

The logo for iris AperTO, featuring the text "iris" in a stylized font and "AperTO" in a serif font, all in white on a red rectangular background.

UNIVERSITÀ
DEGLI STUDI
DI TORINO

This Accepted Author Manuscript (AAM) is copyrighted and published by Elsevier. It is posted here by agreement between Elsevier and the University of Turin. Changes resulting from the publishing process - such as editing, corrections, structural formatting, and other quality control mechanisms - may not be reflected in this version of the text. The definitive version of the text was subsequently published in INTERNATIONAL JOURNAL OF PHARMACEUTICS, 508 (1-2), 2016, 10.1016/j.ijpharm.2016.05.018.

You may download, copy and otherwise use the AAM for non-commercial purposes provided that your license is limited by the following restrictions:

- (1) You may use this AAM for non-commercial purposes only under the terms of the CC-BY-NC-ND license.
- (2) The integrity of the work and identification of the author, copyright owner, and publisher must be preserved in any copy.
- (3) You must attribute this AAM in the following format: Creative Commons BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/deed.en>), 10.1016/j.ijpharm.2016.05.018

The publisher's version is available at:

<http://linkinghub.elsevier.com/retrieve/pii/S037851731630388X>

When citing, please refer to the published version.

Link to this full text:

<http://hdl.handle.net/2318/1571841>

This full text was downloaded from iris - AperTO: <https://iris.unito.it/>

iris - AperTO

University of Turin's Institutional Research Information System and Open Access Institutional Repository

Do surface-based match solution-based techniques? The case of drug-liposome interaction

Giuseppe Ermondi and Giulia Caron*

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

E-mail: giulia.caron@unito.it

Telephone: +39 011 6708337

Reviewer: Myszka, Puglisi, Franco (bari, Sonja), Musumeci, Comer

Abstract

The aim of the study is to check if the information about drug/liposome interactions provided by SPR is comparable with that provided by methods (here potentiometry) in which liposomes are not immobilized on a solid support. To reach our aim we apply QSPR and BR analysis to data extracted from the literature and carefully inspected for their reliability. Results show that $\log K_D$ (SPR) is governed by a different balance of intermolecular interactions than $\log D_{lip}$ (potentiometry).

Introduction

Interactions of drugs and biological compounds with biomembranes are complex phenomena of paramount importance in both drug discovery and drug delivery. (van Balen et al., 2004) (Pignatello et al., 2011). The understanding of drug membrane interaction is crucial both from a pharmacodynamic (PD) and a pharmacokinetic (PK) point of view. Firstly, drug membrane interactions govern drug interactions with membrane-bound transporters, metabolizing enzymes and receptors, which have the binding sites located in the bilayer (Lukacova et al., 2013). Secondly, high and intermediate rates of trans-bilayer transport are responsible for good permeability properties, whereas too strong or too weak interactions lead to poor ADME profiles (Balaz, 2009).

Up to date a number of experimental methods have been proposed to investigate the affinity of drugs for biomembranes or artificial membranes models (e.g. liposomes): potentiometry, dialysis, ultracentrifugation, ultrafiltration, calorimetry, NMR and spectroscopic techniques, (Van Balen et al., 2004) (Lukacova et al., 2013), chromatography (Taillardat-Bertschinger et al., 2003), Surface Plasmon Resonance (SPR) (Abdiche and Myszka, 2004). An exhaustive review of all these methods is beyond the scope of this study. Notably, no computational method is available today to predict drug/liposome interaction.

The potentiometric technique (Avdeef et al., 1998) was shown to yield satisfactory estimates of lipophilicity in the liposome/water system by independent researchers (Escher, 2000) (van Balen et al., 2004). Moreover, once optimized the experimental conditions, it is faster than the reference method, i.e. equilibrium dialysis.

