**Diukendjieva, A, Alov, P, Pencheva, T, Tsakovska, I, Richarz, A, Kren, V, Cronin, MTD and Pajeva, I**

# In vitro and in silico studies of the membrane permeability of natural flavonoids from Silybum marianum (L.) Gaertn. and their derivatives

http://researchonline.ljmu.ac.uk/id/eprint/9137/

**Article**

**Citation** (please note it is advisable to refer to the publisher's version if you intend to cite from this work)

**Diukendjieva, A, Alov, P, Pencheva, T, Tsakovska, I, Richarz, A, Kren, V, Cronin, MTD and Pajeva, I (2018) In vitro and in silico studies of the membrane permeability of natural flavonoids from Silybum marianum (L.) Gaertn. and their derivatives. Phytomedicine, 53. pp. 79-85. ISSN 0944-7113**

# *In vitro* and *in silico* studies of the membrane permeability of natural flavonoids from *Silybum marianum* (L.) Gaertn. and their derivatives

Antonia Diukendjieva[a], Petko Alov[a], Tania Pencheva[a], Ivanka Tsakovska[a], Andrea Richarz[b], Vladimir Kren[c], Mark T.D. Cronin[b], Ilza Pajeva[a] *

[a] Institute of Biophysics and Biomedical Engineering, Bulgarian Academy of Sciences, Acad. G. Bonchev Str., Block 105, 1113 Sofia, Bulgaria

[b] School of Pharmacy and Biomolecular Sciences, Liverpool John Moores University, Byrom Street, Liverpool, L3 3AF, England

[c] Laboratory of Biotransformation, Institute of Microbiology, Czech Academy of Sciences, Videnska 1083, CZ 14220 Prague, Czech Republic

* Corresponding author

Ilza Pajeva, Institute of Biophysics and Biomedical Engineering, Bulgarian Academy of Sciences, Acad. G. Bonchev Str., Block 105, 1113 Sofia, Bulgaria

Tel.: +359 2 979 3605

E-mail address: pajeva@biomed.bas.bg

## Abstract

*Background*: In recent years the number of natural products used as pharmaceuticals, components of dietary supplements and cosmetics has increased tremendously requiring more extensive evaluation of their pharmacokinetic properties.

*Purpose*: This study aims at combining *in vitro* and *in silico* methods to evaluate the gastrointestinal absorption (GIA) of natural flavonolignans from milk thistle (*Silybum marianum* (L.) Gaertn.) and their derivatives.

*Methods:* A parallel artificial membrane permeability assay (PAMPA) was used to evaluate the transcellular permeability of the plant main components. A dataset of 269 compounds with measured PAMPA values and specialized software tools for calculating molecular descriptors were utilized to develop a quantitative structure-activity relationship (QSAR) model to predict PAMPA permeability.

*Results*: The PAMPA permeabilities of 7 compounds constituting the main components of the milk thistle were measured and their GIA was evaluated. A freely-available and easy to use QSAR model predicting PAMPA permeability from calculated physico-chemical molecular descriptors was derived and validated on an external dataset of 783 compounds with known GIA. The predicted permeability values correlated well with obtained *in vitro* results. The QSAR model was further applied to predict the GIA of 31 experimentally untested flavonolignans.

*Conclusions:* According to both *in vitro* and *in silico* results most flavonolignans are highly permeable in the gastrointestinal tract, which is a prerequisite for sufficient bioavailability and use as lead structures in drug development. The combined *in vitro/in silico* approach can be used for the preliminary evaluation of GIA and to guide further laboratory experiments on pharmacokinetic characterization of bioactive compounds, including natural products.

## Keywords

PAMPA, QSAR, gastrointestinal absorption, *Silybum marianum,* flavonolignans.

## Abbreviations

ABL – aqueous boundary layer; AP – sum of atomic polarizations; DS – double sink; F – F-ratio; GIA – gastrointestinal absorption; LOO $q^2$ – leave-one-out cross-validation correlation coefficient; MW – molecular weight; NP – natural product; PAMPA – parallel artificial membrane permeability assay; PSA – polar molecular surface area; QSAR – quantitative structure-activity relationship; $r^2$ – multiple correlation coefficient; SEE – standard error of estimate; TPSA – topological polar surface area; TSA – total surface area; VABC – sum of atomic and bond contributions volume.

## Introduction

In recent years the number of natural products (NPs) used as pharmaceuticals, components of dietary supplements and cosmetics has increased tremendously. In particular, there is strong interest in research on flavonoids from plant sources due to their potential health benefits as reported from various epidemiological studies (Kumar and Pandey, 2013). Flavonoids have been shown to exhibit antioxidant (Chen et al., 2018), antidiabetic (Xiao and Hogger, 2014), hypocholesterolaemiac (Thilakarathna et al., 2012), antiplatelet (Khan et al., 2018), antibacterial (Xiao, 2015) and antiinflammatory effects (Chen et al., 2017) as well as the ability to modulate cell signaling and gene expression (Noll et al., 2009) related to infectious and cardiovascular diseases and different forms of cancer (Sak, 2014). Their low toxicity in general is considered a further major advantage of these compounds. However, most bioactivities of flavonoids have been reported from *in vitro* cell experiments, whereas the poor systemic bioavailability may limit their beneficial effects *in vivo* (Xiao and Högger, 2015; Xiao, 2018). Phase 2 metabolism is known to affect the bioavailability of flavonoids and, in general, metabolites of flavonoids show reduced bioactivity in comparison to parent compounds (Thilakarathna and Rupasinghe, 2013). Thus, bioavailability is an important pharmacokinetic property and should be considered as early as possible when NPs and their derivatives are considered for medicinal and drug discovery purposes.

Among the flavonoids, flavonolignans are a relatively small subclass of compounds where the flavonoid part of the molecule is attached to a lignan (Biedermann et al., 2014). Flavonolignans were originally discovered in the seeds of milk thistle (*Silybum marianum* (L.) Gaertn.), a medicinal plant used from ancient times for the treatment of liver and gallbladder disorders of different etiologies. The herb active component, silymarin, is a mixture of flavonolignans, mainly silybin A and silybin B; other phenolic compounds such as isosilybin, dehydrosilybin, silychristin, silydianin and taxifolin are also found in its fruit and seeds (Chambers et al., 2017; Pyszková et al., 2016). Silybin, as the major flavonolignan component of silymarin, is present as a quasi-equimolar mixture of the two diastereomers A and B (natural racemic silybin is noted below as silybin AB). Nowadays, silymarin is the best known for its antioxidant and chemoprotective effects on the liver (Křen and Walterová, 2005) and is often prescribed or self-prescribed as a complementary hepatoprotective medicine (Testino et al., 2013). Furthermore, its use has been broadened to other organs in addition to the liver, e.g. in the treatment of pancreatic diseases and balancing glycaemia, lung and kidney diseases, in dermatological and cosmetic preparations. Other beneficial effects include hypocholesterolaemic, cardioprotective, neuroactive and neuroprotective properties (Křen and Walterová, 2005). Despite the frequent therapeutic use of silybin and its congeners, many of their pharmacokinetic properties affecting bioavailability, including gastrointestinal absorption (GIA), have not been well investigated.

