Journal of Pharmaceutical and Biomedical Analysis 203 (2021) 114218

Contents lists available at ScienceDirect

Journal of Pharmaceutical and Biomedical Analysis

journal homepage: www.elsevier.com/locate/jpba

A corneal-PAMPA-based in silico model for predicting corneal permeability

Anna Vincze[a], Gergő Dargó[a], Anita Rácz[c,**], György T. Balogh[a,b,*]

[a] Department of Chemical and Environmental Process Engineering, Budapest University of Technology and Economics, Műegyetem Rakpart 3., 1111, Budapest, Hungary
[b] Department of Pharmacodynamics and Biopharmacy, University of Szeged, Eötvös u. 6., 6720, Szeged, Hungary
[c] Institute of Materials and Environmental Chemistry, Research Centre for Natural Sciences, Magyar Tudósok Krt. 2., 1117, Budapest, Hungary

ARTICLE INFO

Article history:
Received 3 February 2021
Received in revised form 14 June 2021
Accepted 15 June 2021
Available online 17 June 2021

Keywords:
Corneal permeability
In vitro
Non-cell-based model
PAMPA
In silico model
QSPR
Quantitative structure-property relationships
Lipophilicity
Polar surface area

ABSTRACT

The capability to predict corneal permeability based on physicochemical parameters has always been a desirable objective of ophthalmic drug development. However, previous work has been limited to cases where either the diversity of compounds used was lacking or the performance of the models was poor. Our study provides extensive quantitative structure-property relationship (QSPR) models for corneal permeability predictions. The models involved in vitro corneal permeability measurements of 189 diverse compounds. Preliminary analysis of data showed that there is no significant correlation between corneal-PAMPA (Parallel Artificial Membrane Permeability Assay) permeability values and other pharmacokinetically relevant *in silico* drug transport parameters like Caco-2, jejunal permeability and blood-brain partition coefficient (logBB). Two different QSPR models were developed: one for corneal permeability and one for corneal membrane retention, based on experimental corneal-PAMPA permeability data. Partial least squares regression was applied for producing the models, which contained classical molecular descriptors and ECFP fingerprints in combination. A complex validation protocol (including internal and external validation) was carried out to provide robust and appropriate predictions for the permeability and membrane retention values. Both models had an overall fit of $R^2 > 0.90$, including R^2-values not lower than 0.85 for validation runs, and provide quick and accurate predictions of corneal permeability values for a diverse set of compounds.

Abbreviations: ADME, absorption, distribution, metabolism, excretion; API, active pharmaceutical ingredient; Caco-2, cell line of heterogeneous human epithelial colorectal adenocarcinoma cells; clog$D_{7.4}$, log of calculated distribution coefficient at pH 7.4; clog$P$, log of calculated partition coefficient; clog$P_{e\ Caco\text{-}2}$, log of calculated Caco-2 permeability; clog$P_{e\ jejunal}$, log of calculated jejunal permeability; DOOD, D-optimal onion design algorithm; f_{bu}, fraction unbound in the brain; GA, genetic algorithm; HBD, number of hydrogen bond donors; HPLC, high performance liquid chromatography; k_a, absorption rate constant; logBB, blood-brain partition coefficient; log$D_{7.4}$, log of distribution coefficient at pH 7.4; log$DI_{7,65}$, degree of ionization at pH 7.65; log$P$, log of partition coefficient; logP_e, log of effective permeability; log$S$, log aqueous solubility measured in mol/L; MR, membrane retention; MW, molecular weight; NN, neural networks; PAMPA, parallel artificial membrane permeability assay; PBS, phosphate buffered saline; PC, phosphatidylcholine; PCA, principal component analysis; P-gp, P,-glycoprotein; PLS, partial least squares; PLS, COMP. number of PLS components in the model; PPB, plasma protein binding; PSA, polar surface area; Q^2_{CV}, goodness of fit for the cross-validation; R^2_P, goodness of fit for the external validation; QSAR, quantitative structure-activity relationship; QSPR, quantitative structure-property relationship; R^2, goodness of fit for the calibration; RMSEC, root mean squared error of calibration; RMSECV, root mean squared error of cross-validation; RMSEP, root mean squared error of external validation; TPSA, topological polar surface area; VIP, variable importance projection; λ_{max}, the wavelength of maximum absorbance; Y-SCR, CV goodness of fit for the cross validation using Y scrambling; Y-SCRP., goodness of fit for the test prediction using Y-scrambling.