Since SPR is emerging as an informative medium-throughput technology for hit validation (Patching, 2014) a method to detect drug/liposome interactions based on this technology deserves peculiar attention. The conventional SPR technique requires one binding component to be immobilised on a sensor chip whilst the other binding component in solution is flowed over the sensor surface; a binding interaction is detected using an optical method that measures small changes in refractive index at the sensor surface (Patching, 2014). To measure drug/liposome interactions, liposomes are attached to a sensor surface, the drug is flowed over the sensor surface and the interactions between drugs and liposomes are monitored (Danelian et al., 2000). A few papers based on selected compounds demonstrate that when the SPR biosensor experiments are performed with care, the equilibrium, thermodynamic, and kinetic constants determined from this surface-based technique match those acquired in solution (Rich et al., 2001)(Day et al., 2002)(Swanenburg et al., 2005). However, this match is strongly dependent on the nature of the immobilized receptor and cannot be generalized.

To our knowledge, no relationship between drug/liposome interactions determined by SPR and liposome/water distribution coefficients is reported in the literature. SPR dissociation data were in fact only compared with lipid retention measurements obtained from parallel artificial membranes permeability assays (PAMPA) (Abdiche and Myszka, 2004).

To fill this gap, in this study we deconvolute the balance of the intermolecular forces governing a) the logarithm of the apparent binding affinities for drug interactions with liposome surfaces ($\log K_D$) and b) the logarithm of the liposomes/water distribution coefficients ($\log D_{lip}$) at pH 7.0.

To do that we apply a computational approach (named BR analysis) developed by us in 2012 (Ermondi and Caron, 2012). BR analysis allows the analysis of the balance of intermolecular interactions governing a given system using common 3D-QSAR/QSPR descriptors. These descriptors are aggregated into property-related groups (blocks), thus providing a convenient framework for comparison and interpretation of descriptors determined in different systems (Caron et al., 2013)(Ermondi et al., 2014)(Potter et al., 2014)(Caron et al., 2015)(Caron et al., 2016).

Materials and Methods

The datasets

Experimental values of drug/liposomes interactions were taken from the literature as described below.

Dataset 1 refers to SPR data. The interaction with liposomes at pH 7.4 determined via SPR were expressed as the logarithm of the apparent binding affinities for drug interactions with liposome surfaces ($\log K_D$, see the original paper for details about K_D determination) (Abdiche and Myszka, 2004). K_D values were obtained

from the histogram reported in Fig. 1 of the original paper and thus weak binders were not included in the study. The length of the bars was measured with a ruler and then converted in numerical values. The conversion was validated by comparing values cited in the original paper and values obtained by our conversion tool (for example tamoxifen: paper $K_D=20$, our value $K_D=20.14$; dibucaine: paper $K_D=163$, our value $K_D=163.35$). According to the definition, the lower K_D , the more bound the drug. The final dataset consists of 41 drugs. The drug panel was analyzed against 1,2-dioleoyl-sn-glycero-3-phosphocholine (DOPC) liposomes that were immobilized on Series S Sensor Chip L1.

Dataset 2 refers to the logarithm of the liposomes/water distribution coefficients ($\log D_{lip}$) at pH 7.0. Data were taken from four different papers which report $\log D_{lip}$ using a similar potentiometric equipment and method. Most compounds (16) were taken from the study by Balon and coworkers (Balon et al., 1999). Rifabutin and paromomycin were discarded since potentiometry has some known limitations in the determination of $\log P$ of multiprotic substances and zwitterions. Eight compounds were extracted from the Avdeef's paper (Avdeef et al., 1998). Four small organic molecules lipophilicity data were taken from the study by Escher (Escher, 2000) but nitro compounds were discarded since they need a peculiar computational treatment in VS+ which was beyond the scope of the study. Finally nine drugs were extracted by the paper of Taillardat-Bertschinger et al. (Taillardat-Bertschinger et al., 2002). All data refer to PhC liposomes. Propranolol was reported in 3 out of 4 papers and all $\log D$ values were very similar. When more than one value was present for the same compounds the Avdeef's value was chosen.