The aim of this study was to address this paucity of pharmacokinetic information by combining *in vitro* and *in silico* methods to evaluate the gastrointestinal absorption of natural flavonoids from milk thistle (*Silybum marianum* (L.) Gaertn.) and their derivatives with a particular focus on flavonolignans. The GIA of several flavonolignans was estimated using the parallel artificial membrane permeability assay (PAMPA). The PAMPA is an *in vitro* model of passive, transcellular permeation. It was introduced by Kansy et al. (1998) to predict oral absorption in a simple, reproducible and high-throughput manner. PAMPA is particularly advantageous in early stages of the drug discovery process. It is cost-effective and easy to automate, additionally it has proved to have good reproducibility and small variability. PAMPA permeability correlates well with GIA *in vivo* and it is now considered to be a good screening system to evaluate the permeability by the passive transcellular route (Ano et al., 2004; Verma et al., 2007). In combination with a high-throughput solubility assay it enables biopharmaceutical classification in the early drug discovery stage. Data from PAMPA have been subject to numerous quantitative structure-activity relationship

(QSAR) studies (Nakao et al., 2009; Leung et al., 2012). Here we report on an *in silico* evaluation of GIA for a broader set of silybin congeners using a QSAR model for the prediction of PAMPA permeability. The model was intentionally developed using descriptors calculated from open-source or free software tools or obtainable from free online resources (Cronin et al., 2012) and is freely available in the COSMOS KNIME WebPortal (http://knimewebportal.cosmostox.eu). It has also been included in the DataBase service on Alternative Methods of the European Union Reference Laboratory for alternatives to animal testing (https://ecvam-dbalm.jrc.ec.europa.eu). Whilst there is variability, the results of the analysis suggest that most of the flavonolignans studied may be considered as being highly permeable in the gastrointestinal tract, implying their potential good bioavailability and appropriateness for using as medicines and lead structures for drug development.

## Materials and methods

### ***Chemicals***

Seven compounds (Fig. 1), provided by the Laboratory of Biotransformation, Institute of Microbiology, Czech Academy of Sciences were investigated *in vitro:* silybin AB (Biedermann et al., 2014), isosilybin A (Gažák et al., 2013a), silychristin A, silydianin (Křenek et al., 2014),  2-3-dehydrosilybin AB (Gažák et al., 2013b), taxifolin and quercetin. This set of compounds was selected empirically to allow analysis of the structural features and physico-chemical properties that can influence permeability. Purity of the flavonolignans was above 96% (HPLC/PDA) and of taxifolin and quercetin above 99% (Sigma-Aldrich).

In addition, the membrane permeability of another 31 silybin derivatives (Džubák et al., 2006; Gažák et al., 2009, 2011; Kosina et al., 2002) were predicted *in silico* (see Supplementary Table 1 for their structures and SMILES codes).

### ***PAMPA***

Double-Sink$^{TM}$ (DS) PAMPA (Avdeef, 2012) measurements were performed in the PAMPA Explorer Test System from Pion Inc. PAMPA “sandwiches” were formed from a Stirwell$^{TM}$ 96-well donor and acceptor plates with a polyvinylidene difluoride filter bottom, coated with a 20% (w/v) dodecane solution of lecithin (Pion Inc., PN 110669). The initial donor sample concentrations were ca. 20 μM. The acceptor compartment was filled with a surfactant-containing buffer at pH 7.4 (Pion Inc., PN 110139); the donor compartment contained buffers at pH 5.0, 6.2, and 7.4 (Pion Inc., PN 110238). The sandwiches were incubated in a water vapor-saturated atmosphere at room temperature for 4 h in the Gut-Box$^{TM}$ module with stirring to adjust the thickness of the aqueous boundary layer (ABL) to 60 μm.

Sample concentrations in acceptor and donor wells were determined by UV spectrophotometry with an Epoch plate reader instrument (BioTek Inc). The effective permeability coefficient, Pe [$cm.s^{-1}$], defined as the number of molecules (mol) diffusing through unit cross-section of the membrane ($cm^2$) per unit of time (s) under a unit of concentration ($mol \cdot cm^{-3}$) gradient, was determined using the PAMPA Explorer software according to Avdeef (2012) (equations A7.28a,b).

Three parallel measurements were made for each sample. Carbamazepine, ketoprofen and ranitidine were used as reference compounds; their measured PAMPA values reproduced those reported in the PAMPA Explorer documentation.

### ***Calculation of the pKa values***

The pKa values of the main components of silymarin were calculated in the ACD/Percepta software, v. 2016.1 (Advanced Chemistry Development, Inc., http://www.acdlabs.com) using the classical algorithm for pKa calculations under standard conditions (25°C and zero ionic strength, aqueous solution) for every ionizable group. Additionally, the pKa values for silybin B, quercetin and taxifolin were calculated using the empirical and quantum-chemical pKa prediction modules in the Schrodinger software, release 2016-1 (http://www.schrodinger.com).

### QSAR model development

The data to construct the DS PAMPA Pe-predicting QSAR model were obtained from "Database of Double-Sink PAMPA log $P_0$, log $Pm^{6.5}$, and log $Pm^{7.4}$" (Avdeef, 2012). The structural information was collected from the NCI/CADD Chemical Identifier Resolver service and from the NCBI PubChem project. Mixtures, compounds with zero permeability and compounds with permeability measured in the presence of a co-solvent were omitted, thus reducing the initial dataset from 292 to 269 compounds. After geometry optimization of the structures (MOPAC2012, http://openmopac.net), the total and polar water-accessible molecular surface areas were calculated in MOE, v. 2015.10 (MOE, http://www.chemcomp.com). Octanol-water distribution-related molecular descriptors (log D at pH 7.4) were calculated by ACD/Percepta or by the calculator plugins of ChemAxon Marvin v. 14.8.25 (http://chemaxon.com). Molecular size-related descriptors were calculated by the KNIME-integrated Chemistry Development Kit (CDK, v. 1.5.1) and Indigo (v. 1.1.4) nodes. The multiple linear regression models were derived and refined in the KNIME Analytics Platform v. 2.12.2 (http://www.knime.com).

## Results and discussion

### Measurement of PAMPA Permeability

The compounds subjected to PAMPA permeability measurements were selected intentionally based on their plant distribution and structural relations: silybin AB (Biedermann et al., 2014), isosilybin A (Gažák et al., 2013a), silychristin A and silydianin (Křenek et al., 2014) are the main components of *Silybum marianum*; 2-3-dehydrosilybin AB (Gažák et al., 2013b) is an NP derivative but also occurs in silymarin as a minor component – up to 1–2% (Chambers et al., 2017); taxifolin and quercetin are structurally identical to the flavonoid part of silybin and dehydrosilybin, respectively, and can be found in many fruits, vegetables, leaves, and grains.

The logarithms of the effective membrane permeability values (log Pe) of the compounds studied are reported in Table 1. Good agreement is observed between the log Pe values of silybin and quercetin reported by Avdeef (2012) and those measured in the present study: –5.08 vs. –5.25±0.05 for silybin, and –4.77 vs. –5.02±0.07 for quercetin.

According to the high/low-to-moderate log Pe classification threshold of –6 (explained in section QSAR model for PAMPA prediction below) and the analysis of the measured log Pe values, the main active component of *Silybum marianum*, silybin, its 2,3-dehydro-derivative and isosilybin A can be considered to be highly permeable in the gastrointestinal tract. At pH 7.4 taxifolin and quercetin demonstrate a similar permeability profile. Silydianin and silychristin A, the second most abundant flavonolignans (after silybin) have lower log Pe values, suggesting lower absorption in the gastrointestinal tract.

The results demonstrate clear dependence of the permeability of the compounds studied with pH. There is a difference of more than one log unit in log Pe at pH 7.4 between silybin and dehydrosilybin; however there is no significant variation at pH 5.0 and / or 6.2. Conversely, the difference in the permeability values between taxifolin and quercetin is higher at the lower pH (6.2 and 5.0). It may be assumed that dehydrogenation in the flavonoid core increases permeability of the flavonolignans at pH 7.4, but does not affect the permeability of the related flavonoids (quercetin and taxifolin), possibly related to the lignan part that is absent in taxifolin and quercetin. Regarding the influence of isomerism, comparison of the permeability values for silybin and isosilybin shows no significant difference with pH.

