Updating the portfolio of physicochemical descriptors related to permeability in the beyond the rule of 5 chemical space

Updating the portfolio of physicochemical descriptors related to permeability in the beyond the rule of 5 chemical space

Giuseppe Ermondi 1, Maura Vallaro 1, Gilles Goetz 2, Marina Shalaeva 3, Giulia Caron 4

1Department of Molecular Biotechnology and Health Sciences, University of Torino, Quarello 15, 10135 Torino, Italy.

2Hit Discovery and Optimization, Discovery Sciences, WWRD, Pfizer Inc, Eastern point Road, Groton, CT 06340, USA.

3Independent Scientist, Dana Point, CA 92629, USA.

4Department of Molecular Biotechnology and Health Sciences, University of Torino, Quarello 15, 10135 Torino, Italy. Electronic address: giulia.caron@unito.it.

Abstract

Beyond rule of 5 (bRo5) molecules are attracting significant interest in modern drug discovery mostly because many novel targets require large and more flexible structures.

The main aim of this paper is the identification of ad hoc bRo5 physicochemical descriptors of ionization, lipophilicity, polarity and chameleonicity and their measurement. We used different methods to collect ionization (pKa measures and log k'80 PLRP-S trends), lipophilicity (in octanol/water, in apolar systems and in biomimetic environments), polarity (Δlog Poct–tol, EPSA and Δlog KWIAM) and chameleonicity (ChameLogD) descriptors for 26 bRo5 drugs. A second aim was to check the relationship between physicochemical descriptors and permeability for a subset of compounds for which solid permeability values are reported in the literature.

Results showed that the physicochemical profile in the bRo5 chemical space is often experimentally accessible, albeit more tools are required to overcome limitations of individual methods. For the investigated compounds, permeability is governed by Δlog Poct–tol and preliminary data support that chameleonicity could also have an impact.

Keywords

Beyond rule of 5ChameleonicityIonizationLipophilicityPermeabilityPolarityProperty-based drug design

1. Introduction

Beyond rule of 5 (bRo5) molecules are attracting significant interest in modern drug discovery for at least two main reasons. First they have a potential to interact effectively with difficult-to-drug targets which cannot be modulated by Ro5 compliant small molecules i.e. ligands that reside in the chemical space defined by Lipinski´s rule of 5 (Ro5) (Lipinski et al., 1997; Doak et al., 2016; Doak and Kihlberg, 2016). And new promising compounds in medicinal chemistry such as proteolysis targeting chimeras (PROTACs) reside in the bRo5 chemical space (Zou et al., 2019; Edmondson et al., 2019).

According to a recent paper (Caron et al., 2019a) the following classes of compounds are included in the bRo5 chemical space: a) cyclic peptides (which are macrocyclic if their ring contains four or more amino acids), b) non-peptidic macrocycles (macrocycles are those molecular structures that contain one or more rings of at least 12 atoms) and c) non-macrocyclic compounds. Most of them are large and flexible structures which often require long synthetic routes and are associated with low yields. Therefore an efficient property-based drug design is essential for researchers who are called to design bRo5 candidates with the right solubility and permeability profiles (Leeson and Young, 2015; Shalaeva et al., 2018)

Various sequential processes affect passive drug permeation across cell membranes, i.e. desolvation, interactions with phospholipid head groups and access to the hydrophobic membrane interior (Krämer et al., 2009; Guimarães et al., 2012). Each of these steps is likely differently affected by the drug's molecular properties. For instance, polarity is expected to be involved in the desolvation step, whereas lipophilicity probably could roughly describe both interaction with phospholipids and passive membrane diffusion. Although not clearly stated, a different balance of polarity and lipophilicity is required for various series of compounds to be permeable. Routinely drug discovery projects are applying log P/log D in the octanol/water system and the topological polar surface area (TPSA) as lipophilicity and polarity descriptors, respectively. bRo5 molecules are large and flexible (often ionizable) and capable of adjusting their properties to the environment, thus acting as molecular chameleons (Whitty et al., 2016; Caron et al., 2019a) i.e., they can adopt a more apolar conformation when crossing a cell membrane and a more hydrophilic conformation in the aqueous intra and extra cellular environment. Several weak intramolecular interactions can in principle be formed and contribute to provide dynamic, environment-dependent shielding of the polar groups (Matsson et al., 2016). The best known are intramolecular hydrogen bonds (IMHBs) (Caron et al., 2019a) but recently the shielding of an amide bond by formation of an NH–π interaction with an aromatic group was investigated to design molecular chameleons (Tyagi et al., 2018; Nielsen et al., 2014).

Standard physicochemical tools used for Ro5 compliant compounds are often unsuitable for such complex structures. For instance, Shultz recently stated that descriptors such as TPSA,

HBA, and HBD overestimate solvent accessible polarity and thus underestimate permeability (Shultz, 2019). Therefore, there is an emerging need (Caron and Ermondi, 2017) for updating the portfolio of physicochemical descriptors specific to the bRo5 chemical space.

Ionization is accepted as a fundamental property in medicinal chemistry (Manallack et al., 2013; Charifson and Walters, 2014; Young and Leeson, 2018), but sometimes is poorly determined in many drug discovery programs which could present a major issue for series including compounds with different ionization profiles. A new chromatographic descriptor, log k'80 PLRP-S (Caron and Ermondi, 2017; Caron et al., 2016) at different pHs provides a reliable estimation of the ionization properties of the bRo5 drug candidate if pKa could not be measured for instance due to insufficient sample quantity, purity or low solubility. Secondly, lipophilicity should also be determined in systems different from octanol/water to better consider the membrane complexity. The Immobilized Artificial Membranes (IAM) chromatographic systems (Russo et al., 2017; Tsopelas et al., 2016) in which columns are prepared by covalent binding a monolayer of phospholipids to silica particles (Pidgeon et al., 1995) are of potential interest to this respect. Then, the experimental polarity descriptors, key determinants for permeability prediction (Caron and Ermondi, 2018), should be integrated in bRo5 drug discovery projects: Δlog Poct–tol, i.e. the difference between log Poct and log Ptol, (Caron and Ermondi, 2017; Ermondi et al., 2018) and EPSA that quantifies exposed polarity by retention time using controlled Supercritical Fluid Chromatography (SFC) conditions (Goetz et al., 2014a; Goetz et al., 2014b; Goetz et al., 2017) have attracted interest as polarity descriptors (Degoey et al., 2018; Hopkins et al., 2019; Vorherr et al., 2018). Notably, the Lipophilic Permeability Efficiency (LPE) that has been recently introduced in the literature as a new efficiency metric is a surrogate of Δlog Poct−tol. (Naylor et al., 2019) Also Δlog KWIAM introduced by Grumetto et al. (Grumetto et al., 2016) has been applied by some of us as a polarity descriptor involved in permeability prediction (Ermondi et al., 2018). Finally, as previously mentioned, most bRo5 compounds show a 'chameleonic' ability to change their conformation by exposing the polar groups in aqueous solution and burying them in apolar environments. Chameleonicity deserves therefore to be considered as an additional molecular property specific to large and flexible molecules. We recently demonstrated that the difference between the two chromatographic log P values (BRlogP and ElogP) provides a lipophilicity range which could be a first attempt to obtain an experimental chameleonicity index (Ermondi et al., 2019).