If needed $\log D$ was calculated from $\log P^N$ and $\log P^I$ using the following equations

$$D = P^N * \left(\frac{1}{1+10^{pKa-pH}} \right) + P^I * \left(\frac{10^{pKa-pH}}{1+10^{pKa-pH}} \right) \text{ for bases}$$

$$D = P^N * \left(\frac{1}{1+10^{pH-pKa}} \right) + P^I * \left(\frac{10^{pH-pKa}}{1+10^{pH-pKa}} \right) \text{ for acids}$$

The SMILES codes of the investigated compounds (dataset 1 and 2) can be downloaded from www.cassmedchem.unito.it

PLS models

SMILES codes were obtained using the ChemCell plugin (2010 Collaborative Drug Discovery, Inc.) for Excel and, then, submitted to VS+ (version 1.0.7, 198 <http://www.moldiscovery.com>) using default settings (protonation normalized at pH =7) and four probes (OH2, DRY, N1, and O probes that mimic, respectively, water, hydrophobic, HBA, and HBD properties of the environment). The VS+ matrix of descriptors was exported and then submitted to PLS analysis using SIMCA13 (SIMCA13, ver. 13.0.3.0; Umetrics, Umea, Sweden).

BR analysis

BR analysis was performed as described elsewhere (Caron et al., 2013)(Ermondi et al., 2014)(Potter et al., 2014)(Caron et al., 2015)(Caron et al., 2016).

Processing was done on a two 8 cores Xeon E5 at 3.3GHz CPUs and 128GB of RAM.

Results and Discussion

The first step of the study consisted in retrieving reliable $\log K_D$ and $\log D_{lip}$ values. Actually, there is a paucity of published data about drug/liposome interactions. In particular, we verified that it was not possible to set-up a dataset of compounds for which both values were available. Therefore for $\log K_D$ we focused on the dataset provided by Abdiche and Myszka (Abdiche and Myszka, 2004) (dataset 1, see Materials and Methods), whereas for $\log D_{lip}$ we decided to focus on the potentiometric technique and extract data from four different papers which used the same instrument and similar experimental conditions (Balon et al., 1999), (Avdeef et al., 1998), (Taillardat-bertschinger et al., 2002), (Escher, 2000) (dataset 2, see Materials and Methods).

The two datasets share only five compounds (ibuprophen, imipramine, propranolol, tetracaine and warfarin) and thus we firstly verified that they cover a similar chemical space (Fig. S1 in the Supporting Information).

$\log K_D$ values (38 compounds, Table S1 (Supporting Information), see Materials and Methods for details) were used to build a QSPR model by splitting up the datasets into a training set ($n=28$) and a test set ($n=10$). Since lipophilicity data (see below) were determined at pH 7.0 we verified that the ionization state of compounds did not vary when passing from pH 5.5 to pH 7.0 (data not shown). Experimental $\log K_D$ values were then imported into VS+ as response variables (Y, the property) and a relation between Y and the 82 VS+ molecular descriptors (X, the structure) was sought using the PLS algorithm implemented in the software. Three compounds were identified as outliers (auramine O, octyl gallate and tamoxifen). Tamoxifen and auramine O contain substructures characterized by large delocalization, whereas octyl gallate has a flexible chain. The presence of these structural elements would require a careful reevaluation of the force field empirical parameters on which VolSurf+ descriptors are based. This reevaluation is beyond the scope of this paper and thus the three outliers were discarded from the study. Statistics of the best model are reported in Table 1