Analysis of the pH dependence of permeability of individual compounds shows other significant variations. For silybin, isosilybin A, silychristin A and taxifolin there is a difference of ca. one log unit between log Pe values measured at pH 6.2 and 7.4 (Table 1). However, such a difference was not observed for dehydrosilybin and quercetin. We assumed that these variations may be related to the ionization states of the compounds influencing the ratio between their neutral and ionized forms and thus their permeability. As an indicator of relative ionization, which would affect passive diffusion, the ACD/Percepta pKa values of the compounds were calculated. The lowest calculated acidic pKa values are presented in Fig. 1 and vary between 6.3 and 7.4, implying that at pH 7.4 the proportion of their ionized forms is higher compared to that at pH 6.2 and that should result in a lower permeability of the compounds. However, such a tendency has not been observed. Similar results have been recorded using more sophisticated pKa calculations by the specialized modules in Schrodinger software (data not shown). Three compounds with different profiles of log Pe dependence on pH have been studied: silybin B, quercetin and taxifolin. Again, the observed differences in their log Pe could not be referred to the differences in their pKa values. Thus, the calculated pKa values alone are unlikely to explain the effect of pKa on the pH-dependent log Pe of the studied compounds.

### ***QSAR model for PAMPA prediction***

*In silico* estimation of the GIA of the flavonoids was performed using a QSAR model for the prediction of PAMPA permeability. The model was developed using DS PAMPA data (Avdeef, 2012) obtained under experimental conditions equivalent to the PAMPA measurements performed in this study. The dataset of 269 compounds was characterized by a broad distribution of the Pe values. The sink conditions of DS PAMPA (lowering the active concentration of free permeant in the acceptor compartment) together with the ABL control (40-60 μm ABL achieved by in-well stirring) allowed for elimination of non-linearity of the Pe data across a broad range of lipophilicity.

Molecular descriptors similar to those suggested by Kansy et al. (2001) – the logarithm of the apparent octanol/water distribution coefficient (log D), and the ratio of polar to total molecular surface area (PSA/TSA) – were utilized in the QSAR. Log D values were calculated by ACD/Percepta or calculator plugins of ChemAxon Marvin. These log D estimates are readily available from http://www.chemspider.com (calculated by ACD/Percepta for compounds already included in the ChemSpider database) or from http://chemicalize.com (calculated by ChemAxon tools for any submitted compound).

Substitution of the PSA/TSA ratio was considered to allow for the calculation of all descriptors with freely available software tools. As such PSA was substituted by TPSA (topological polar surface area (Ertl et al., 2000). To find an appropriate structural descriptor to substitute for TSA, polar and total surface areas and their ratio were calculated in MOE for all the compounds in the PAMPA dataset. Sixty-two descriptors related to molecular size were obtained and their relationships with TSA assessed (Table 2A), as were the relationships of TPSA/descriptor ratios to PSA/TSA (Table 2B). Following identification of the top-ranked TPSA/descriptor ratios, they were tested in the development of QSAR models.

In order to increase the QSAR models' stability, high leverage compounds and the response outliers were filtered out. To evaluate the external predictivity of the models the datasets were split into training and test sets (4:1 stratified splitting). The goodness-of-fit ($r^2$, SEE, F) and the internal leave-one-out cross-validation (LOO $q^2$) statistics of the models were very close to those using PSA/TSA (Table 3), thus the substitution of any of the three top-ranked TPSA/descriptor ratios – TPSA/VABC (sum of atomic and bond contributions volume), TPSA/MW (molecular weight) and TPSA/AP (sum of atomic polarizations) for PSA/TSA – is well justified. The very close values of $r^2$ and LOO $q^2$ for all models demonstrate high model stability. The external predictivity coefficients are also in a narrow range (0.69-0.79) and similar to those using PSA/TSA (0.68 and 0.79 for ACD/Percepta and ChemAxon tools calculated log D-based models, respectively). Therefore, the use of descriptors from freely available sources does not decrease the quality of the models and is justified for future use.

Considering that MW is the most fundamental descriptor of the molecular size, and that the statistical parameters of the models using it were among the best, MW was selected to substitute for TSA. The two implementations of the model based on log D at pH 7.4 as estimated by the ACD/Percepta or ChemAxon tools are presented in equations 1 and 2, respectively:

$$\log Pe = -2.20(\pm 0.21) + 0.49(\pm 0.04)\log D - 10.14(\pm 0.74)TPSA/MW \quad (1)$$

$n = 251$, $r^2 = 0.75$, $SEE = 1.10$, $F = 371.3$,

LOO $q^2 = 0.74$, external validation $q^2 = 0.79$ (200/51)

$$\log Pe = -2.11(\pm 0.22) + 0.47(\pm 0.05)\log D - 10.71(\pm 0.78)TPSA/MW \quad (2)$$

$n = 248$, $r^2 = 0.74$, $SEE = 1.11$, $F = 345.1$,

LOO $q^2 = 0.73$, external validation $q^2 = 0.77$ (198/50)

The ability of these models to predict GIA was assessed using an external dataset (accessible at http://biomed.bas.bg/qsarmm) of 783 compounds (1227 distinct values) with reported GIA collected from the literature, 167 of them (383 distinct GIA values) with DS PAMPA Pe in the training set of the model developed. The data collected did not distinguish low and medium GIA, due to the low percentage of compounds with low and medium GIA (Fig. 2A). However, a rapid decrease in the percentage of observations belonging to the highest GIA class (>80%) is evident for compounds with PAMPA log Pe lower than –6 (Fig. 2B), which confirms the recommendation in Avdeef (2012) to use log Pe < –6 as an indication for possible low GIA. The model classified the remaining 616 compounds into high or medium-to-low GIA classes and the accuracy, sensitivity and specificity of the classification were calculated (Table 4).

### *In silico prediction of Pe for the flavonoids*

The results from the *in silico* prediction of PAMPA permeability for the compounds studied *in vitro* using the QSAR model are reported in Table 5. Fig. 3 represents their positions within the space defined by the physico-chemical parameters used for the development of the model for the compounds in the training set. The figure demonstrates that the compounds fall into the applicability domain of the model thus confirming the reliability of the predictions.

The predicted log Pe values of the silybin congeners (silybin AB, 2,3-dehydrosilybin AB and isosilybin A, Table 5) correspond well to the measured PAMPA permeability at pH 7.4 (Table 1). For these compounds, there is a difference of less than one log unit between the measured and calculated permeability values. For silychristin A and silydianin the predicted values are higher than those measured by more than 1.5 log units. Log D and TPSA/MW for these compounds are similar to those of silybin and 2,3-dehydrosilybin, suggesting the presence of specific structural features not accounted for by the model that result in higher than predicted membrane permeability.

Fig. 4 illustrates the plot of experimental log Pe values vs. those calculated by the QSAR model for the flavonoids studied. Among the main components of milk thistle, silybin and its congeners show higher *in vitro* and *in silico* permeability. These findings are in agreement with previously reported *in vivo* data which indicate that silybin is absorbed rapidly in the gastrointestinal tract, although its low solubility and fast elimination remain major concerns with regard to bioavailability (Wu et al., 2009). 2,3-dehydrosilybin AB possesses the highest *in vitro* log Pe and close to that obtained by the QSAR model. The predicted permeabilities of taxifolin and quercetin differ from the experimental values by ca. one log unit and place these compounds close to the high/low permeability threshold.