The main aim of this paper is therefore the identification of ad hoc bRo5 physicochemical descriptors of ionization, lipophilicity, polarity and chameleonicity, their measurement with different methods; to highlight their application domain, and their relationship with permeability. Therefore, in this study, we set-up a dataset of 26 bRo5 compounds including both cyclic and acyclic drugs, neutral and ionizable structures. For these compounds we collected a comprehensive list of ionization, lipophilicity, polarity and chameleonicity descriptors and discussed their experimental accessibility and relation with permeability.

2. **Experimental**

Source of drugs. All compounds were commercially available and had a purity >98%.

pKa determination. Acid dissociation constants were measured potentiometrically using a SiriusT3 instrument (Sirius Analytical Instruments) equipped with a Ag/AgCl double junction reference pH electrode and a turbidity sensing device. Titration experiments were conducted in 0.15 M KCl solution under nitrogen atmosphere at a temperature of 25 ± 1 °C. All tests were performed using standardized 0.5 M KOH and 0.5 M HCl as titration reagents. For low-solubility compounds, pKa values were determined in water–methanol mixtures (30–70% methanol v/v) and extrapolated to pure aqueous conditions using the Yasuda-Shedlovsky method. Reported experimental values are means of three determinations.

The pH-metric method to measure lipophilicity in octanol/water and toluene/water. Log P/D in octanol/water and toluene/water were measured potentiometrically with a SiriusT3 instrument (Sirius Analytical Instruments). Log P values were obtained from the difference between the aqueous pKa and the apparent pKa determined from dual phase titrations (Avdeef, 2001). About 1 mg of the samples were titrated as in aqueous pKa, in presence of various amounts of the partitioning solvent, water-saturated n-octanol and toluene. The phase ratio applied was varied depending on the expected log P but generally high log P conditions were used. Log P values were first estimated and then refined by a weighted non-linear least-squares, where the aqueous pKa values were used as unrefined contributions. log Ps from different phase volume ratios were finally averaged by taking into account the contribution of the ionized species.

Shake-flask. To measure log Ds a shake flask method was used. 2.5 mg of compound were dissolved in 3 mL of buffer solution with 0.15 M KCl. 1 mL of this solution was put in a separate vial to which 1 mL of octanol/toluene was added. The vials were vortexed for 10 min and the two phases were separated and analyzed by HPLC. Measurements were performed in triplicate.

BRlogD. The original method was described in a previous paper (Ermondi et al., 2019). Shortly we correlated log k'60 determined on a XBridgeTM Shield RP18 (Waters, 5 cm x 4.6 mm, 5 µm particle size, 130 Å pore size) with 60%%acetonitrile, 40%%buffer with log P of compounds. Here, we extend the method reliability to basic compounds using log D7.0 values.

The PLRP-S system to measure log k'80 PLRP-S. A PLRP-S column, 100 A, 5 uM, 50 × 4.6 mm; part no. PL1512–1500 from Agilent was used with the mobile phase (80%%acetonitrile−20%%buffer) flow rate 1 mL/min. Details of the method are reported elsewhere (Caron et al., 2016).

EPSA. Experimental details are reported elsewhere (Goetz et al., 2014a; Goetz et al., 2014b) and in Annex 3 (S.I).

IAM chromatography. The analyses were performed at 30 °C with 20 mM ammonium/acetate at pH 7.0 in mixture with acetonitrile at various percentages. The stationary phase was IAM PC.DD2 (Regis Technologies, 10 cm × 4.6 cm 10 µm packing 300 Å pore size). The flow rate was 1.0 ml/min. All samples were dissolved in the mobile phase. The adopted methodology to obtain log KWIAM and Δlog KWIAM was already described elsewhere (Ermondi et al., 2018). Δlog KWIAM was calculated using the shake-flask log D value.

A HPLC Varian ProStar instrument equipped with a 410 autosampler, a PDA 335 LC Detector and Galaxie Chromatography Data System Version 1.9.302.952 was used for all chromatographic analysis.

Molecular descriptors calculation. Most descriptors were calculated with Dragon v. 7.0.10 (Kode, srl). pKa was calculated with MoKa v.3.0.0 (www.moldiscovery.com).

3. **Results and discussion**

3.1. Dataset and 2D computed descriptors

The dataset is comprised of 26 compounds and is divided in two subsets. The first subset includes 15 molecules reported by Rossi Sebastiano et al. (Rossi Sebastiano et al., 2018) for which consistent permeability data are available. These are the training set of 10 compounds (asunaprevir, atazanavir, azithromycin, clarithromycin, erythromycin, rifampicin, ritonavir, roxithromycin, telaprevir, telithromycin) and the test set of 5 compounds (cyclosporin A (CsA), indinavir, paclitaxel, rifaximin, vinblastine). The second subset includes 11 additional drugs which were included in the study for several reasons: a) three rapamycins (everolimus, temsirolimus and sirolimus) as examples of poorly soluble compounds, b) rifapentine and rifabutin to extend the pool of ionization profiles exhibited by the rifamycin class, c) nelfinavir and saquinavir to extend the number of HIV-1 protease inhibitors, d) danoprevir to have at least two HCV NS3/4A protease inhibitors and e) three drugs which are not orally available (lanreotide, somatostatin 14, and vasopressin).

The SMILES codes of the 26 compounds (Table S1 in the Supporting Information) were submitted to the Dragon software to obtain the following molecular descriptors, i.e. the molecular weight (MW), the number of rotatable bonds (RBN), the Kier's flexibility index (PHI), the number of hydrogen bond donor groups (HBD), the number of hydrogen bond acceptor groups (HBA), the topological polar surface area (TPSA) and a calculated lipophilicity value (ALOGP) (Table 1).

Table 1. Data bars are used to visualize molecular descriptors of the investigated dataset (all the values are in Table S2).

Although ubiquitously reported, the descriptors in Table 1 should be cautiously interpreted in the description of the bRo5 datasets for several reasons. First, they refer to the 1D/2D structures of the molecules. Moreover, they can only be calculated for the neutral species of the drugs. As far as flexibility is concerned, PHI, introduced by Kier (Kier, 1989), is the most judicious descriptor to be used in the bRo5 chemical space, while RBN (Veber et al., 2002) is not suited for macrocyclic structures (Caron et al., 2019b). HBD and HBA counts may vary depending on the software used. Finally, as mentioned in the Introduction, TPSA overestimates polarity (Shultz, 2019) and lipophilicity values strongly depend on the calculation method (see below). Overall Table 1 should only be used to have a general idea of the dataset diversity, for instance somatostatin, vasopressin and lanreotide (not orally available drugs) are very different from the other drugs.