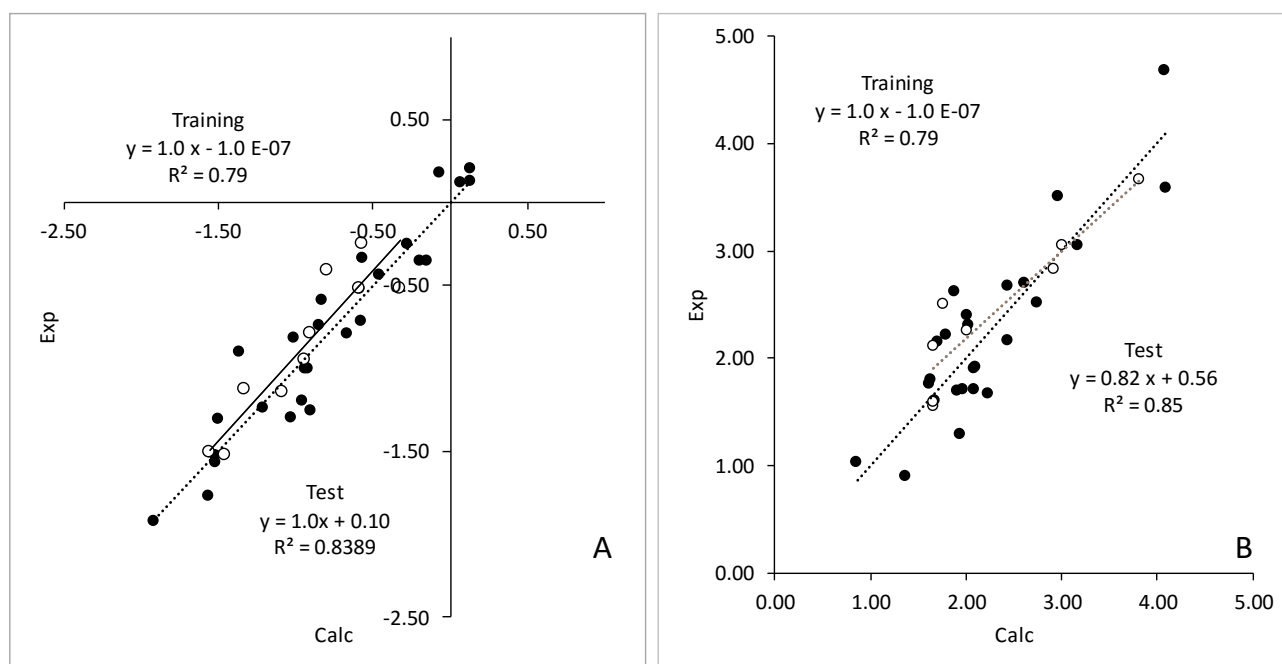
Table 1. PLS analysis results (n = number of observations, R^2 = cumulative determination coefficient, Q^2 = cross-validated correlation coefficient, LV = number of latent variables).

System	N	R^2	Q^{2*}	LV
$\log K_D$	28	0.91	0.52	4
$\log D_{lip}$	25	0.79	0.53	3

* Partial cross-validation adopted by SIMCA

The correlations between calculated and experimental values are shown in Figure 1A. for the training ($R^2 = 0.79$) and the test sets ($R^2 = 0.84$) and support the statistical stability of the model.

Figure 1. Correlation between calculated and experimental values: A) $\log K_D$ and B) $\log D_{lip}$