Based on the good correspondence between the observed and calculated permeability of the silybin congeners (silybin AB, dehydrosilybin AB and isosilyibn A), the permeability of a further 31 silybin derivatives, with structural skeleton similar to those of the studied silybins and unknown permeability, was also predicted (data shown in Table 6 and Supplementary Table 1). As demonstrated in Fig. 5, high GIA can be expected for most of these compounds. Only four flavonolignans (silybinic acid, 2,3-dehydrosilybinic acid, silybin 23-O-β-lactoside and silybin 23-O-β-maltoside) have log Pe values lower than –6. This could be attributed to the presence of highly polar carboxyl groups in the two acids and the bulky polar disaccharide moiety in the two glycosides. The majority of the compounds have log Pe values between –4 and –5, which classifies them as highly permeable. Additional experimental studies are necessary to confirm these predictions.

## Conclusions

In the present study the PAMPA methodology has been applied to estimate the membrane permeability of all major components of *Silybum marianum* (L.) Gaertn. A QSAR model for PAMPA has been developed and combined with the *in vitro* results to predict the GIA of all major components of the milk thistle and their derivatives. The QSAR model uses descriptors calculated by open-source or free software tools or those obtainable from free online resources that makes it appropriate for a broader application. According to both *in vitro* and *in silico* methods most flavonolignans are highly permeable in the gastrointestinal tract, which is a good prerequisite for sufficient bioavailability. The estimated permeability of the studied flavonoids makes them appropriate lead structures for drug development purposes. The results confirm that the combined interdisciplinary approach based on *in silico* QSAR predictions and *in vitro* PAMPA measurements can be used for preliminary evaluation of GIA and can guide further laboratory experiments for characterization of bioactive compounds, including NPs.

## Conflict of interest

The authors declare no competing financial interest.

## Acknowledgments

This work is supported by the “Program for career development of young scientists, BAS” projects DFNP-141/2016 and DFNP-134/2017, the National Science Fund of Bulgaria project DCOST 01/11-2016 and the project from Czech Science Foundation No. 18-00150S. The QSAR model development has been funded by the EC 7th Framework Program and Cosmetics Europe COSMOS project (grant No. 266835). IP and VK acknowledge the networking contribution from the COST Action CM1407 “Challenging organic syntheses inspired by nature – from natural products chemistry to drug discovery”. Authors are grateful to Sofia TechPark for the kindly provided access to PAMPA Explorer equipment.

## Supplementary materials

Structures, molecular structural descriptors, predicted log Pe and GIA permeability estimations of 31 silybin derivatives studied *in silico*.

## References

Ano, R., Kimura, Y., Shima, M., Matsuno, R., Ueno, T., Akamatsu, M., 2004. Relationships between structure and high-throughput screening permeability of peptide derivatives and related compounds with artificial membranes: application to prediction of Caco-2 cell permeability. Bioorg. Med. Chem. 12, 257–264. https://doi.org/10.1016/j.bmc.2003.10.002

Avdeef, A., 2012. Absorption and drug development: solubility, permeability, and charge state, 2nd ed. ed. John Wiley & Sons, Hoboken, N.J.

Biedermann, D., Vavříková, E., Cvak, L., Křen, V., 2014. Chemistry of silybin. Nat. Prod. Rep. 31, 1138. https://doi.org/10.1039/C3NP70122K

Chambers, C.S., Holečková, V., Petrásková, L., Biedermann, D., Valentová, K., Buchta, M., Křen, V., 2017. The silymarin composition… and why does it matter??? Food Res. Int. https://doi.org/10.1016/j.foodres.2017.07.017

Chen, L., Teng, H., Jia, Z., Battino, M., Miron, A., Yu, Z., Cao, H., Xiao, J., 2017. Intracellular signaling pathways of inflammation modulated by dietary flavonoids: The most recent evidence. Crit. Rev. Food Sci. Nutr. 1–17. https://doi.org/10.1080/10408398.2017.1345853

Chen, L., Teng, H., Xie, Z., Cao, H., Cheang, W.S., Skalicka-Woniak, K., Georgiev, M.I., Xiao, J., 2018. Modifications of dietary flavonoids towards improved bioactivity: An update on structure–activity relationship. Crit. Rev. Food Sci. Nutr. 58, 513–527. https://doi.org/10.1080/10408398.2016.1196334

Cronin, M.T.D., Madden, J.C., Richarz, A.-N., 2012. The COSMOS Project: A Foundation for the Future of Computational Modelling of Repeat Dose Toxicity [WWW Document]. URL http://alttox.org/the-cosmos-project-a-foundation-for-the-future-of-computational-modelling-of-repeat-dose-toxicity/ (accessed 8.8.17).

Džubák, P., Hajdúch, M., Gažák, R., Svobodová, A., Psotová, J., Walterová, D., Sedmera, P., Křen, V., 2006. New derivatives of silybin and 2,3-dehydrosilybin and their cytotoxic and P-glycoprotein modulatory activity. Bioorg. Med. Chem. 14, 3793–3810. https://doi.org/10.1016/j.bmc.2006.01.035

Ertl, P., Rohde, B., Selzer, P., 2000. Fast calculation of molecular polar surface area as a sum of fragment-based contributions and its application to the prediction of drug transport properties. J. Med. Chem. 43, 3714–3717. https://doi.org/10.1021/jm000942e

Gažák, R., Fuksová, K., Marhol, P., Kuzma, M., Agarwal, R., Křen, V., 2013a. Preparative method for isosilybin isolation based on enzymatic kinetic resolution of silymarin mixture. Process Biochem. 48, 184–189. https://doi.org/10.1016/j.procbio.2012.11.006

Gažák, R., Sedmera, P., Vrbacký, M., Vostálová, J., Drahota, Z., Marhol, P., Walterová, D., Křen, V., 2009. Molecular mechanisms of silybin and 2,3-dehydrosilybin antiradical activity—role of individual hydroxyl groups. Free Radic. Biol. Med. 46, 745–758. https://doi.org/10.1016/j.freeradbiomed.2008.11.016

Gažák, R., Trouillas, P., Biedermann, D., Fuksová, K., Marhol, P., Kuzma, M., Křen, V., 2013b. Base-catalyzed oxidation of silybin and isosilybin into 2,3-dehydro derivatives. Tetrahedron Lett. 54, 315–317. https://doi.org/10.1016/j.tetlet.2012.11.049

Gažák, R., Valentová, K., Fuksová, K., Marhol, P., Kuzma, M., Medina, M.Á., Oborná, I., Ulrichová, J., Křen, V., 2011. Synthesis and antiangiogenic activity of new silybin galloyl esters. J. Med. Chem. 54, 7397–7407. https://doi.org/10.1021/jm201034h

Kansy, M., Fischer, H., Kratzat, K., Senner, F., Wagner, B., Parrilla, I., 2001. High-throughput artificial membrane permeability studies in early lead discovery and development, in: Testa, B., Waterbeemd, H. van de, Folkers, G., Guy, R. (Eds.), Pharmokinetic Optimization in Drug Research. Verlag Helvetica Chemica Acta, Zurich (Switzerland), pp. 447–464.

Kansy, M., Senner, F., Gubernator, K., 1998. Physicochemical high throughput screening: parallel artificial membrane permeation assay in the description of passive absorption processes. J. Med. Chem. 41, 1007–1010.