3.2. Ionization

The acid-base dissociation constant (pKa) of a drug is a key physicochemical parameter influencing many biopharmaceutical characteristics; however, for several reasons its experimental determination is not a very common practice in drug discovery programs. Predictions are largely preferred even though high calculation errors or calculation failures are often observed (Balogh et al., 2012) (see Annex S1).

We believe that the ionization profile of bRo5 compounds is key to understand their behavior. The first column in Table 2 shows the nature of the dominant (>50%) species of the compounds at pH 7.0. These assignments were obtained by combining pKa with chromatographic measurements as described below (all data are in Annex S1). pKa potentiometric determinations by T3 instrument provided results for most ionizable compounds, although solubility limits were found for atazanavir, ritonavir, telaprevir and paclitaxel. Moreover, the refinement procedure for lanreotide and somatostatin 14 was particularly difficult to perform because of the high number of ionization centers present on both molecules and no definitive results could be obtained. The chromatographic descriptor log k'80 PLRP-S (Caron et al., 2016; Caron and Ermondi, 2017) was also used to estimate the ionization properties of compounds. Briefly log k'80 PLRP-S at four different pHs (2, 5, 7, 10) are plotted against the pH of the buffer. The ionization profile of the considered compound is determined from the log k'80 PLRP-S trend (Caron et al., 2016). For instance, through log k'80 PLRP-S we could experimentally verify the basic nature of low soluble atazanavir (as mentioned above experimentally inaccessible to potentiometry) since log k'80 PLRP-S increases when passing from acidic to basic pHs. As a further example, in Fig. 1 we used the log k'80 PLRP-S trend to compare ionization behavior of the rifampicin class drugs (rifampicin, rifapentine, rifaximin, rifabutin). Rifampicin and rifapentine showed a similar ionization profile (zwitterionic), very different from rifaximin (neutral) and rifabutin (basic). Notably compounds belonging to the same class may have different ionization profiles.

According to experimental data, Table 2 shows that the dataset includes neutral, basic and acidic compounds (all the details are in Annex S1). Furthermore, rifampicin and rifapentine are zwitterions whereas lanreotide and somatostatin 14 are drugs with a complex and uncertain ionization profile.

3.3. Physicochemical descriptors and permeability

In Table 2 we report all the data (i.e. lipophilicity, polarity, chameleonicity and permeability) collected for the investigated series of compounds.

3.3.1. Lipophilicity

The lipophilicity of large and flexible molecules (i.e. bRo5) is problematic to predict as it pertains to the assessment of multiple conformations and their relative populations. log D is even more challenging to predict than log P because of the uncertainty added by pKa predictions (Goetz and Shalaeva, 2018). To illustrate log Poct calculation limits in the bRo5 chemical space we calculated log P with 8 calculators (MoKa, ACD, Marvin, MlogP, Xlogp3, IlogP, WlogP, Silicos-IT) for the subseries of five erythronolides (azithromycin, clarithromycin, erythromycin, roxitromycin, telithromycin) and for a series of beta-blockers (Ro5 compliant

drugs). A box and whisker plot is used to compare data (Fig. 2). Larger boxes represent variation in the lipophilicity values across methods used to calculate log P. Boxes found for erythronolides (Fig. 2B) are significantly larger than those obtained for beta-blockers (Fig. 2A) and demonstrate that log P of bRo5 compounds strongly depend on the calculation method. Similar results have been recently discussed by Schultz (Tyagi et al., 2018). Because of these findings we decided to avoid calculated log P in this study altogether (MoKa values and additional observations along these lines are reported in the Annex S2).

Several methods have been described in the literature to measure lipophilicity in the octanol/water system and the pros and cons of such methods have been discussed (their review is beyond the scope of this publication). In this study, we are interested in evaluating the relationship between permeability and lipophilicity. To do that we used several tools. First, we measured log P/log D values using the pH-metric method as implemented in the automatic titrator Sirius T3. Then shake-flask log D values were retrieved from the literature (Rossi Sebastiano et al., 2018) when available or measured in our laboratory. Because of the limits of direct methods, many research organizations have developed their own chromatographic tools for lipophilicity measurements, e.g. chromatographic hydrophobicity index (CHI) and Chrom Log D at GlaxoSmithKline (Valkó et al., 1997), or ElogD at Pfizer (Lombardo et al., 2001). We recently reported our own chromatographic method named BRlogD (Ermondi et al., 2019). log KWIAM has also been described as a valid surrogate of lipophilicity in octanol/water for neutral compounds (Tsopelas et al., 2016). We therefore measured ElogD, BRlogD and log KWIAM for the investigated dataset and reported their values in Table 2.

Obtaining reliable permeability values for bRo5 compounds is a complex endeavor, due to the nature of the analytes and because of the significant impact of active transport mechanisms (Matsson et al., 2016). To optimize the quality of permeability data we used log Papp of the training set reported by Rossi Sebastiano et al. (Rossi Sebastiano et al., 2018) (in blue in Table 2) since these permeability values were consistently obtained in a Caco-2 system. Notably we did not consider permeability data of the test set (in red in Table 2) since they were retrieved from the literature and present inconsistencies in the composition of the inhibitor cocktails implemented.

The relationships between permeability expressed as log Papp and lipophilicity indexes are reported in Fig. 3.

Fig. 3A shows a significant correlation (R2 = 0.63) between log Papp and log D7.4 as already described in the original paper (Rossi Sebastiano et al., 2018) (lipophilicity values were obtained using a standard shake–flask method). Fig. 3B and 3C shows the relationship between permeability and BRlogD and ElogD, respectively. Overall the lipophilicity descriptors are linearly correlated with permeability and similar equations are found. Fig. 3D shows that for this series of compounds log KWIAM does not significantly correlate with permeability. This could be related to the fact that log KWIAM is strongly dependent on the

presence of charges (Ermondi et al., 2018) and thus unsuitable for datasets of compounds with different ionization profiles (Table 1).

Overall the plots in Fig. 3 share a positive slope coefficient that supports the reliability of the found regressions (i.e. permeability tracks with lipophilicity).

Finally, we also tested log Poct (Table 2, potentiometric determinations) for modeling permeability. The low R2 value (0.38, plot not shown) suggests that log P is a poorer descriptor of permeability than log D.