* Corresponding author at: Department of Chemical and Environmental Process Engineering, Budapest University of Technology and Economics, Műegyetem rakpart 3., 1111, Budapest, Hungary.
** Corresponding author.
E-mail addresses: racz.anita@ttk.hu (A. Rácz), balogh.gyorgy@vbk.bme.hu (G.T. Balogh).

https://doi.org/10.1016/j.jpba.2021.114218

1. Introduction

The eye is a challenging target of therapy. Because of its complicated anatomic structure and small absorptive surface, it is difficult to achieve therapeutic drug levels in the interior of the eye by topical dosage forms (eye drops, ointments, in-situ gel-forming systems, etc.). More invasive ones such as intraocular injections or implants can be very useful for maintaining high drug levels in both the anterior and posterior segment [1–3]. Topical administration is the most common way of anterior segment treatment as it is non-invasive, easy and very comfortable. However, it has many drawbacks including low bioavailability despite the fact that the active pharmaceutical ingredient (API) is usually applied in high doses [4]. Blinking, tear formation and nasolacrimal drainage cause a huge drop in drug level as the solutions are washed away within the first 15–30 seconds from the eye surface [5,6]. Typically no more than 5% bioavailability can be reached in the anterior segment by topical administration due to these defense mechanisms, and membrane penetration issues [4–7]. By using dosage forms with high viscosity and mucoadhesivity the residence time of drugs can be increased, however, these ointments and in situ gelling systems often cause irritation and blurred vision [1,4,5,8]. Higher drug levels can be reached in the posterior segment of the eye by systemic administration or intraocular injections and implants [2,3,5,9]. While the first may result in systemic side effects due to high dosage the latter does not require frequent use, although, it may cause endophthalmitis, lens damage or retinal detachment [4,10,11]. Despite low patient compliance, the intraocular injections (usually targeting the vitreous) are still the most commonly used posterior segment therapies since there are no other efficient delivery routes for reaching high drug levels in the posterior with minimal systemic side effects [9].

The possible routes of absorption are the corneal and non-corneal pathways in terms of topical administration [5,6,10]. In the case of non-corneal routes the drug penetrates through the conjunctiva and/or sclera which is not very effective due to the presence of local capillary beds that continuously remove the drug, transferring it into the systemic circulation. On the other hand, the corneal route represents the main absorption path for most ophthalmic therapeutics (small, lipophilic compounds) and for this reason, it is the most studied area both in terms of drug discovery and drug formula development [4,5]. The cornea is a clear and avascular tissue consisting of 5 layers: corneal epithelium, Bowman's layer, stroma, Descemet's membrane and endothelium. Although the relative thicknesses of corneal epithelium, stroma and endothelium are about 0.1:1:0.01, the thinnest corneal epithelium is the main barrier of drug absorption into the eye [7,10].

Like in the case of other topical dosage forms, the in vivo human pharmacokinetic characterization of APIs in ophthalmic treatment is greatly hampered by the difficulty of directly examining topical drug concentrations. Furthermore, due to technical difficulties no suitable in vivo animal testing solution is available either [12]. For preclinical ophthalmic absorption studies *ex vivo* and in vitro cellular models are available. For *ex vivo* studies excised rabbit cornea is used most commonly because of easy access and anatomic similarities between rabbit and human eyes [10,13]. However, there are also some differences between the two species and thus the penetration is generally much higher and cannot be correlated adequately to humans [10]. In vitro studies apply rabbit or human cell cultures which are considered as a more ethical alternative. In this case, mostly corneal epithelial cells are used considering the main role of corneal epithelium in drug transport, however, it is difficult to isolate ocular tissues for this purpose and the life span of these primary cells is very short [13,14]. Although immortalized cell lines provide an inexhaustible supply of cells they may exhibit abnormal gene expression or biological functions especially after several divisions. Also, penetration across cell cultures is only a crude estimate at best since a single type of cell cannot represent the complexity of the corneal barrier as a whole [13].

The capability to predict corneal permeability based on physicochemical parameters has always been a desirable objective of ophthalmic drug development. For this purpose, great efforts have been made to develop predictive models starting with the work of Schoenwald and Ward [15]. A summary of the early development stages can be found in Table 1. These models were based on just a few physicochemical properties but were typically not generally applicable – their diversity and performance were limited to only certain groups of molecules.