Khan, H., Jawad, M., Kamal, M.A., Baldi, A., Xiao, J., Nabavi, S.M., Daglia, M., 2018. Evidence and prospective of plant derived flavonoids as antiplatelet agents: Strong candidates to be drugs of future. Food Chem. Toxicol. https://doi.org/10.1016/j.fct.2018.02.014

Kosina, P., Kren, V., Gebhardt, R., Grambal, F., Ulrichova, J., Walterova, D., 2002. Antioxidant properties of silybin glycosides. Phytother. Res. 16, 33–39. https://doi.org/10.1002/ptr.796

Křen, R., Walterová, D., 2005. Silybin and silymarin-new and emerging applications in medicine. Biomed. Pap. 149, 29–41.

Křenek, K., Marhol, P., Peikerová, Ž., Křen, V., Biedermann, D., 2014. Preparatory separation of the silymarin flavonolignans by Sephadex LH-20 gel. Food Res. Int., 7th World Congress on Polyphenols Applications 65, 115–120. https://doi.org/10.1016/j.foodres.2014.02.001

Kumar, S., Pandey, A.K., 2013. Chemistry and biological activities of flavonoids: An overview. Sci. World J. 1–16. https://doi.org/10.1155/2013/162750

Leung, S.S.F., Mijalkovic, J., Borrelli, K., Jacobson, M.P., 2012. Testing physical models of passive membrane permeation. J. Chem. Inf. Model. 52, 1621–1636. https://doi.org/10.1021/ci200583t

Nakao, K., Fujikawa, M., Shimizu, R., Akamatsu, M., 2009. QSAR application for the prediction of compound permeability with *in silico* descriptors in practical use. J. Comput. Aided Mol. Des. 23, 309–319. https://doi.org/10.1007/s10822-009-9261-8

Noll, C., Hamelet, J., Matulewicz, E., Paul, J.-L., Delabar, J.-M., Janel, N., 2009. Effects of red wine polyphenolic compounds on paraoxonase-1 and lectin-like oxidized low-density lipoprotein receptor-1 in hyperhomocysteinemic mice. J. Nutr. Biochem. 20, 586–596. https://doi.org/10.1016/j.jnutbio.2008.06.002

Pyszková, M., Biler, M., Biedermann, D., Valentová, K., Kuzma, M., Vrba, J., Ulrichová, J., Sokolová, R., Mojović, M., Popović-Bijelić, A., Kubala, M., Trouillas, P., Křen, V., Vacek, J., 2016. Flavonolignan 2,3-dehydroderivatives: Preparation, antiradical and cytoprotective activity. Free Radic. Biol. Med. 90, 114–125. https://doi.org/10.1016/j.freeradbiomed.2015.11.014

Sak, K., 2014. Cytotoxicity of dietary flavonoids on different human cancer types. Pharmacogn. Rev. 8, 122.

Testino, G., Leone, S., Ansaldi, F., Borro, P., 2013. Silymarin and S-adenosyl-L-methionine (SAMe): two promising pharmacological agents in case of chronic alcoholic hepathopathy. A review and a point of view. Minerva Gastroenterol. Dietol. 59, 341–356.

Thilakarathna, S., Rupasinghe, H., 2013. Flavonoid bioavailability and attempts for bioavailability enhancement. Nutrients 5, 3367–3387. https://doi.org/10.3390/nu5093367

Thilakarathna, S.H., Wang, Y., Rupasinghe, H.P.V., Ghanam, K., 2012. Apple peel flavonoid- and triterpene-enriched extracts differentially affect cholesterol homeostasis in hamsters. J. Funct. Foods 4, 963–971. https://doi.org/10.1016/j.jff.2012.07.004

Verma, R.P., Hansch, C., Selassie, C.D., 2007. Comparative QSAR studies on PAMPA/modified PAMPA for high throughput profiling of drug absorption potential with respect to Caco-2 cells and human intestinal absorption. J. Comput. Aided Mol. Des. 21, 3–22. https://doi.org/10.1007/s10822-006-9101-z

Wu, J.-W., Lin, L.-C., Tsai, T.-H., 2009. Drug–drug interactions of silymarin on the perspective of pharmacokinetics. J. Ethnopharmacol. 121, 185–193. https://doi.org/10.1016/j.jep.2008.10.036

Xiao, J., 2015. Dietary flavonoid aglycones and their glycosides: which show better biological significance? Crit. Rev. Food Sci. Nutr. 00–00. https://doi.org/10.1080/10408398.2015.1032400

Xiao, J., 2018. Stability of dietary polyphenols: It’s never too late to mend? Food Chem. Toxicol. https://doi.org/10.1016/j.fct.2018.03.051

Xiao, J., Högger, P., 2015. Stability of dietary polyphenols under the cell culture conditions: Avoiding erroneous conclusions. J. Agric. Food Chem. 63, 1547–1557. https://doi.org/10.1021/jf505514d

Xiao, J.B., Hogger, P., 2014. Dietary polyphenols and type 2 diabetes: Current insights and future perspectives. Curr. Med. Chem. 22, 23–38. https://doi.org/10.2174/0929867321666140706130807