3.3.2. Polarity

Polarity, as discussed in the Introduction, is one of the major determinants of permeability. Experimental polarity is quantified by three main descriptors (Caron and Ermondi, 2018). EPSA is an exposed polarity measurement which assesses polarity by retention time using controlled Supercritical Fluid Chromatography (SFC) conditions (Goetz et al., 2014a). Δlog Poct−tol is the difference between log Poct (the logarithm of the partition coefficient P in the octanol/water system) and log Ptol (the logarithm of the partition coefficient P in the toluene/water system). Δlog KWIAM is the difference between the logarithm of the experimental chromatographic retention factor (log KWIAM) and the value here named clog KWIAM (Ermondi et al., 2018) calculated using eq. (1).

$$\mathrm{clog}\, K_{\mathrm{W}}{}^{\mathrm{IAM}} = 0.92 * \log D_{\mathrm{oct}} \; - 1.03$$

Polarity data are in Table 2. Whereas most EPSA values were experimentally accessible, more limitations were found for Δlog KWIAM and Δlog Poct−tol since their determination (see above) needs a reliable value of log P/log D in octanol/water. Solubility issues in toluene also limited the experimental accessibility of Δlog Poct–tol. For the considered dataset we could determine Δlog Poct−tol for about 50% of the investigated compounds. Moreover, the experimental determination of Δlog Poct−tol requires some additional considerations. Since Δlog Poct−tol is the difference between two experimental values (Shalaeva et al., 2013), it is important to use the same method to obtain the single values. In this study, the pH-metric method was mostly used to measure both log Poct and log Ptol (Table 2). Shake-flask was used to determine Δlog Poct−tol of unionizable compounds and the zwitterionic rifampicin (for which a Δlog Doct−tol and not a Δlog Poct−tol was therefore determined). Lipophilicity of some compounds was beyond the experimental range of our shake-flask method (−2.5 <log P < 2.5). Finally, for indinavir and telithromycin we used the shake-flask method to confirm potentiometric values.

A relationship between permeability and polarity has been reported in many papers (their review is beyond the scope of this paper). In particular, the linear trend found between permeability and polarity (expressed as SA 3D PSA, the lowest polar surface area calculated on different crystallographic structures in the Rossi Sebastiano's training set is very strong (R2 = 0.89) (Rossi Sebastiano et al., 2018). When replacing SA 3D PSA with EPSA for the same

dataset (Table 2, compounds with permeability values in blue) the relationship between permeability and polarity is almost lost (Fig. 4A. R2 = 0.22). The relationship between permeability and Δlog KWIAM is shown in Fig. 4B (R2 = 0.48), higher than what was found for EPSA, but not excellent. The relationship between permeability and Δlog Poct–tol is very high (R2 = 0.89, Fig. 4C). Rifampicin is excluded from this regression since the balance of forces governing Δlog Doct–tol could be significantly different from Δlog Poct–tol.

Overall, the plots in Fig. 4 show that the increase of polarity is always linked to a decrease in permeability which supports the reliability of the found regressions. Moreover, Δlog Poct–tol has experimental limits but is the best descriptor of permeability. This last finding is rationalized using two different strategies, as discussed below. The Block Relevance (BR) analysis (Caron et al., 2013) is a computational tool that has been used to characterize the balance of intermolecular forces governing polarity descriptors (Caron and Ermondi, 2018). BR analysis showed that Δlog Poct–tol is the cleanest descriptor of exposed solutes HBD groups (Ermondi et al., 2014) whereas Δlog KWIAM is governed by both HBD and HBA groups and also influenced by the solutes dimensions. Moreover, BR analysis showed that for EPSA the presence of HBA groups could be considered as an interference and thus EPSA can fail to model permeability when many and variable HBA groups are present in the chemical structure, which is the case in the bRo5 chemical space (Goetz et al., 2017). Besides BR analysis, we also used a second strategy to evaluate the significance of the three polarity descriptors. We reasoned that polarity calculated as SA 3D PSA (i.e. on the crystallographic structures (Rossi Sebastiano et al., 2018)) is a credible polarity value. Then we selected the subset of erythronolides for which all experimental data were obtained and normalized all the values (data in Table S3). Finally, we plotted on the Y axis SA 3D PSA and on the X axis the three experimental descriptors, i.e. EPSA, Δlog Poct–tol and Δlog KWIAM (Fig. 5).

3.3.3. Combining lipophilicity and polarity to model permeability

At first glance, Fig. 3 and 4 seem to suggest that both lipophilicity and polarity play a role in governing permeability in the investigated dataset of bRo5 compounds. However, the strong correlation found between Δlog Poct–tol (as well as SA 3D PSA) and permeability suggest that polarity is the main determinant, while lipophilicity is less important. This result is probably explained by the fact that compounds in this dataset have similar size and shapes (their radius of gyration covered a relatively narrow interval) (Rossi Sebastiano et al., 2018). The weak but non-negligible relationship with lipophilicity could be ascribed to the polarity component included in log Doct (Lombardo et al., 2000; Abraham et al., 2002; Caron et al., 2013).

3.4. Chameleonicity

Molecular chameleons do exist as shown by CsA (Alex et al., 2011) and some other examples (Tyagi et al., 2018). But how can one experimentally determine the extent of chameleonicity of a given compound? And how relevant is chameleonicity in permeability modulation? In a recent paper (Ermondi et al., 2019) we showed that two validated chromatographic methods

with different experimental conditions provide very similar log P for small molecules but give significantly different log P (but linearly correlated) for large and flexible molecules. This trend has been confirmed here also for basic compounds for which log D values were originally reported (Table 2 and Fig. 6A).

The difference between the two chromatographic log D values (BRlogD and ElogD) provides an experimental log D range (called ChameLogD, Table 2) which is a first attempt to experimentally quantify the chameleonic properties of drugs.

Fig. 6B shows the relationship between permeability and ChameLogD. Although the number of compounds is limited and we cannot generalize, the higher the lipophilicity range, the higher the permeability. In other words, the higher the capacity of a compound to behave as a chameleon, the higher the permeability.

4. **Conclusion**

To increase the accessibility to bRo5 chemical space as a source of new drugs we need to use ad hoc tools, not easily transferable from the Ro5 world. In particular the fact that "…there is no substitute for generating measured data…"(Young and Leeson, 2018) recently outlined by Young and Leeson is essential for bRo5 compounds where calculations often fail.

In this paper we outlined that physicochemical descriptors are experimentally accessible for most bRo5 drugs. In particular, for neutral and basic bRo5 compounds, the determination of lipophilicity and polarity is relatively straightforward whereas more difficulties are encountered with acidic, zwitterionic and compounds with a complex ionization profile (parenteral drugs in this study). That highlights the necessity of thorough ionization properties assessment before measuring other physicochemical descriptors. To aid pKa determination, a chromatographic tool such as log k'80 PLRP-S could be used as a valid alternative to assess the ionization properties of bRo5 compounds.