In the past few decades, quantitative structure-property relationship (QSPR) techniques have become a well-established method for correlating molecular structure and physicochemical properties of chemical compounds with the use of several descriptors and nowadays they are commonly used for the fast prediction of different ADME (absorption, distribution, metabolism, excretion) related molecular properties [21,22]. There have been a few attempts to develop QSPR models also for the prediction of corneal permeability. Kidron et al. used partial least squares (PLS) regression for 69 non-congeneric compounds and concluded that the total number of putative hydrogen bonds (HB_{TOT}) and logarithmic distribution coefficients ($\log D_{pH7.4}$, $\log D_{pH8.0}$) were the most significant descriptors for corneal permeability [12]. The significance of lipophilicity in corneal penetration has been recognized long ago, while for ionizable compounds (most of the ophthalmic drugs) the distribution coefficient measured at the pH of aqueous humor and tear film (pH 7.2–7.6) has been found to be a more proper descriptor in terms of membrane permeability [5,8]. In the work of Ghorbanzade et al. the same dataset was analyzed by stepwise multivariate linear regression (MLR), and neural networks (NN) as well [23]. With the multilayer perceptron-NN technique, they established a nonlinear model for corneal permeability with moderate performance. Although Kidron and Ghorbanzade analyzed more compounds compared to their predecessors, the data came from several sources with different experimental conditions.

In 2018, Ramsay et al. produced a corneal and conjunctival QSPR model by PLS and principal component analysis (PCA) based on an *ex vivo* dataset of 32 small molecules. They found that corneal permeability has an inverse relationship with the polar surface area (PSA) and the number of hydrogen bond donors (HBD) [24]. However, the predictive power of their model was relatively modest.

Up to the present day, there have been several models developed for predicting corneal permeability, but there is still no reliable QSPR model based on a large and homogeneous dataset. Of course, it is not an easy mission since *ex vivo* experiments have serious constraints like the demand for numerous animals, appropriate tissue viability and high costs. In vitro cell culture models are better in terms of cost-effectiveness and throughput, however, the use of cell-based assays is not a robust process.

To solve this problem, we have recently developed a novel in vitro non-cellular PAMPA-based method for investigating the corneal permeability of APIs [25]. Corneal-PAMPA is a robust, high-throughput assay using easily accessible materials like phosphate buffered saline (PBS), phosphatidylcholine, dodecane, hexane and chloroform. For developing the method, *ex vivo* rabbit corneal permeability data were used from heterogeneous data sets in terms of experimental conditions. Despite the use of heterogeneous data, the PAMPA model with standardized experimental conditions provided highly reproducible permeability values that correlated well with the *ex vivo* data ($R^2 = 0.88$). The corneal-PAMPA methodology can provide many corneal permeability values as a homogeneous dataset in a short time period and therefore it is a suitable basis for the development of an *in silico* model.

Table 1
Detailed summary of the early models for corneal permeability determination.

Equation	Type of molecules (and numbers)	Molecular properties	Ref.
parabolic	steroids (11)	log*P*	[15]
parabolic	*N*-alkyl-*p*-aminobenzoate esters (6)	log*P*	[16]
parabolic	β-blocking agents (12)	log*P*, log$D_{pH7.65}$, logMW, log$DI_{7.65}$*	[17]
parabolic/ sigmoidal	β-blocking agents (13)	log*P*	[18]
parabolic	β-blockers, steroids, miscellaneous compounds (32)	log*D*, Δlog*P* (log$P_{octanol}$-logP_{alkane})	[19]
MLR model	β-blockers, steroids, miscellaneous compounds (30)	quantum chemical descriptors (3)	[20]

* Degree of ionization at pH 7.65.

In the present work, the corneal permeability of 189 APIs selected from a diverse chemical space was determined by the corneal-PAMPA method [25] experimentally and used to build a robust QSPR model for predicting corneal permeability, which could be a useful tool in the development of future ophthalmic agents.

2. Materials and methods

2.1. Materials

Analytical grade solvents like acetonitrile, chloroform, hexane, dodecane and formic acid were purchased from Merck KGaA (Darmstadt, Germany). Phosphate buffered saline (PBS) powder, L-α-phosphatidylcholine (PC) and all investigated compounds (*See Table A1., Appendix A*) were purchased from Sigma Aldrich Co. Ltd. (Budapest, Hungary) except for betaxolol hydrochloride, bevantolol hydrochloride, bromfenac sodium, bufuralol hydrochloride, 11-deoxycorticosterone, ethoxzolamide, flurbiprofen, nepafenac and penbutolol, which were purchased from Toronto Research Chemicals Inc. (North York, Toronto, Canada). Distilled water was purified by the Millipore Milli-Q® 140 Gradient Water Purification System.