## Figures

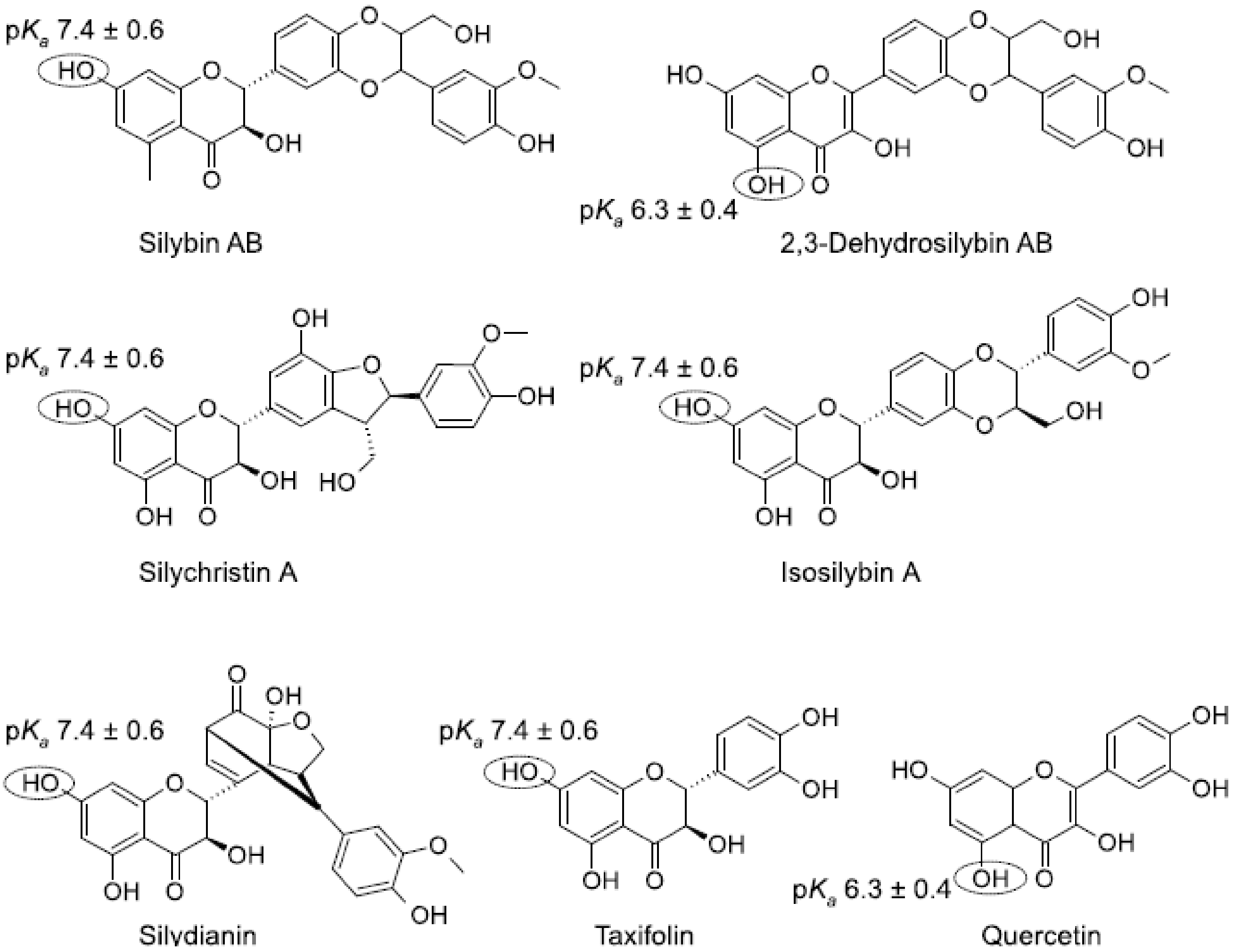

**Fig. 1.** Chemical structures of the flavonoids investigated *in vitro* and their calculated lowest acidic pKa values shown next to the corresponding hydroxyl group.

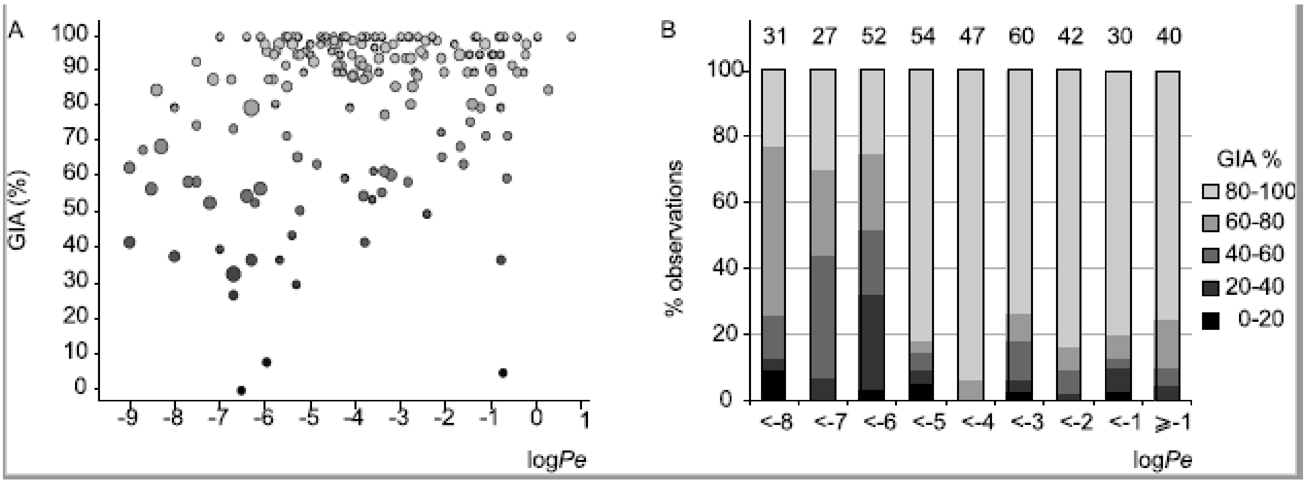

**Fig. 2.** Correspondence between log Pe and GIA (%) for 167 compounds present in both PAMPA Pe and GIA datasets: A – mean GIA values vs. PAMPA log Pe; size of the circles corresponds to the number of averaged GIA values for the compound. B – distribution of GIA classes among PAMPA Pe classes (numbers on top of the columns correspond to the number of distinct GIA values in each PAMPA Pe class).

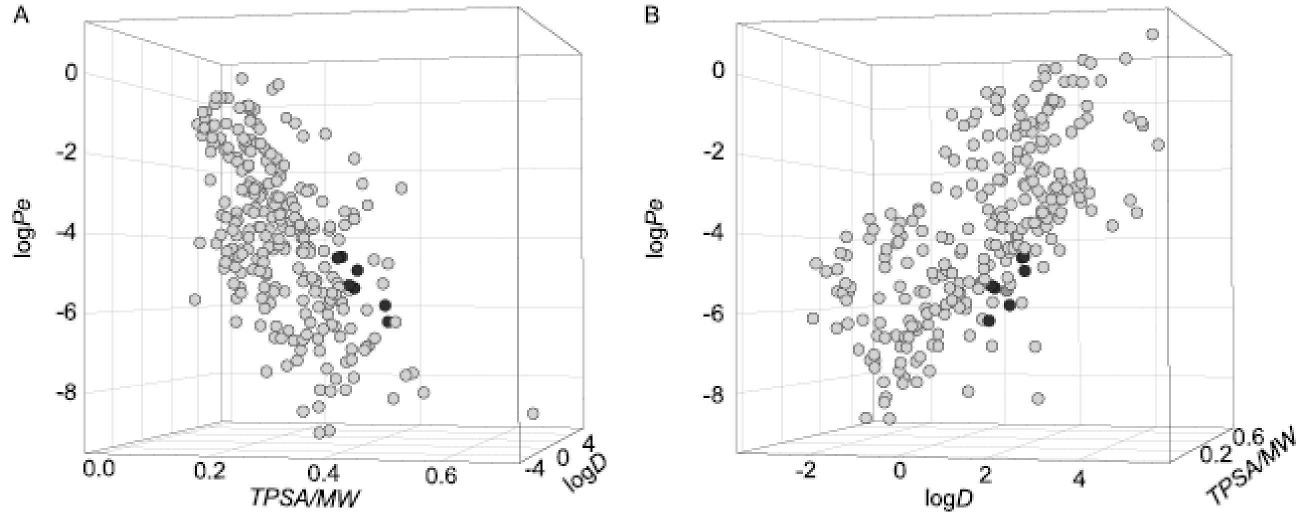

**Fig. 3.** 3-D plots of experimental log Pe vs. calculated structural descriptors TPSA/MW (A) and log D at pH 7.4 (B) obtained by the ACD/Percepta model (equation 1) as the x-axis respectively for the training set of compounds (◦) and the predicted flavonoids (•). The parameters' intervals are: –9÷0.78 for log Pe; 0.011÷0.695 for TPSA/MW and –3.16÷5.51 for log D (pH 7.4).