Availability of more than one experimental method for measuring log Doct is crucial. (i.e. chromatography alone is not sufficient). Notably, although log Doct is not the major determinant of permeability by itself, its determination is essential to obtain polarity and chameleonicity descriptors.

Polarity has been confirmed to be a crucial descriptor for permeability. However different polarity descriptors (EPSA, Δlog Poct–tol and Δlog KWIAM) have different experimental limitations. Furthermore, the balance of intermolecular forces need to be known in order to define their applicability domain for a particular set of compounds. In this paper we have shown that Δlog Poct–tol is the best polarity descriptor.

Since bRo5 compounds can adapt their properties to the environment and act as molecular chameleons, here we defined a new experimental descriptor called ChameLogD to quantify chameleonic properties of compounds. Preliminary results suggest a good relationship between ChameLogD and permeability.

CRediT authorship contribution statement

Giuseppe Ermondi: Conceptualization, Software. Maura Vallaro: Data curation. Gilles Goetz: Data curation, Writing - review & editing. Marina Shalaeva: Writing - original draft. Giulia Caron: Conceptualization, Writing - original draft.

References

Abraham, M.H., Ibrahim, A., Zissimos, A.M., Zhao, Y.H., Comer, J.E., Reynolds, D., 2002. Application of hydrogen bonding calculations in property based drug design. Drug Discov. Today 7, 1056–1063.

Alex, A., Millan, D.S., Perez, M., Wakenhut, F., Whitlock, G., 2011. Intramolecular hy- drogen bonding to improve membrane permeability and absorption in beyond rule of five chemical space. Med. Chem. Comm. 2, 669–674.

Avdeef, A., 2001. Physicochemical profiling (Solubility, permeability and charge state).

Curr. Top Med. Chem. 1, 277–351.

Balogh, G.T., Tarcsay, Á., Keseru, G.M., 2012. Comparative evaluation of pKa prediction tools on a drug discovery dataset. J. Pharm. Biomed. Anal. 67–68, 63–70.

Caron, G., Digiesi, V., Solaro, S., Ermondi, G., 2019a. Flexibility in early drug discovery: focus on the beyond-Rule-of-5 chemical space. Drug Discov. Today. https://doi.org/10.1016/j.drudis.2020.01.012. in press.

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

Caron, G., Ermondi, G., 2017. Updating molecular properties during early drug discovery.

Drug Discov. Today 22, 835–840.

Caron, G., Kihlberg, J., Ermondi, G., 2019b. Intramolecular hydrogen bonding: an op- portunity for improved design in medicinal chemistry. Med. Res. Rev. 39, 1–23.

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

Caron, G., Vallaro, M., Ermondi, G., Goetz, G.H., Abramov, Y.A., Philippe, L., Shalaeva, M., 2016. A fast chromatographic method for estimating lipophilicity and ionization in nonpolar membrane-like environment. Mol. Pharm. 13, 1100–1110.

Charifson, P.S., Walters, W.P., 2014. Acidic and basic drugs in medicinal chemistry: a perspective. J. Med. Chem. 57, 9701–9717.

Degoey, D.A., Chen, H.J., Cox, P.B., Wendt, M.D., 2018. Beyond the rule of 5: lessons learned from Abbvie's drugs and compound collection. J. Med. Chem. 61, 2636–2651.

Doak, B.C., Kihlberg, J., 2016. Drug discovery beyond the rule of 5 - opportunities and challenges. Expert Opin. Drug Discov. 12, 1–5.

Doak, B.C., Zheng, J., Dobritzsch, D., Kihlberg, J., 2016. How beyond rule of 5 drugs and clinical candidates bind to their targets. J. Med. Chem. 59, 2312–2327.

Edmondson, S.D., Yang, B., Fallan, C., 2019. Proteolysis targeting chimeras (PROTACs)in 'beyond rule-of-five' chemical space: recent progress and future challenges. Bioorg.

Med. Chem. Lett. 29, 1555–1564.

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

Ermondi, G., Vallaro, M., Goetz, G., Shalaeva, M., Caron, G., 2019. Experimental lipo- philicity for beyond rule of 5 compounds. Futur. Drug Discov. 1, 1–12.

Ermondi, G., Visconti, A., Esposito, R., Caron, G., 2014. The block relevance (BR) analysis supports the dominating effect of solutes hydrogen bond acidity on Δlog Poct-tol. Eur.

J. Pharm. Sci. 53, 50–54.

Goetz, G.H., Farrell, W., Shalaeva, M., Sciabola, S., Anderson, D., Yan, J., Philippe, L., Shapiro, M.J., 2014a. High throughput method for the indirect detection of in- tramolecular hydrogen bonding. J. Med. Chem. 57, 2920–2929.

Goetz, G.H., Philippe, L., Shapiro, M.J., 2014b. EPSA: a novel supercritical fluid chro- matography technique enabling the design of permeable cyclic peptides. ACS Med. Chem. Lett. 5, 1167–1172.

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

Goetz, G.H., Shalaeva, M., Caron, G., Ermondi, G., Philippe, L., 2017. Relationship be- tween passive permeability and molecular polarity using block relevance analysis. Mol. Pharm 14, 386–393.

Grumetto, L., Russo, G., Barbato, F., 2016. Polar interactions drug/phospholipids esti- mated by IAM–HPLC vs cultured cell line passage data: their relationships and comparison of their effectiveness in predicting drug human intestinal absorption. Int.

J. Pharm. 500, 275–290.

Guimarães, C.R.W., Mathiowetz, A.M., Shalaeva, M., Goetz, G., Liras, S., 2012. Use of 3D properties to characterize beyond rule-of-5 property space for passive permeation. J. Chem. Inf. Model. 52, 882–890.

Hopkins, B.A., Lee, H., Ha, S., Nogle, L., Sauvagnat, B., Mcminn, S., Smith, G.F., Sciammetta, N., 2019. Development of a platform to enable efficient permeability evaluation of novel organo-peptide macrocycles. ACS Med. Chem. Lett. 10, 874–879.

Kier, L.B., 1989. An index of molecular flexibility from kappa shape attributes. Quant.

Struct.-Act. Relat. 8, 218–221.

Krämer, S.D., Lombardi, D., Primorac, A., Thomae, A.V., Wunderli-Allenspach, H., 2009. Lipid-bilayer permeation of drug-like compounds. Chem. Biodivers. 6, 1900–1916.

Leeson, P.D., Young, R.J., 2015. Molecular property design: does everyone get it? ACS Med. Chem. Lett 6, 722–725.

Lipinski, C.A., Lombardo, F., Dominy, B.W., Feeney, P.J., 1997. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv. Drug Deliv. Rev. 23, 3–26.