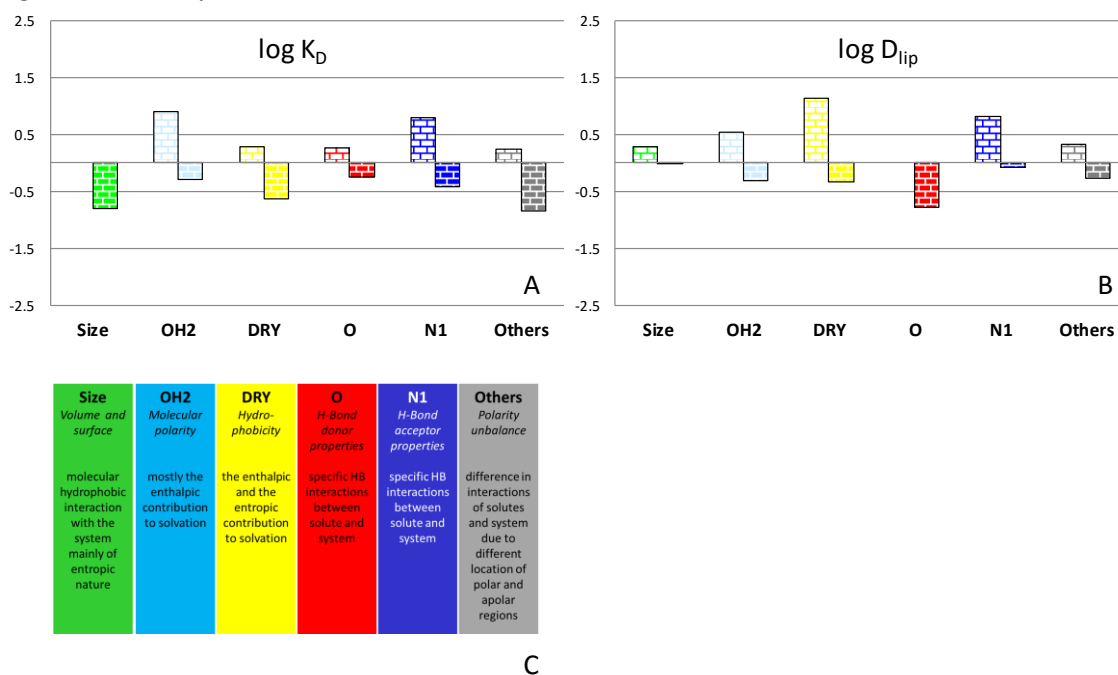
BR analysis graphical output is reported in Figure 2A. The meaning and significance of the blocks is reported in Figure 2C to help with interpretation of parameters. In Fig 2A blocks with positive weighting (e.g. the light blue block) show how much the property described by the block contributes to increase $\log K_D$ and thus to decrease the interaction with liposomes (see Methods for the definition of K_D). Conversely, blocks with negative weighting (e.g. the green block) indicate how much the property contributes to decrease $\log K_D$. Moreover, blocks with comparable positive and negative contributions (e.g. the red block) indicate the high noise and inter-correlation of the descriptors of the block itself and thus are poorly relevant in the description of the investigated phenomenon.

Fig. 2A shows that $\log K_D$ increases, and thus compounds are weakly bound to liposomes, when the solute has hydrogen bonding acceptor (HBA, blue block) groups. Conversely, the presence of hydrophobic moieties favors the interaction with liposomes (green and yellow blocks). The Others group is also relevant in decreasing $\log K_D$. The presence of hydrogen bonding donor groups in the chemical structure (HBD) poorly affects the interactions with liposomes.

The same approach used for modeling $\log K_D$ was applied to $\log D_{lip}$. The full list of $\log D_{lip}$ (33 compounds, see Materials and Methods for details) is reported in Table S2 (Supporting Information). Preliminary analysis revealed that nizatidine and famotidine are outliers. The reasons are probably related to the presence of sulfur atoms that are known to be not properly parametrized by Volsurf+ and thus, as discussed above, we decided to discard the two drugs from the study without additional investigation. 33 compounds were therefore used to build the QSPR model and they were then split into a training set ($n=25$) and a test set ($n=8$). Statistics are reported in Table 1, whereas the correlations between calculated

and experimental values are shown in Figure 1B. Table 1 and Figure 1B support the statistical stability of the model.

Figure 2. BR analysis



BR analysis graphical output is reported in Figure 2B in which positive blocks contribute to increase the interaction with liposomes whereas the reverse is true for negative blocks. Fig. 2B shows that hydrophobicity (yellow block) and HBA (blue block) favor partitioning in liposomes whereas HBD solutes properties (red block) are detrimental for the interaction.

Taken together Fig. 2A and 2B outline that $\log K_D$ is governed by a balance of intermolecular forces very different from that governing $\log D_{lip}$. In particular, the interaction with immobilized liposomes described by $\log K_D$ is mainly driven by hydrophobic interactions, whereas in solution the interaction with liposomes is also favoured by HB between HBA of the drugs and HBD of liposomes (probably the positive phosphatidylcholine headgroups).

We reasoned that the source of this difference could be related to several factors. The first point concerns the major liposome density in SPR than in potentiometric experiments: lipid/drug ratios used in the SPR and potentiometric experiments are in fact very different, being up to 100 times higher in SPR experiments (data not shown). Moreover, liposomes could fuse on the chip surface (Besenicar et al., 2006) and DMSO could play a role in SPR experiments (Abdiche and Myszk, 2004).