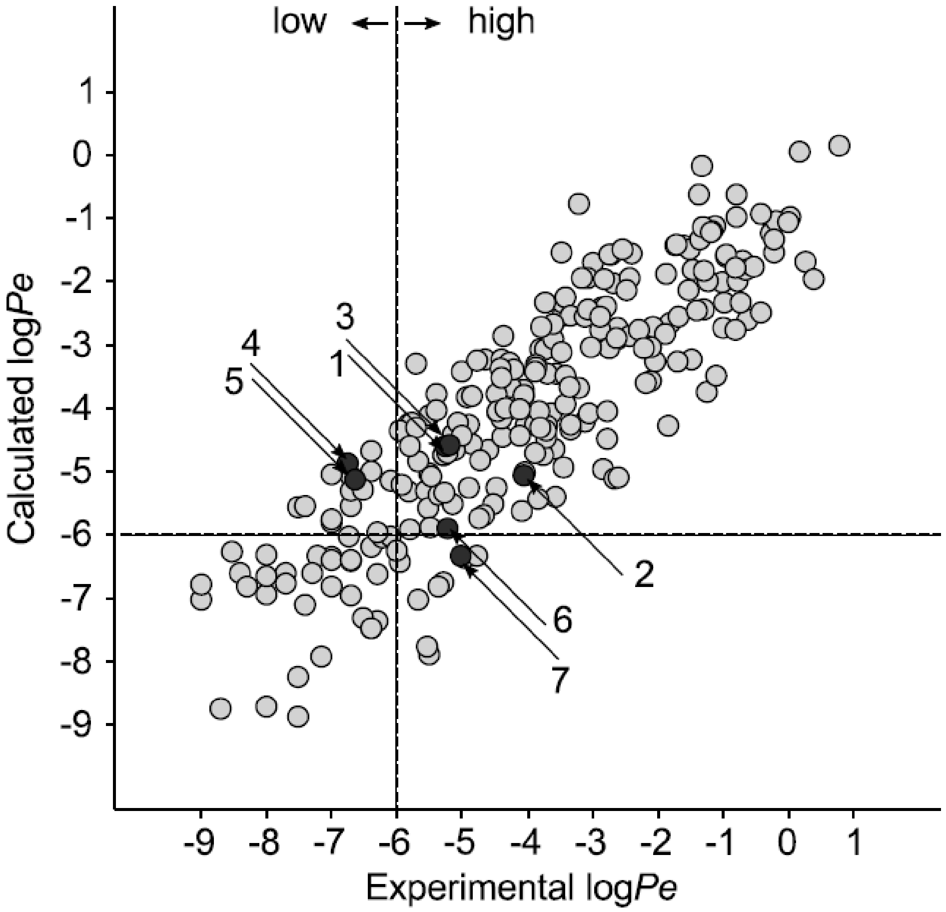

**Fig. 4.** Plot of experimental vs. calculated log Pe values for the flavonoids studied: ○ – compounds used to derive the PAMPA QSAR model; ● – compounds studied: silybin AB (**1**), 2,3–dehydrosilybin AB (**2**), isosilybin A (**3**); silychristin A (**4**), silydianin (**5**), taxifolin (**6**), quercetin (**7**); the dashed line represents the border between low and high permeability.

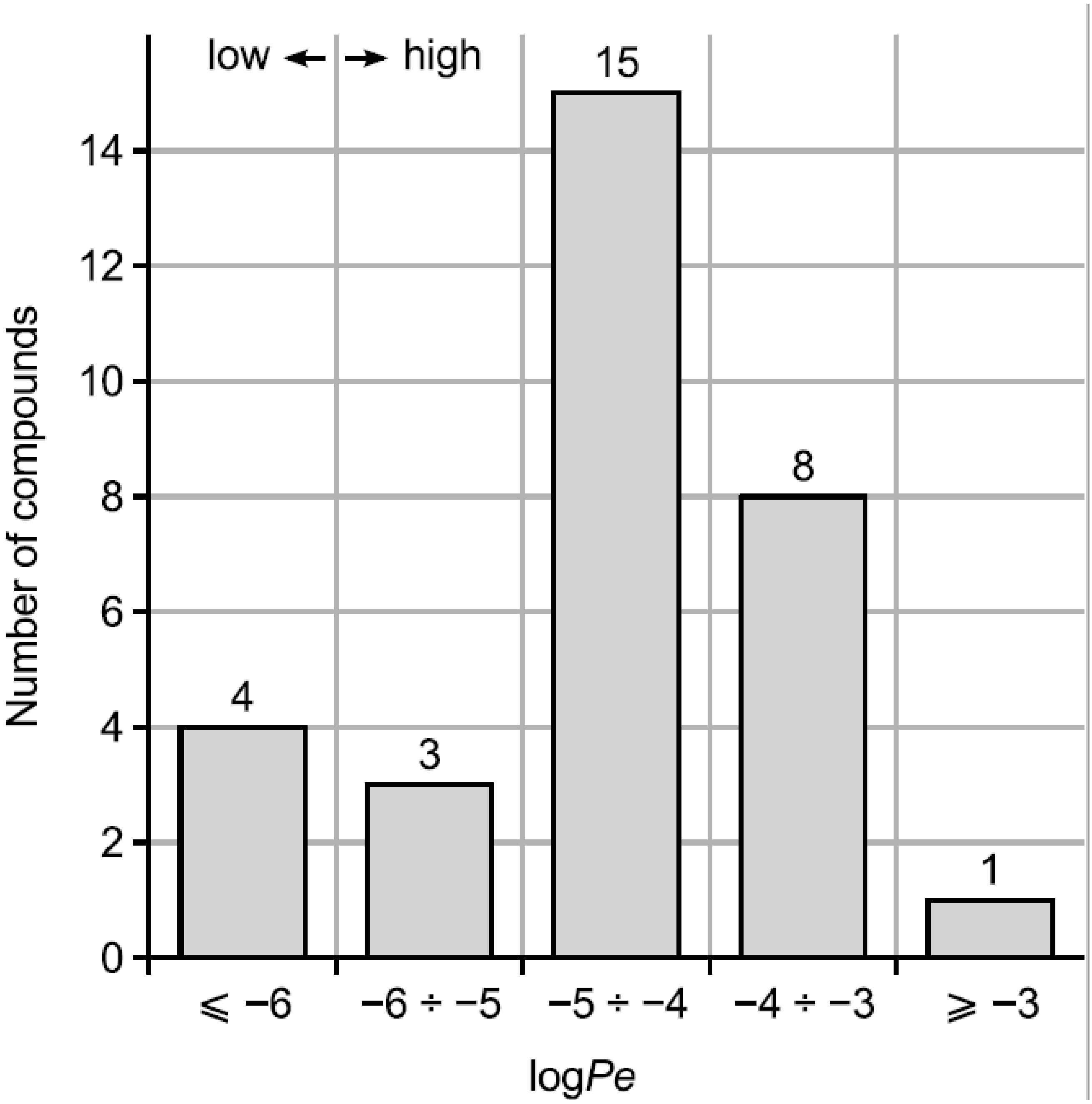

**Fig. 5.** Distribution of the flavonolignans according to their predicted log Pe.

## Tables

**Table 1.** Effective membrane permeability log Pe ± SD of the compounds studied. The SD values have been calculated based on 3 parallel experiments.

| **pH**<br>**Compound** | **5.0** | **6.2** | **7.4** |
|---|---|---|---|
| Silybin AB | –4.11 ± 0.03 | –4.14 ± 0.03 | –5.25 ± 0.05 |
| 2,3-Dehydrosilybin AB | –4.11 ± 0.06 | –4.17 ± 0.03 | –4.06 ± 0.03 |
| Isosilybin A | –4.32 ± 0.09 | –4.31 ±0.06 | –5.19 ± 0.02 |
| Silychristin A | –6.14 ± 0.08 | –6.09 ± 0.05 | –6.75 ± 0.11 |
| Silydianin | –5.76 ± 0.05 | –5.79 ± 0.04 | –6.64 ± 0.09 |
| Taxifolin | –5.95 ± 0.10 | –5.93 ± 0.02 | –5.23 ± 0.01 |
| Quercetin | –5.14 ± 0.42 | –5.10 ± 0.17 | –5.02 ± 0.07 |

**Table 2.** The CDK and Indigo calculated molecular descriptors and TPSA/descriptor ratios with the highest correlation to TSA (A) and to PSA/TSA (B).

| **A** | | **B** | |
|---|---|---|---|
| **Descriptors** | **r** | **Ratios** | **r** |
| TSA | 1.000 | PSA/TSA | 1.000 |
| Atomic polarizabilities | 0.959 | TPSA/VABC volume descriptor | 0.881 |
| Number of heavy atoms | 0.957 | TPSA/Molecular weight | 0.880 |
| VABC volume descriptor | 0.954 | TPSA/Atomic polarizabilities | 0.878 |
| Number of bonds | 0.951 | TPSA/Number of heavy atoms | 0.876 |
| Number of carbons | 0.950 | TPSA/Total number of atoms | 0.873 |
| Molecular weight | 0.946 | TPSA/Bond polarizabilities | 0.848 |
| Total number of atoms | 0.941 | TPSA/Number of bonds | 0.842 |
| Zagreb index | 0.923 | TPSA/Zagreb index | 0.801 |
| Vertex adjacency information magnitude | 0.917 | TPSA/Number of carbons | 0.740 |
| Bond polarizabilities | 0.913 | TPSA/Vertex adjacency information magnitude | 0.686 |

r – correlation coefficient, TSA – total surface area, PSA – polar surface area, TPSA – topological polar surface area, VABC – sum of atomic and bond contributions volume.