Lombardo, F., Shalaeva, M.Y., Tupper, K.a, Gao, F., Abraham, M.H., 2000. ElogPoct: a tool for lipophilicity determination in drug discovery. J. Med. Chem. 43, 2922–2928.

Lombardo, F., Shalaeva, M.Y., Tupper, K.A., Gao, F., 2001. ElogD oct : a tool for lipo- philicity determination in drug discovery. 2. Basic and neutral compounds.

J.Med.Chem. 44, 2490–2497.

Manallack, D.T., Prankerd, R.J., Yuriev, E., Oprea, T.I., Chalmers, D.K., 2013. The sig- nificance of acid/base properties in drug discovery. Chem. Soc. Rev. 42, 485–496.

Matsson, P., Doak, B.C., Over, B., Kihlberg, J., 2016. Cell permeability beyond the rule of

5. Adv. Drug Deliv. Rev. 101, 42–61.

Naylor, M.R., Ly, A.M., Handford, M.J., Ramos, D.P., Pye, C.R., Furukawa, A., Klein, V., Noland, R.P., Edmondson, Q., Turmon, A.C., Hewitt, W.M., Schwochert, J., Townsend, C.E., Kelly, C.N., Blanco, M.-.J., Lokey, R.S., 2019. Lipophilic permeability efficiency (LPE) reconciles the opposing roles of lipophilicity in membrane perme- ability and aqueous solubility. J. Med. Chem. 61, 11169–11182.

Nielsen, D.S., Hoang, H.N., Lohman, R.J., Hill, T.A., Lucke, A.J., Craik, D.J., Edmonds,

D.J., Griffith, D.A., Rotter, C.J., Ruggeri, R.B., Price, D.A., Liras, S., Fairlie, D.P., 2014. Improving on nature: making a cyclic heptapeptide orally bioavailable. Angew. Chem. Int. Ed. Engl. 53, 12059–12063.

Pidgeon, C., Ong, S., Liu, H., Qiu, X., Pidgeon, M., Dantzig, a H., Munroe, J., Hornback, W.J., Kasher, J.S., Glunz, L., 1995. IAM chromatography: an in vitro screen for predicting drug membrane permeability. J. Med. Chem. 38, 590–594.

Rossi Sebastiano, M., Doak, B.C., Backlund, M., Poongavanam, V., Over, B., Ermondi, G., Caron, G., Matsson, P., Kihlberg, J., 2018. Impact of dynamically exposed polarity on permeability and solubility of chameleonic drugs beyond the rule of 5. J. Med. Chem. 61, 4189–4202.

Russo, G., Grumetto, L., Barbato, F., Vistoli, G., Pedretti, A., 2017. Prediction and me- chanism elucidation of analyte retention on phospholipid stationary phases (IAM- HPLC) by in silico calculated physico-chemical descriptors. Eur. J. Pharm. Sci. 99, 173–184.

Shalaeva, M., Caron, G., Abramov, Y.A., Connell, T.N.O., Plummer, M.S., Yalamanchi, G., Farley, K.A., Goetz, G.H., Philippe, L., Shapiro, M.J., 2013. Integrating intramolecular hydrogen bonding (IMHB) considerations in drug discovery using Δ logP as a tool. J. Med. Chem. 56, 4870–4879.

Shalaeva, M., Goetz, G.H., Philippe, L., 2018. Physical chemical properties for drug design in beyond rule-of-five chemical space. Med. Chem. Rev. 53.

Shultz, M.D., 2019. Two decades under the influence of the rule of five and the changing properties of approved oral drugs. J. Med. Chem. 62, 1701–1714.

Tsopelas, F., Vallianatou, T., Tsantili-Kakoulidou, A., 2016. Advances in immobilized artificial membrane (IAM) chromatography for novel drug discovery. Expert Opin. Drug Discov. 0441, 1–16.

Tyagi, M., Poongavanam, V., Lindhagen, M., Pettersen, A., Sjö, P., Schiesser, S., Kihlberg, J., 2018. Toward the design of molecular chameleons: flexible shielding of an amide

bond enhances macrocycle cell permeability. Org. Lett. 20, 5737–5742.

Valkó, K., Bevan, C., Reynolds, D., 1997. Chromatographic hydrophobicity index by fast-gradient RP-HPLC: a high-throughput alternative to log P/log D. Anal. Chem. 69, 2022–2029.

Veber, D.F., Johnson, S.R., Cheng, H.Y., Smith, B.R., Ward, K.W., Kopple, K.D., 2002.

Molecular properties that influence the oral bioavailability of drug candidates. J.Med.Chem. 45, 2615–2623.

Vorherr, T., Lewis, I., Berghausen, J., Desrayaud, S., Schaefer, M., 2018. Modulation of oral bioavailability and metabolism for closely related cyclic hexapeptides. Int. J. Pept. Res. Ther. 24, 35–48.

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

Young, R.J., Leeson, P.D., 2018. Mapping the efficiency and physicochemical trajectories of successful optimizations. J. Med. Chem. 61, 6421–6467.

Zou, Y., Ma, D., Wang, Y., 2019. The protac technology in drug development. Cell Biochem. Funct. 37, 21–30.

Table 1
Data bars are used to visualize molecular descriptors of the investigated dataset (all the values are in Table S2).

Name	MW	RBN	PHI	HBD	HBA	TPSA	ALOGP
Asunaprevir							
Atazanavir							
Azithromycin							
Clarithromycin							
CsA							
Danoprevir							
Erythromycin							
Everolimus							
Indinavir							
Lanreotide							
Nelfinavir							
Paclitaxel							
Rifabutin							
Rifampicin							
Rifapentine							
Rifaximin							
Ritonavir							
Roxithromycin							
Saquinavir							
Sirolimus							
Somatostatin 14							
Telaprevir							
Telithromycin							
Temsirolimus							
Vasopressin							
Vinblastine							
Average	*855*	*12*	*17.4*	*6*	*16*	*231*	*3.4*

Table 2
Physicochemical data of the investigated series of bRo5 compounds.