2.2. In vitro permeability studies

The corneal-PAMPA measurements were based on our previous report [25]. Briefly all compounds were dissolved in PBS buffer (pH 7.4) to make solutions with 100 μM nominal concentration. Due to the negative effect of using DMSO as a cosolvent on the goodness of prediction in the corneal-PAMPA assay system (according to our previous report [25]), no DMSO was used for this study either. 300 μL of the initial PBS solutions were placed into a 96-well polypropylene plate (A) (Agilent, Waldbronn, Germany). For the corneal-PAMPA membrane, 16 mg PC was dissolved in a solvent mixture of chloroform, hexane and dodecane 5:70:25 % (v/v) in a total volume of 600 μL. Each well of the acceptor plate (B) (Multiscreen Acceptor Plate, MSSACCEPTOR; Millipore) contained 300 μL PBS and each well of the donor plate (MultiscreenTM-IP, MAIPN4510, pore size 0.45 μm; Millipore) (C) was coated with 5 μL of the lipid solution. After evaporation of hexane and chloroform, the donor plate was inserted into the acceptor plate, and each well of the donor plate was filled with 150 μL PBS solutions of compounds from plate A. After that, the donor plate was covered with a sheet of wet tissue paper and plate lid to prevent evaporation of the solvent during the 4 h incubation (35 °C, Heidolph Titramax 1000). In each assay, two reference compounds were measured: buspirone and hydrocortisone with high and low corneal permeability values, respectively.

For HPLC analysis at least 120 μL aliquots were taken out of every well of plates B and C and transferred to 96-well polypropylene plates and sealed. The initial plate A was analyzed as well. The effective permeability and membrane retention of drugs were calculated using the following equations [26]:

$$P_e = \frac{-2.303}{A \times (t - \tau_{ss})} \times \left(\frac{1}{1+r_v}\right) \times \lg\left[-r_v + \left(\frac{1+r_v}{1-MR}\right) \times \frac{c_D(t)}{c_D(0)}\right] \quad (1)$$

where P_e is the effective permeability (cm/s), A is the filter area (0.3 cm^2), t is the incubation time of PAMPA assay (14400 s), τ_{ss} is the time to reach steady-state (s), r_v is the volume ratio of aqueous compartments (V_D/V_A), V_D and V_A are the volumes in the donor (0.15 cm^3) and acceptor phase (0.3 cm^3), $c_D(t)$ is the concentration of the compound after incubation on plate C (mol/cm^3), $c_D(0)$ is the concentration of the compound on plate A (mol/cm^3) and MR is the membrane retention factor, defined as:

$$MR = 1 - \frac{c_D(t)}{c_D(0)} - \frac{V_A c_A(t)}{V_D c_D(t)} \quad (2)$$

where $c_A(t)$ is the concentration of the compound after incubation on plate B (mol/cm^3). In tables and figures we usually use this term as 100*MR (%).

2.3. HPLC analysis

Quantitative chromatographic analyses were carried out using an Agilent 1260 liquid chromatography system (Agilent Technologies, Palo Alto, CA, USA) on a Kinetex® 2.6 μm C18 100 Å LC column (30 × 3 mm) with a mobile phase flow rate of 1.1 mL/ min. Mobile phase A and B consisted of 0.1 % (v/v) formic acid in water, and acetonitrile/water (95/5) with 0.1 % (v/v) formic acid, respectively. A 3.91 min long linear gradient program was used: in the first 0.3 min 0% B, between 0.3 and 1.8 min 0–100 % B, then 100 % B was kept for 0.6 min and finally, at 2.41 the 100 % B was dropped to 0%. This was followed by an equilibration period of 1.5 min prior to the next injection. The injection volume was 6 μL. Chromatograms were recorded by a diode array detector in the wavelength range of 200–500 nm, and integration was carried out at the λ_{max} of each compound. Data acquisition and analysis were carried out with ChemStation B.04.03.

2.4. Dataset generation

The applied dataset contained the experimental corneal-PAMPA permeability and the membrane retention values of 189 APIs based on the previously mentioned experimental model [15]. The dominant protonation state at pH 7.4 was assigned to each compound with the ChemAxon Calculator (cxcalc) [27] and Schrödinger (LigPrep) [28,29]. Consequently, the appropriate 3D structures were used for the molecular descriptor generation.

Classical 1, 2 and 3D molecular descriptors and extended connectivity fingerprints (ECFP) were generated for the APIs with the DRAGON 7.0 software [30]. The definition of the applied molecular descriptors can be found in the work of Todeschini et al. [31]. Moreover, an additional 80 physicochemical properties were calculated by the ACD/Percepta software [32]. An inter-correlation limit of 0.997 was used for the filtering of the classical descriptors [33]. Constant descriptors were also excluded from the dataset. For the ECFP

fingerprints, default parameters were applied in DRAGON 7.0, with a maximum radius of 4 and a fingerprint length of 1024 [34]. The final number of generated molecular descriptors was 3302, which were further used for the QSPR model building.