**Table 3.** Statistical parameters of a set of tested DS-PAMPA Pe models based on two differently calculated log D estimates, on PSA/TSA, and on three different substitutes for the PSA/TSA ratio.

**A**

| **log D** | **surface descriptors** | **N** | **$r^2$** | **SEE** | **F** | **LOO $q^2$** |
|---|---|---|---|---|---|---|
| ACD/Percepta-calculated | PSA/TSA | 259 | 0.69 | 1.20 | 286 | 0.68 |
| | TPSA/VABC | 254 | 0.74 | 1.11 | 354 | 0.73 |
| | TPSA/MW | 251 | 0.75 | 1.10 | 371 | 0.74 |
| | TPSA/AP | 253 | 0.74 | 1.10 | 350 | 0.73 |

**B**

| **log D** | **surface descriptors** | **N** | **$r^2$** | **SEE** | **F** | **LOO $q^2$** |
|---|---|---|---|---|---|---|
| ChemAxon tools-calculated | PSA/TSA | 245 | 0.75 | 1.08 | 370 | 0.75 |
| | TPSA/VABC | 245 | 0.74 | 1.09 | 348 | 0.74 |
| | TPSA/MW | 248 | 0.74 | 1.11 | 345 | 0.73 |
| | TPSA/AP | 245 | 0.74 | 1.09 | 351 | 0.74 |

N – number of compounds in the model set (starting number of compounds was 269), $r^2$ – multiple correlation coefficient, SEE – standard error of estimate, F – F-ratio, LOO $q^2$ – leave-one-out cross-validation correlation coefficient, VABC – sum of atomic and bond contributions volume, MW – molecular weight, AP – atomic polarizabilities.

**Table 4.** Statistical parameters for the classification power of the PAMPA Pe, predicted by TPSA/MW-based models, with respect to GIA.

| **Model implementation** | **accuracy** | **sensitivity** | **specificity** | **% outliers** |
|---|---|---|---|---|
| ACD/Percepta-calculated log D | 76.1 | 83.9 | 58.3 | 11.6 |
| ChemAxon tools-calculated log D | 77.1 | 84.4 | 60.0 | 14.6 |

**Table 5.** Calculated molecular descriptors and log Pe values predicted by the QSAR model for the flavonoids studied.

| **Compound** | **log D at pH 7.4** | **TPSA/MW** | **Predicted log Pe** |
|---|---:|---:|---:|
| Silybin AB | 1.77 | 0.322 | –4.60 |
| 2,3-Dehydrosilybin AB | 1.03 | 0.331 | –5.06 |
| Isosilybin A | 1.82 | 0.322 | –4.57 |
| Silychristin A | 1.70 | 0.345 | –4.86 |
| Silydianin | 1.03 | 0.338 | –5.12 |
| Taxifolin | 1.15 | 0.419 | –5.89 |
| Quercetin | 0.59 | 0.435 | –6.32 |

**Table 6.** Calculated values of the molecular descriptors and log Pe values predicted by the QSAR model of 31 silybin congeners.

| **Name** | **log D at pH 7.4** | **TPSA/MW** | **Predicted log Pe at pH 7.4** |
|---|---:|---:|---:|
| 7-O-Benzylsilybin [a] | 3.89 | 0.252 | –2.85 |
| 5,7,20-tri-O-Methylsilybin [a] | 2.92 | 0.233 | –3.13 |
| 7-O-Benzoylsilybin [a] | 3.65 | 0.275 | –3.20 |
| 5,7,20-tri-O-Methyl-2,3-dehydrosilybin [a] | 2.71 | 0.241 | –3.32 |
| 23-O-Pivaloylsilybin [a] | 3.40 | 0.285 | –3.42 |
| 7-O-Benzyl-2,3-dehydrosilybin [a] | 2.87 | 0.260 | –3.43 |
| 3,7,20-tri-O-Methyl-2,3-dehydrosilybin [a] | 2.28 | 0.241 | –3.53 |
| 7,20-di-O-Methylsilybin [a] | 2.68 | 0.261 | –3.54 |
| 19-O-Demethyl-19-O-benzyl-2,3-dehydrosilybin [a] | 2.45 | 0.286 | –3.90 |
| 7,20-di-O-Methyl-2,3-dehydrosilybin [a] | 1.93 | 0.270 | –3.99 |
| 3-O-Methyl-silybin [b] | 2.35 | 0.291 | –4.00 |
| 7-O-Methylsilybin [a] | 2.26 | 0.291 | –4.04 |
| 20-O-Methylsilybin [a] | 2.21 | 0.291 | –4.07 |
| 3,7-di-O-Methyl-2,3-dehydrosilybin [a] | 1.73 | 0.270 | –4.09 |
| 3,20-di-O-Methyl-2,3-dehydrosilybin [a] | 1.59 | 0.270 | –4.16 |
| 23-O-Galloylsilybin [c] | 2.85 | 0.350 | –4.35 |
| 23-O-Methyl-2,3-dehydrosilybin [b] | 1.70 | 0.300 | –4.41 |
| 7-O-Methyl-2,3-dehydrosilybin [a] | 1.62 | 0.300 | –4.45 |
| 3-O-Galloylsilybin [c] | 2.60 | 0.350 | –4.48 |
| 20-O-Methyl-2,3-dehydrosilybin [a] | 1.52 | 0.300 | –4.49 |
| 20-O-Galloylsilybin [c] | 2.55 | 0.350 | –4.50 |
| 5-O-Methyl-dehydrosilybin [b] | 1.46 | 0.300 | –4.52 |
| 3-O-Methyl-2,3-dehydrosilybin [a] | 1.36 | 0.300 | –4.57 |
| 7-O-Galloylsilybin [c] | 1.86 | 0.350 | –4.84 |
| 19-O-Demethyl-2,3-dehydrosilybin [a] | 0.88 | 0.365 | –5.47 |
| Silybin 23-O-β-galactoside [d] | –0.12 | 0.364 | –5.95 |
| Silybin 23-O-β-glucoside [d] | –0.12 | 0.364 | –5.95 |
| Silybinic acid [a] | –1.75 | 0.347 | –6.58 |
| Silybin 23-O-β-lactoside [d] | –1.00 | 0.389 | –6.63 |
| Silybin 23-O-β-maltoside [d] | –1.00 | 0.389 | –6.63 |
| 2,3-Dehydrosilybinic acid [a] | –2.28 | 0.356 | –6.93 |

Structures taken from: [a] Džubák et al. (2006), [b] Gažák et al. (2009), [c] Gažák et al. (2011), [d] Kosina et al. (2002).