Name	Charge 7.0	log $D^{7.4}$	BRlogD	ElogD	ChameLogD	log K_W^{IAM}	Δlog K_W^{IAM}	EPSA	log P_{oct}	log P_{tol}	Δlog $P_{oct-tol}$	log P_{app}
Asunaprevir	1-	3.1	NA	NA	NA	3.02	1.20	87	3.3	NA	NA	-4.32
Atazanavir	0	4.2	3.1	4.7	1.6	2.74	-0.09	68	4.4	3.9	0.5	-4.14
Azithromycin	2+	1.1	0.8	1.4	0.6	0.56	0.58	82	4.0	3.0	1.0	-5.55
Clarithromycin	1+	1.6	2.0	ND	NA	1.89	1.45	73	3.4	2.9	0.5	-4.55
CsA	0	NA	5.9	8.2	2.3	4.71	NA	71	NA	NA	NA	-5.6
Danoprevir	1-	ND	NA	NA	NA	2.99	NA	NA	2.9	3.2	-0.3	ND
Erythromycin	1+	0.9	0.9	ND	NA	1.33	1.53	77	3.1	1.7	1.4	-5.82
Everolimus	0	ND	5.0	6.7	1.7	4.15	NA	80	NA	NA	NA	ND
Indinavir	0	ND	1.9	3.8	1.9	2.58	NA	81	2.9	0.9*	2.0	-4.82
Lanreotide	≠0	ND	-0.3	NA	NA	2.75	NA	NA	NA	NA	NA	ND
Nelfinavir	1+	ND	4.7	6.1	1.4	2.69	NA	101	4.0	3.5	0.5	ND
Paclitaxel	0	ND	3.1	3.4	0.3	3.03	NA	102	NA	NA	NA	-5.92
Rifabutin	1+	ND	4.5	5.1	0.6	2.83	NA	73	5.8	5.7	0.1	ND
Rifampicin	1-/1+	1.3	NA	NA	NA	2.58	2.41	120	1.3**	1.1**	0.2**	-6
Rifapentine	1-/1+	ND	NA	NA	NA	2.4	NA	125	NA	NA	NA	ND
Rifaximin	0	ND	3.0	4.2	1.2	2.29	NA	109	3.1	2.9	0.2	-5.7
Ritonavir	0	4.6	3.3	4.9	1.6	3.02	-0.18	80	NA	NA	NA	-3.63
Roxithromycin	1+	1.8	2.0	ND	NA	2	1.37	75	3.5	2.8	0.7	-4.92
Saquinavir	1+	4.7	3.6	5.9	2.3	3.44	0.15	99	4.1	>2.5	NA	ND
Sirolimus	0	ND	5.0	6.4	1.4	3.96	NA	80	NA	NA	NA	ND
Somatostatin 14	≠0	ND	NA	NA	NA	2.1	NA	NA	NA	NA	NA	ND
Telaprevir	0	3.8	3.5	4.4	0.9	3.02	0.55	67	>2.5	>2.5	NA	-5.15
Telithromycin	1+	2.1	0.8	1.6	0.8	3.12	2.22	90	3.2	2.2*	1.0	-5.37
Temsirolimus	0	ND	4.9	6.9	2	3.78	NA	92	NA	NA	NA	ND
Vasopressin	0	ND	NA	NA	NA	1.55	NA	143	NA	NA	NA	ND
Vinblastine	1+	ND	3.1	6.9	3.8	2.69	NA	90	3.1	3.2	0.0	-4.92

ND = not determined; NA = not experimentally accessible (more details are available in the text).
* Validated by shake-flask

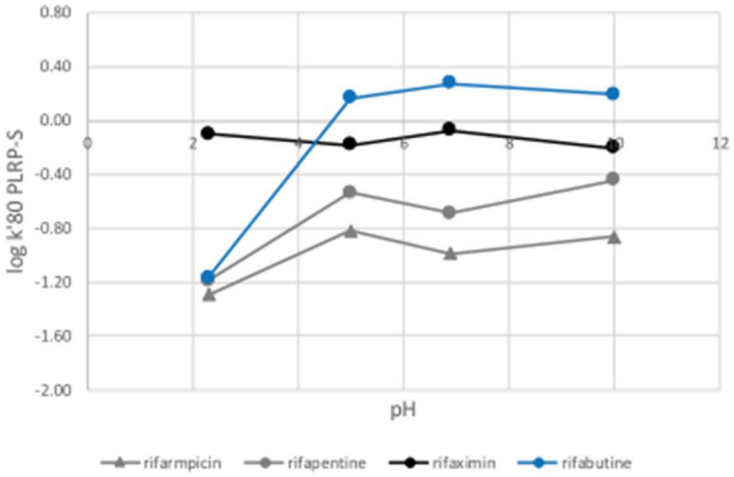

Fig. 1. The log *k*'80 PLRP-S descriptor is designed to experimentally determine the ionization properties of a series of rifampicin. Rifampicin and rifapentine showed a similar ionization profile (gray: zwitterionic), very different from rifaximin (black: neutral) and rifabutin (blue: basic). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

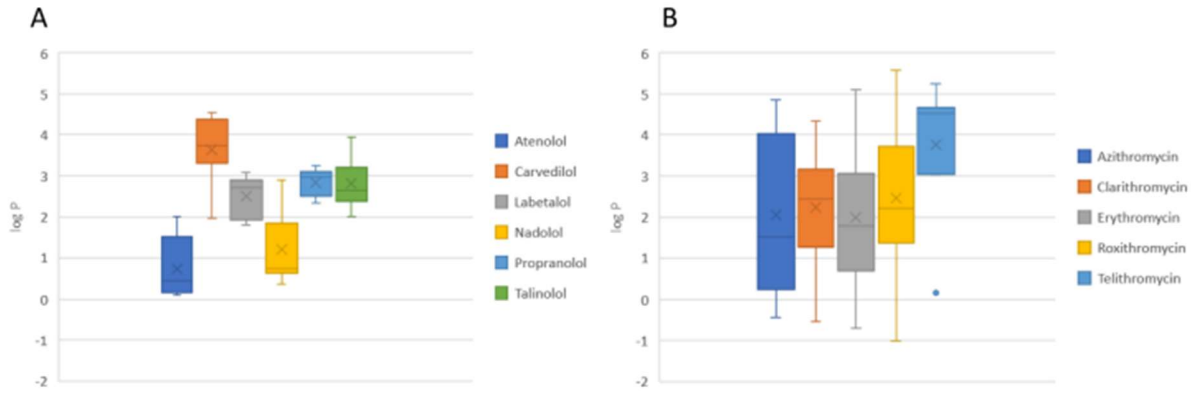

Fig. 2. Eight log *P* calculators (MoKa, ACD, Marvin, *MlogP, Xlogp3, IlogP, WlogP*, Silicos-IT) were used to verify log *P* distribution in two classes of drugs: (A) beta-blockers (Ro5 compliant) and (B) eythronolides (beyond Ro5).