Conclusion

Technologies immobilizing receptors on solid supports and measuring drug/receptor interactions are largely appreciated by industrial researchers since they provide high/medium-throughput tools for hit

validation. However, surface-based techniques could provide different information than solution-based tools.

Here we verified how the interaction between drugs and liposomes (= models of biological membranes) is governed by a different balance of intermolecular forces according to the immobilization or not of the liposomes. The study provides an alert message: SPR and potentiometry are not exchangeable to measure drug/liposomes interactions but from their combination more insight in the phenomenon could be obtained. Therefore, it seems convenient to follow the suggestions by some authors (Hämäläinen and Frostell-Karlsson, 2004) and use several approaches to screen early-stage compounds for their properties.

Finally, this result underlined once more the usefulness of a computational tool, named BR analysis, specifically designed to provide an overview of the main factors governing molecular interactions related to ADME properties.

References

- Abdiche, Y.N., Myszka, D.G., 2004. Probing the mechanism of drug/lipid membrane interactions using Biacore. *Anal. Biochem.* 328, 233–43. doi:10.1016/j.ab.2004.01.018
- Avdeef, A., Box, K.J., Comer, J., Hibbert, C., Tam, K.Y., 1998. pH-metric logP 10. Determination of Liposomal Membrane-Water partition coefficients of ionizable compounds. *Pharm. Res* 15, 209–215. Doi: 10.1023/A:1011954332221
- Balaz, S., 2009. Modeling kinetics of subcellular disposition of chemicals. *Chem. Rev.* 109, 1793–1899. doi:10.1021/cr030440j
- Balon, K., Riebesehl, B.U., Müller, B.W., 1999. Drug liposome partitioning as a tool for the prediction of human passive intestinal absorption. *Pharm. Res.* 16, 882–888 doi:10.1023/A:1018882221008
- Besenicar, M., Macek, P., Lakey, J.H., Anderluh, G., 2006. Surface plasmon resonance in protein-membrane interactions. *Chem. Phys. Lipids* 141, 169–78. doi:10.1016/j.chemphyslip.2006.02.010
- Caron, G., Vallaro, M., Ermondi, G., 2013. The Block Relevance (BR) analysis to aid medicinal chemists to determine and interpret lipophilicity. *Medchemcomm* 4, 1376. doi:10.1039/c3md00140g
- 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.* acs.molpharmaceut.5b00910. doi:10.1021/acs.molpharmaceut.5b00910
- Caron, G., Visentin, S., Pontremoli, C., Ermondi, G., 2015. Profile of the intermolecular forces governing the interaction of drugs with mucin. *Int. J. Pharm.* 488, 67–69. doi:10.1016/j.ijpharm.2015.04.058
- Danelian, E., Karlen, A., Karlsson, R., Winiwarter, S., Hansson, A., Lofas, S., Lennernas, H., Hamalainen, M.D. 2000 SPR Biosensor studies of the direct interaction between 27 drugs and a liposome surface: correlation with fraction absorbed in humans. *J.Med. Chem.* 43(11), 2083–2086: doi: 10.1021/jm991156g
- Day, Y.S.N., Baird, C.L., Rich, R.L., Myszka, D.G., 2002. Direct comparison of binding equilibrium, thermodynamic, and rate constants determined by surface- and solution-based biophysical methods. *Protein Sci.* 11, 1017–1025. doi:10.1110/ps.4330102
- Ermondi, G., Caron, G., 2012. Molecular interaction fields based descriptors to interpret and compare chromatographic indexes. *J. Chromatogr. A* 1252, 84–9. doi:10.1016/j.chroma.2012.06.069
- Ermondi, G., Visconti, A., Esposito, R., Caron, G., 2014. The Block Relevance (BR) analysis supports the dominating effect of solutes hydrogen bond acidity on $\Delta\log P(\text{oct-tol})$. *Eur. J. Pharm. Sci.* 53, 50–4. doi:10.1016/j.ejps.2013.12.007
- Escher, B., 2000. Evaluation of liposome-water partitioning of organic acids and bases. 2. Comparison of experimental determination methods. *Environ. Sci.* 34, 3962–3968. doi:10.1021/es0010711
- Esteves, F., Moutinho, C., Matos, C., 2013. Correlation between octanol/water and liposome/water distribution coefficients and drug absorption of a set of pharmacologically active compounds. *J. Liposome Res.* 23, 83–93. doi:10.3109/08982104.2012.742539

