements. Eur. J. Pharm. Sci. 48, 484–493. doi:10.1016/j.ejps.2012.12.008

Pallicer, J.M., Pous-Torres, S., Sales, J., Rosés, M., Ràfols, C., Bosch, E., 2010. Determination of the hydrophobicity of organic compounds measured as log $P_{o/w}$ through a new chromatographic method. J. Chromatogr. A 1217, 3026–3037. doi:10.1016/j.chroma.2010.02.051

Pallicer, J.M., Rosés, M., Ràfols, C., Bosch, E., Pascual, R., Port, A., 2014. Evaluation of log $P_{o/w}$ values of drugs from some molecular structure calculation softwares. ADMET DMPK 2, 107–114. doi:10.5599/admet.2.2.45

Ràfols, C., Bosch, E., Ruiz, R., Box, K.J., Reis, M., Ventura, C., Santos, S., Araújo, M.E., Martins, F., 2012. Acidity and hydrophobicity of several new potential antitubercular drugs: Isoniazid and benzimidazole derivatives. J. Chem. Eng. Data 57, 330–338. doi:10.1021/je200827u

Ràfols, C., Subirats, X., Rubio, J., Rosés, M., Bosch, E., 2017. Lipophilicity of amphoteric and zwitterionic compounds: A comparative study of determination methods. Talanta 162, 293–299. doi:http://dx.doi.org/10.1016/j.talanta.2016.10.038

Ritchie, T.J., Ertl, P., Lewis, R., 2011. The graphical representation of ADME-related molecule properties for medicinal chemists. Drug Discov. Today 16, 65–72. doi:10.1016/j.drudis.2010.11.002

Rogers, D., Hahn, M., 2010. Extended-Connectivity Fingerprints. J. Chem. Inf. Model. 50, 742–754. doi:10.1021/ci100050t

Rosés, M., Bosch, E., 2002. Influence of mobile phase acid-base equilibria on the chromatographic behaviour of protolytic compounds. J. Chromatogr. A 982, 1–30. doi:10.1016/S0021-9673(02)01444-9

Subirats, X., Bosch, E., Rosés, M., 2006. Retention of ionisable compounds on high-performance liquid chromatography. XVI. Estimation of retention with acetonitrile/water mobile phases from aqueous buffer pH and analyte pK_a. J. Chromatogr. A 1121, 170–177. doi:10.1016/j.chroma.2006.03.126

Subirats, X., Rosés, M., Bosch, E., 2007. On the effect of organic solvent composition on the pH of buffered HPLC mobile phases and the pK_a of analytes - A review. Sep. Purif. Rev. 36, 231–255. doi:10.1080/15422110701539129

Takács-Novák, K., Avdeef, A., 1996. Interlaboratory study of log *P* determination by shake-flask and potentiometric methods. J. Pharm. Biomed. Anal. 14, 1405–1413. doi:10.1016/0731-7085(96)01773-6

Todeschini, R., Consonni, V., 2000. Handbook of molecular descriptors. Wiley-VCH, Weinheim.

Ulrich S.; Brown, T.N.; Watanabe, N.; Bronner, G.; Abraham, M.H.; Goss, K.-U., 2017. UFZ-LSER database v 3.2 [Internet].

Valko, K., Du, C., Bevan, C., Reynolds, D., Abraham, M., 2001. Rapid method for the estimation of octanol/water partition coefficient from gradient RP-HPLC retention and a hydrogen bond acidity term. Curr. Med. Chem. 8, 1137–1146. doi:10.2174/0929867013372643

Völgyi, G., Ruiz, R., Box, K., Comer, J., Bosch, E., Takács-Novák, K., 2007. Potentiometric and spectrophotometric pK_a determination of water-insoluble compounds: Validation study in a new cosolvent system. Anal. Chim. Acta 583, 418–428. doi:10.1016/j.aca.2006.10.015

Wager, T.T., Hou, X., Verhoest, P.R., Villalobos, A., 2010. Moving beyond rules: The development of a central nervous system multiparameter optimization (CNS MPO) approach to enable alignment of druglike properties. ACS Chem. Neurosci. 1, 435–449. doi:10.1021/cn100008c

Walker, M.A., 2014. Novel tactics for designing water-soluble molecules in drug discovery. Expert Opin. Drug Discov. 9, 1421–1433. doi:10.1517/17460441.2014.960839

APPENDIX 1

List of molecular descriptors calculated in this study (Todeschini and Consonni, 2000):

Num_H_Donors, Num_H_Acceptors, Num_RotatableBonds, Num_Rings, Num_AromaticRings, a_aro, a_count, a_donacc, a_heavy, a_hyd, a_IC, a_ICM, b_1rotN, b_1rotR, b_ar, b_count, b_double, b_heavy, b_max1len, b_rotN, b_rotR, b_single, b_triple, VAdjEq, VAdjMa, ALogP, Molecular_Weight, bpol, Apol, SlogP, h_logP, h_pKa, h_pKb, PEOE_PC+, PEOE_PC-, PEOE_RPC+, PEOE_RPC-, PEOE_VSA+0, PEOE_VSA+1, PEOE_VSA+2, PEOE_VSA+3, PEOE_VSA+4, PEOE_VSA+5, PEOE_VSA+6, PEOE_VSA-0, PEOE_VSA-1, PEOE_VSA-2, PEOE_VSA-3, PEOE_VSA-4, PEOE_VSA-5, PEOE_VSA-6, PEOE_VSA_FHYD, PEOE_VSA_FNEG, PEOE_VSA_FPNEG, PEOE_VSA_FPOL, PEOE_VSA_FPOS, PEOE_VSA_FPPOS, PEOE_VSA_HYD, PEOE_VSA_NEG, PEOE_VSA_PNEG, PEOE_VSA_POL, PEOE_VSA_POS, PEOE_VSA_PPOS, diameter, petitjean, petitjeanSC, radius, VDistEq, VDistMa, Molecular_SASA, Molecular_PolarSASA, Molecular_FractionalPolarSASA, Molecular_SAVol, Molecular_SurfaceArea, Molecular_PolarSurfaceArea, Molecular_FractionalPolarSurfaceArea, vsa_acc, vsa_don, vsa_hyd, vsa_other, vsa_pol, TPSA, vdw_area, vdw_vol, JX, JY, balabanJ, Wiener, Zagreb, CHI_0, CHI_1, CHI_2, CHI_3_P, CHI_3_C, CHI_3_CH, CHI_V_0, CHI_V_1, CHI_V_2, CHI_V_3_P, CHI_V_3_C, CHI_V_3_CH, IC, BIC, CIC, SIC, IAC_Total, IAC_Mean, V_ADJ_mag, V_DIST_mag, V_ADJ_equ, V_DIST_equ, E_ADJ_mag, E_DIST_mag, E_ADJ_equ, E_DIST_equ, Kappa_1, Kappa_2, Kappa_3, Kappa_1_AM, Kappa_2_AM, Kappa_3_AM, PHI, KierFlex.