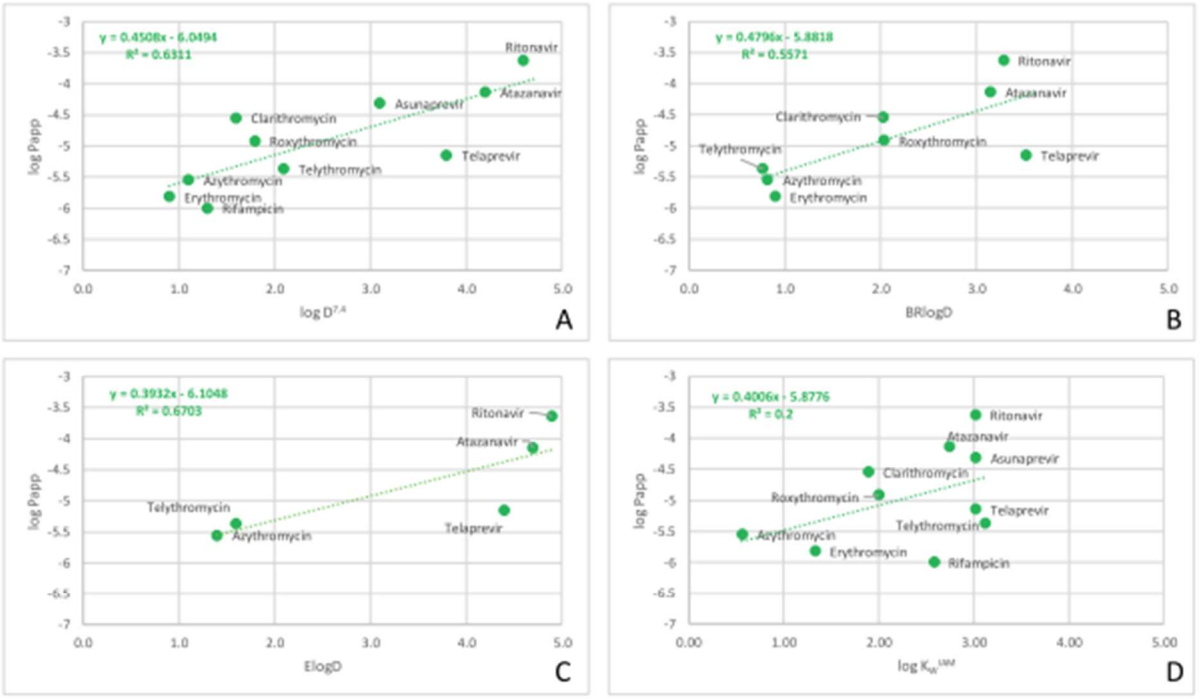

Fig. 3. Permeability vs lipophilicity descriptors: (A) log *D7.4*, (B) *BRlogD*, (C) *ElogD*, (D) log K_W^{IAM}.

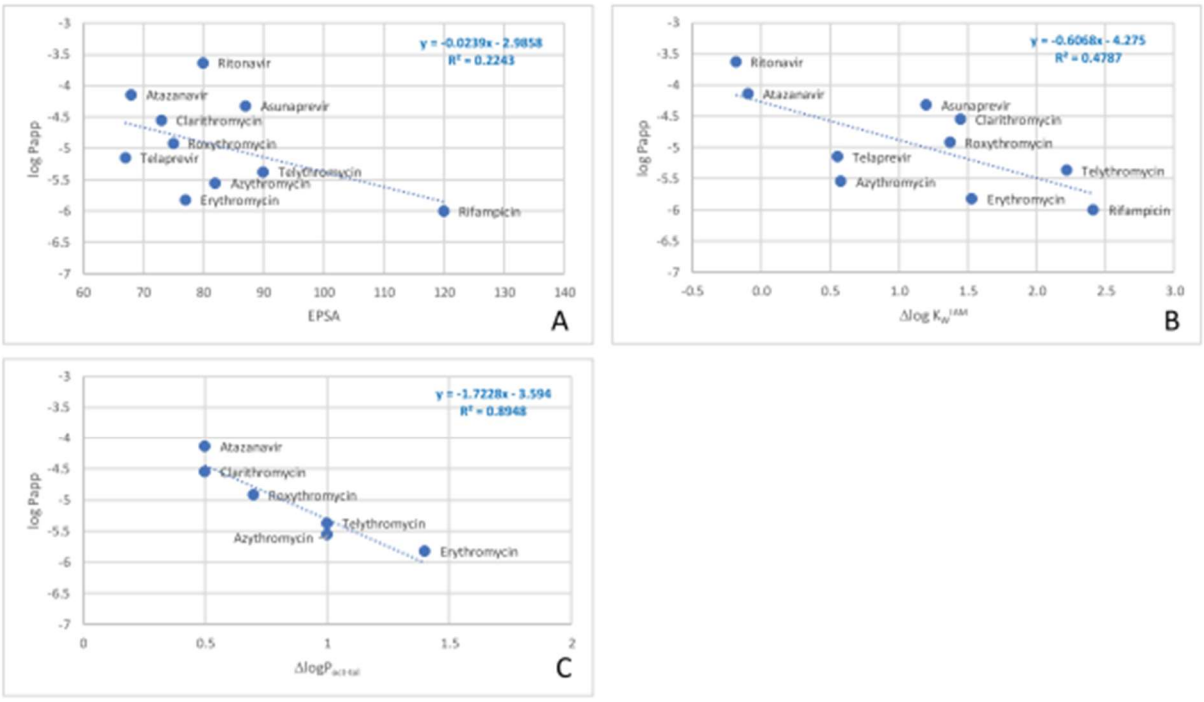

Fig. 4. Permeability vs polarity descriptors. (A) EPSA, (B) Δlog K_W^{IAM}, (C) Δlog $P_{oct-sol}$.

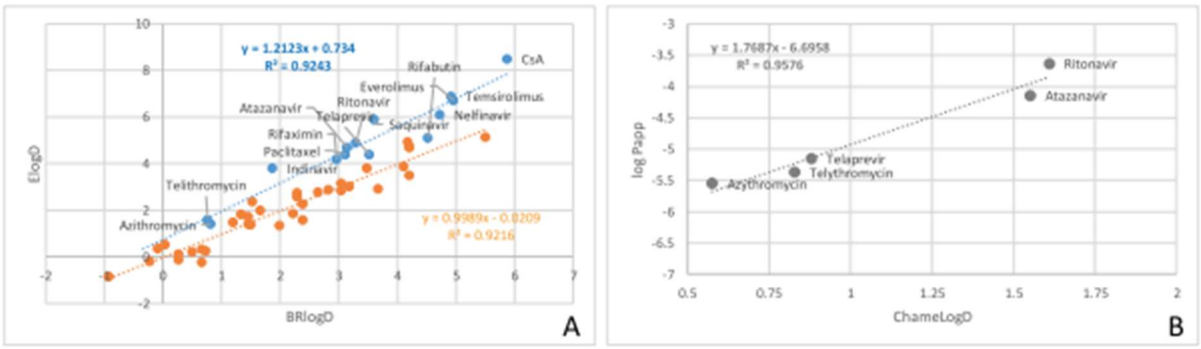

Fig. 6. (A) Relationship between *ElogD* and *BRlogD* (Ro5(Ermondi et al., 2019): orange dots; bRo5; blue dots) and (B) relationship between permeability and ChameLogD. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)