2.5. QSPR modeling

Partial least squares (PLS) regression was applied as a classical and frequently used regression algorithm for the model building. In PLS regression, latent variables are calculated, which can be used for the prediction of the Y dependent variable, in our case the corneal-PAMPA permeability values and the membrane retention. As a well-established tool in QSAR/QSPR analysis, there is no need to explain the PLS algorithm in more detail other than referring to the tutorial paper of Geladi and Kowalski [35]. The optimal number of PLS components for the models was selected based on the global minimum of the root mean squared error of cross-validation (RMSECV). Outlier selection of the samples has been carried out by the plot of the first two latent variables (PLS components). The 95 % confidence ellipse was used for the determination of the outliers. The range of the Y variable (corneal-PAMPA values) was also modified to $0-50 \times 10^{-6}$ cm/s, due to the poor coverage outside of this range.

In each multivariate calibration with thousands of molecular descriptors, the use of variable selection techniques [36] is an essential step for the appropriate model building with good prediction performance. In our case, the variable importance projection (VIP) and genetic algorithm (GA) methods were combined for this purpose. First, the number of variables was decreased to the most important 20 % based on VIP values, then a genetic algorithm was used on the remaining variables. This form of pre-filtering with VIP scores was necessary in order to decrease the calculation time for the more time-demanding GA protocol. The major parameters of the genetic algorithm were the following: the population size was 64, the percentage of the initial terms was 30, the algorithm was run for a hundred generations and double cross-over was used.

Model building was carried out in PLS Toolbox [37]. Validation of the models is indispensable for reliable prediction, therefore our models were validated in three different ways: cross-validation, external validation and Y-scrambling [38]. Training and test splits were carried out by the D-optimal onion design algorithm (DOOD) implemented in PLS toolbox [39]. The data split ratio was 80 % Training – 20 % Test set. The training samples were used for 7-fold cross-validation, which means that 1/7 of the samples were predicted based on the other 6/7 of the samples, and this process was repeated while all the partitions have been predicted exactly once. The folds of samples were selected randomly, thus twenty iterations were implemented into the protocol. Seven-fold cross-validation is a well-known option for cross-validation processes. The optimal number of folds should be between 5 and 10 based on the work of Hastie et al. [40]. The test splits of the samples were used for the external validation, where the Y target values of the molecules were predicted based on the calibration model. Finally, Y-scrambling (randomly reordered Y vector) can verify that our models are far better than the use of random numbers. This can help us to check "chance correlations" and whether the models are overfitted or not. An additional method, called the permutation test was also applied to test the significance of the difference between the actual and the random model. Here the main hypothesis is that our original model is not significantly different from the randomly reordered Y vector ones. Wilcoxon, Sign and randomization T-tests were applied for testing.

Performance parameters such as R^2 (calibration), Q^2_{CV} (cross-validation), R^2_P (external validation) and Q^2_{Fn} [41], the root mean squared error of calibration, cross-validation and external validation (RMSEC/RMSECV/RMSEP respectively), the mean absolute error (for calibration, cross-validation and external validation) the coefficient of concordance (CCC, likewise for all calibration and validation steps) [42] etc. were calculated to appropriately determine the quality of the models.