- Hämäläinen, M.D., Frostell-Karlsson, A., 2004. Predicting the intestinal absorption potential of hits and leads. *Drug Discov. Today Technol.* 1, 397–405. doi:10.1016/j.ddtec.2004.09.004
- Lukacova, V., Natesan, S., Peng, M., Tandlich, R., Wang, Z., Lynch, S., Subramaniam, R., Balaz, S., 2013. Structural Determinants of Drug Partitioning in Surrogates of Phosphatidylcholine Bilayer Strata. *Mol. Pharm.* 10, 3684–3696. doi: 10.1021/mp400204y
- Patching, S.G., 2014. Surface plasmon resonance spectroscopy for characterisation of membrane protein-ligand interactions and its potential for drug discovery. *Biochim. Biophys. Acta* 1838, 43–55. doi:10.1016/j.bbamem.2013.04.028
- Pignatello, R., Musumeci, T., Basile, L., Carbone, C., Puglisi, G., 2011. Biomembrane models and drug-biomembrane interaction studies: Involvement in drug design and development. *J. Pharm. Bioallied Sci.* 3, 4–14. doi:10.4103/0975-7406.76461
- Potter, T., Ermondi, G., Newbury, G., Caron, G., 2014. Relating Caco-2 permeability to molecular properties using block relevance analysis. *Med. Chem. Commun.* 6, 626–629. doi:10.1039/C4MD00470A
- Rich, R.L., Day, Y.S.N., Morton, T.A., Myszka, D.G., 2001. High-Resolution and High-Throughput Protocols for Measuring Drug/Human Serum Albumin Interactions Using BIACORE. *Anal. Biochem.* 296, 197–207. doi:10.1006/abio.2001.5314
- Swanenburg, M., Bloemraad, R., Maassen, K., 2005. Validation of a surface plasmon resonance based assay to detect Salmonella antibodies in serum of pigs. In *Proceedings of the 6th International Symposium on the Epidemiology and Control of Foodborne Pathogens in Pork*, California, 6-9 September 2005. - California : - p. 100 - 102.
- Taillardat-Bertschinger, A., Carrupt, P.-A., Barbato, F., Testa, B., 2003. Immobilized artificial membrane HPLC in drug research. *J. Med. Chem.* 46, 655–65. doi:10.1021/jm020265j
- Taillardat-Bertschinger, A., Martinet, C.A.M., Carrupt, P., Reist, M., Caron, G., Fruttero, R., Testa, B., 2002. Retention on Immobilized Artificial Membranes (IAM) Compared to Partitioning in Liposomes and n-Octanol. *Pharm. Res* 19, 729-737. doi: 10.1023/A:1016156927420
- van Balen, G.P., Martinet, C. a M., Caron, G., Bouchard, G., Reist, M., Carrupt, P.-A., Fruttero, R., Gasco, A., Testa, B., 2004. Liposome/water lipophilicity: methods, information content, and pharmaceutical applications. *Med. Res. Rev.* 24, 299–324. doi:10.1002/med.10063

Supplementary Material

Table S1: SPR data (dataset 1) as extracted by the paper by Abdiche et al. (Abdiche and Myszk, 2004)