3. Results and discussion

3.1. Preliminary analysis of the APIs

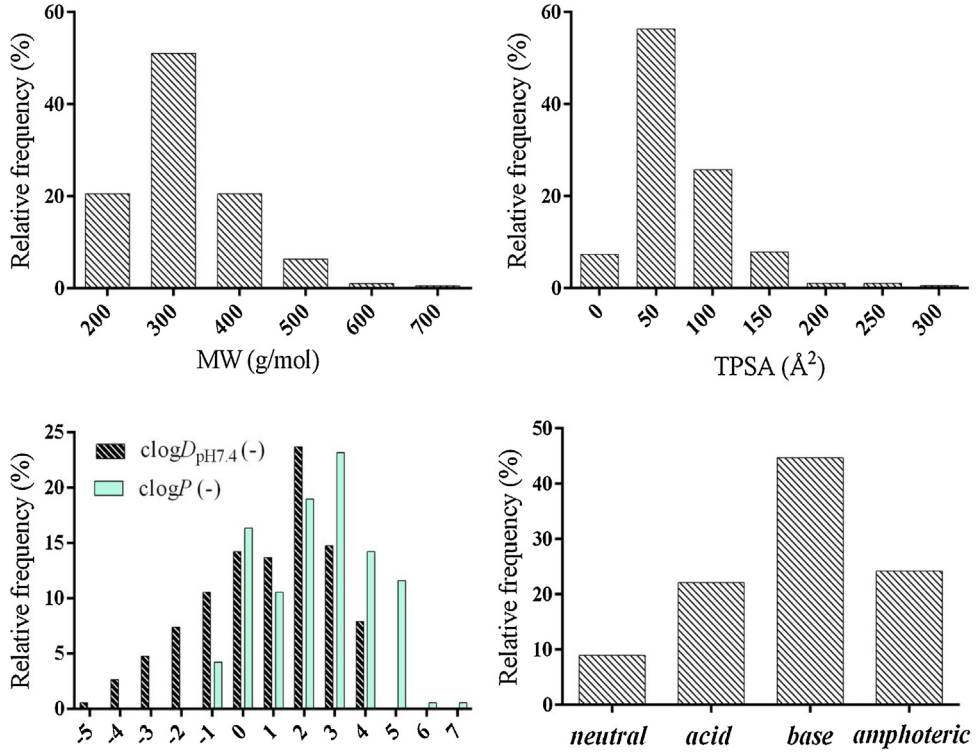
In our previous work, we investigated the relationships between corneal-PAMPA permeability ($\log P_e$) and membrane retention values (MR) in comparison with experimental Caco-2 permeability data and calculated physicochemical properties, such as molecular weight (MW), topological polar surface area (TPSA) and lipophilicity descriptors ($\log P$, $\log D_{pH7.4}$), in the case of 50 structurally and physicochemically diverse compounds [43]. Due to weak correlations ($R^2 = 0.018-0.506$) we concluded that corneal permeability cannot be predicted based on only these physicochemical descriptors and for this reason, a larger measured permeability dataset is needed to be able to carry out a QSPR analysis. As part of our conclusion, we established that the weak correlation between corneal-PAMPA permeability and Caco-2 (gastrointestinal) permeability ($R^2 = 0.125$) confirms that our model is independent of the generally accepted gastrointestinal permeability model and may support its adequacy for cornea-specific in vitro measurements.

In the current work, 189 APIs (from which 50 APIs have already been measured in the previous work [43]) were selected from a diverse chemical space and tested in the corneal-PAMPA model to produce data set for in silico modeling. (See experimental data in Table A1., Appendix A) Before computational modeling, we evaluated the main characteristics of the compounds based on their physicochemical properties like $\log P$, $\log D_{pH7.4}$, MW, TPSA and acid-base characteristics (Fig. 1).

We investigated the relationship between corneal-PAMPA permeability and the above-mentioned physicochemical parameters (Fig. 2). For linear regression analysis, the GraphPad Prism software was used [45]. The result proved to be alike as each of these parameters showed weak correlation with corneal permeability, only the following tendencies could be observed: the corneal permeability decreased with increasing TPSA while it increased with increasing $\log P$ and $\log D_{pH7.4}$ values. However, the molecular weight seemed to have no effect on corneal permeability. The neutral and basic compounds exhibited higher permeability values than the acidic or amphoteric ones, which observation corresponds to the general trend for membrane permeability of drugs it as has also been reported previously [46]. The best squared correlation coefficient ($R^2 = 0.356$) was observed in the case of $\log D_{pH7.4,}$ which is probably due to the fact that the measurements were carried out at iso-pH 7.4 conditions.

The relationship between membrane retention and these physicochemical parameters has also been investigated. In the case of TPSA and $\log P$ only trends could be observed with low squared correlation coefficients ($R^2 = 0.289$ and 0.367 respectively). The MR values increased with increasing $\log P$ and decreased with increasing TPSA, which is in accordance with the fact that the more lipophilic (and less polar) the API is, the more likely it gets into the membrane. The analysis resulted in very poor fits with the other two parameters. (*See Figure B1., Appendix B*). The corneal-PAMPA permeability values were also compared with other permeability-related *in silico* values like Caco-2 permeability (pH 7.4, at 350 rpm stirring rate) jejunal permeability [47] and blood-brain partition coefficient [48] (logBB), which were predicted by the ACD/Percepta software [32] as a gold standard tool in drug discovery (Fig. 3).

The Caco-2 model applies human epithelial colorectal adenocarcinoma cells thus, unlike the PAMPA model, it contains



MW – molecular weight, TPSA – topological polar surface area, clog$P$ – calculated log partition coefficient, clog$D_{pH7.4}$ – calculated log distribution coefficient at pH 7.4

Fig. 1. Histograms of physicochemical properties of 189 compounds. Physicochemical properties include molecular weight, topological polar surface area, calculated log partition coefficient, calculated log distribution coefficient at pH 7.4, and acid-base characteristics. The TPSA, clog$P$ and clog$D_{pH7.4}$ values were calculated with the ACD/Percepta software [32] and the acid-base characteristics were determined with the ChemAxon/MarvinSketch software [44].

active transporters. For comparing the corneal-PAMPA and Caco-2 permeability values, 9 APIs (hydrocortisone, emetine, etoposide, loperamide, quinine, risperidone, sparfloxacin, trimethoprim, verapamil) were labeled with red stars on Fig. 3, which are potential substrates of P-gp efflux pump according to the prediction of ACD/Percepta software [32]. For the proper comparison between the different assays, compounds that are known to be penetrating via active transport should be avoided. To clarify the differences in their penetration behavior, squared correlation coefficients are represented here, which refer to the datasets without (R^2) and with (R^{2*}) the P-gp substrates aided by active transport. Similarly, to our former conclusion, only a weak correlation could be determined between corneal-PAMPA and Caco-2 permeability values. A weak correlation was found in the case of jejunal permeability and blood-brain partition coefficient too, which indicates that the corneal-PAMPA is not a universal permeability model and that the trans-corneal diffusion process is presumably governed by different factors than intestinal or blood-brain-barrier penetration. Thus, the results of the preliminary analysis clearly verified the need for a unique QSPR model for corneal-PAMPA predictions.

3.2. Corneal-PAMPA permeability QSPR prediction

In the model-building phase, the original matrix contained 3302 molecular descriptors (including the predicted physicochemical ones, classical 1, 2 and 3D descriptors and the ECFP fingerprint) as X variables, and corneal-PAMPA experimental results as the Y vector. The total number of APIs where corneal-PAMPA experimental values could be determined was 151. Outlier selection was performed on the dataset: a) first the APIs outside of the well-covered $0–50 \times 10^{-6}$ cm/s range were excluded, b) then the first two PLS components were plotted against each other and 95 % confidence ellipse was calculated to exclude the outliers based on the molecular descriptors. In the final form, 143 APIs were included in the dataset.

In the next step, the matrix was split into training and test set, which was obligatory to carry out a proper validation of the model. With the DOOD algorithm, 80 % of the compounds were selected for the training set. Thus, 114 APIs were applied for the calibration and cross-validation of the model. Autoscaling (standardization) was used as data pretreatment. We have selected 100 molecular descriptors with the combination of VIP scoring and genetic algorithm from the original descriptor set. In the 7-fold cross-validation part, RMSECV values were used for the determination of the number of PLS components. The global minimum of the RMSECV values was at five PLS components. Cross-validation was performed with a randomized selection in 20 iterations. The final calibration curve along with the test validation samples based on 5 PLS components is shown in Fig. 4 ab.

The samples were well-fitted to the calibration curve and the errors were homogenous along the range. It is clearly seen in the performances as well: the R^2 (goodness of fit for the calibration) and Q^2_{CV} (goodness of fit for the cross-validation) were 0.953 and 0.866, respectively. The performance of the test validation was 0.862. The remaining performance parameters of the model are summarized in Table 2.

Out of the selected 100 variables, the most important physicochemical descriptors (based on their VIP scores) were the following: absorption rate constant (ka), jejunal permeability (clog$P_{e\ jejunal}$,pH = 7.4), fraction unbound in the brain (f_{bu}),