Name	Training/test set	K _D (μM)	log K _D
Alprenolol	training	444.162	-0.352
Amitriptyline	training	125.418	-0.902
Auramine O	training	516.553	-0.287
Chlorpromazine	test	29.323	-1.533
Clomipramine	training	48.812	-1.311
Cyproterone	training	100.000	-1.000
Dansylamide	test	71.197	-1.148
Dapsone	training	452.624	-0.344
Desipramine	test	73.935	-1.131
Dexamethasone	training	256.943	-0.590
Dibucaine	test	163.348	-0.787
Ethoxzolamide	training	163.348	-0.787
Fenoterol	training	567.673	-0.246
Flutamide	training	182.934	-0.738
Formoterol	training	444.162	-0.352
Genistein	training	16.963	-1.770
Homochlorcyclizine	training	54.664	-1.262
Ibuprofen	training	1327.244	0.123
Imipramine	training	100.000	-1.000
Indapamide	training	1602.943	0.205
Mesulergine	training	293.233	-0.533
Nafoxidine	training	27.191	-1.566
Naftopidil	training	63.574	-1.197
NDGA	training	49.742	-1.303
Norclomipramine	test	31.032	-1.508
Nortriptyline	training	57.849	-1.238
Octyl gallate	training	60.074	-1.221
Perphenazine	training	29.323	-1.533
Phenolphthalin	training	1514.704	0.180
Propranolol	training	189.972	-0.721
Provitamin D3	test	111.990	-0.951
Pyrimethamine	test	304.514	-0.516
Quinine	test	389.192	-0.410
Raloxifene	training	12.077	-1.918
Spironolactone	training	360.892	-0.443
Sulfadimethoxine	training	1352.532	0.131
Tamoxifen	training	20.104	-1.697
Tetracaine	test	557.060	-0.254
Thioridazine	training	17.951	-1.746
Verapamil	training	151.470	-0.820
Warfarin	test	298.820	-0.525

Table S2. Dataset 2: log D_{lip} determined using the potentiometric approach

Drug	Training/test set	log D _{lip}	Source
1d	test	1.55	(Taillardat-bertschinger et al., 2002)
2,4,6-trimethylaniline	training	2.52	(Escher, 2000)
3,4-dimethylaniline	training	2.16	(Escher, 2000)
Atenolol	training	1.03	(Balon et al., 1999)
Clonidine	training	1.29	(Taillardat-bertschinger et al., 2002)
Diazepam	training	3.58	(Taillardat-bertschinger et al., 2002)
Famotidine	training	2.20	(Balon et al., 1999)
Fluoxetine	training	2.21	(Balon et al., 1999)
Imipramine	test	2.83	(Taillardat-bertschinger et al., 2002)
Lidocaine	test	1.59	(Avdeef et al., 1998)
Miconazole	test	3.66	(Balon et al., 1999)
Nicotine	training	2.30	(Taillardat-bertschinger et al., 2002)
Nizatadine	training	2.94	(Balon et al., 1999)
Olanzapine	training	3.05	(Balon et al., 1999)
Procaine	training	0.90	(Avdeef et al., 1998)
Propranolol	training	2.62	(Avdeef et al., 1998)
Rilmenidine	test	2.11	(Taillardat-bertschinger et al., 2002)
Terbinafine	training	4.68	(Balon et al., 1999)
Tetracaine	test	2.25	(Avdeef et al., 1998)
Zidovudine	training	2.40	(Balon et al., 1999)
Zopiclone	training	1.69	(Balon et al., 1999)
2,3,4,6-tetrachlorophenol	training	3.51	(Escher, 2000)
2,4,6-trichlorophenol	training	2.70	(Escher, 2000)
5-Phenylvaleric acid	training	1.71	(Avdeef et al., 1998)
Acetylsalicylic acid	training	1.60	(Balon et al., 1999)
Acyclovir	training	1.70	(Balon et al., 1999)
Allopurinol	test	2.50	(Balon et al., 1999)
Amiloride	training	1.79	(Balon et al., 1999)
Diclofenac	training	2.67	(Avdeef et al., 1998)
Furosemide	training	1.90	(Balon et al., 1999)
Ibuprofen	training	1.91	(Balon et al., 1999)
Phenobarbital	training	2.15	(Taillardat-bertschinger et al., 2002)
Phenytoin	test	3.05	(Taillardat-bertschinger et al., 2002)
Warfarin	training	1.67	(Avdeef et al., 1998)
Xipamide	training	1.76	(Balon et al., 1999)

Figure S1. The chemical space covered by dataset 1 and dataset 2 as described by the distribution of some common molecular descriptors (MW, HBA, HBD, log P, rotatable bonds)