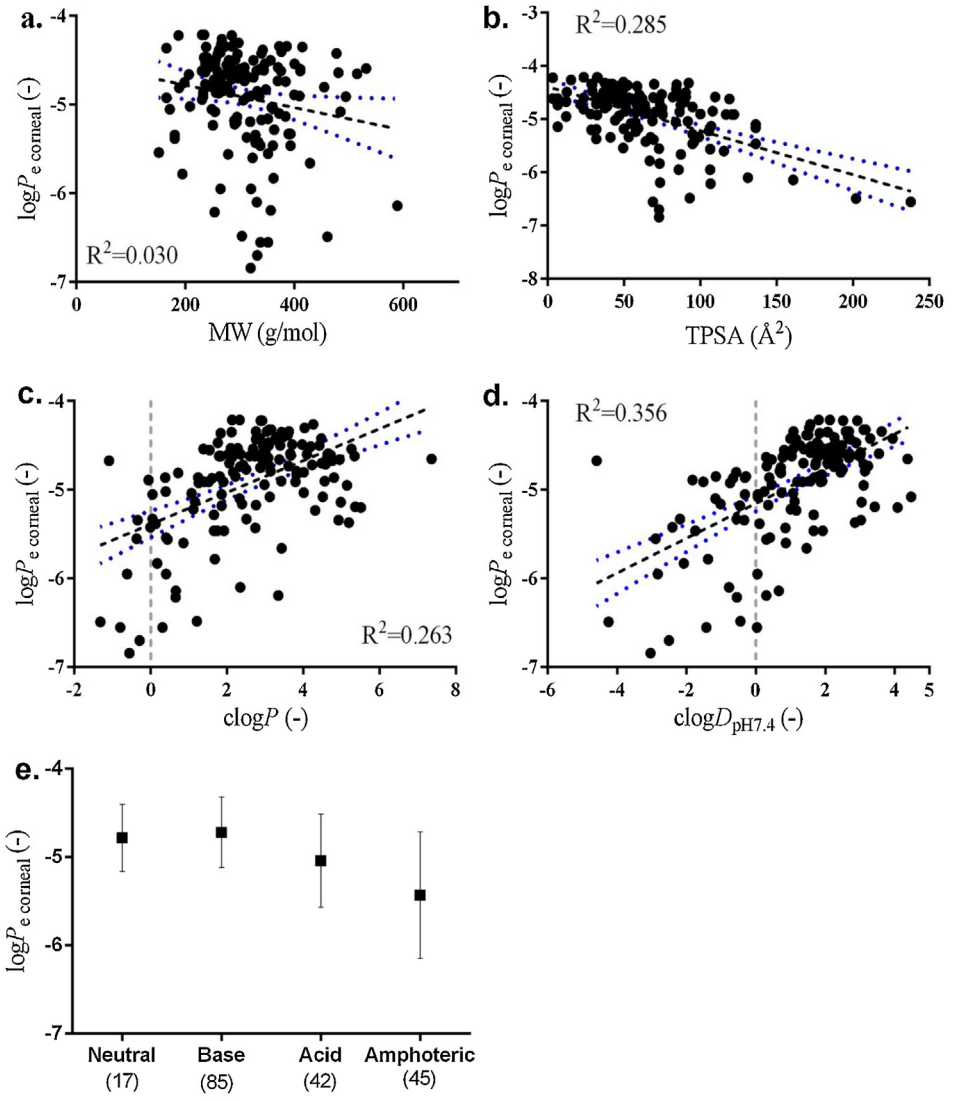


Fig. 2. PAMPA-related corneal permeability as a function of physicochemical properties. Correlation between corneal permeability and molecular weight (a.), topological polar surface area (b.), calculated log partition coefficient (c.), calculated log distribution coefficient at pH 7.4 (d.) and acid-base characteristics (e.) of the compounds with the squared correlation coefficient (R^2). TPSA, clog*P* and clog$D_{pH7.4}$ values were calculated with the ACD/Percepta software [32] and the acid-base characteristics were determined by the ChemAxon/MarvinSketch software [44]. The correlation coefficients were determined using the GraphPad Prism software [45]. The blue dotted lines stand for the 95 % confidence bands.

clog$D_{pH7.4}$, plasma protein binding (PPB). All these descriptors had VIP (variable importance) scores above 1.3. The results correspond to the kinetic behavior of the permeability parameter as these descriptors are related to passive transport of drugs. In addition, the fact that predicted data of neither the gastrointestinal (k_a, clog$P_{e\ jejunal}$), nor the blood-brain barrier transport (f_{bu}) in silico models could describe the corneal-PAMPA permeability process, suggests the complexity of our model, which is also supported by our conclusions in the preliminary analysis. Ionization state and lipophilicity are also well-known determinants of the permeability process [46,49,50], which are also present here in the form of clog$D_{pH7.4}$. Additionally, linear regression analysis was carried out in the case of k_a, PPB and f_{bu} against logP_e and only a weak correlation was found in each case, which supports our previous observation that using linear regression of only one parameter at a time could not result in a good model for prediction (*See Figure B2., Appendix B and Table A2., Appendix A).*

Yscrambling was also used for the validation of the model and to check the possibility of overfitting. The experimental corneal-PAMPA values were randomized in the dataset for this validation. The prediction performances at 5 PLS components were decreased properly to 0.016 and 0.14 for cross-validation and test validation, respectively. The permutation test (with 50 iterations) has found a significant difference between our final model and the randomized one (at $\alpha = 0.05$ level). For the result of the permutation see *Table A3, Appendix A.*

3.3. Corneal-PAMPA membrane retention QSPR prediction

In the case of membrane retention prediction, the same protocol (with the same parameters) was carried out as in the case of permeability prediction. The selected range of membrane retention was between 0 and 100 %. After the outlier selection process, 180 APIs out of 189 were included in the dataset. The remaining

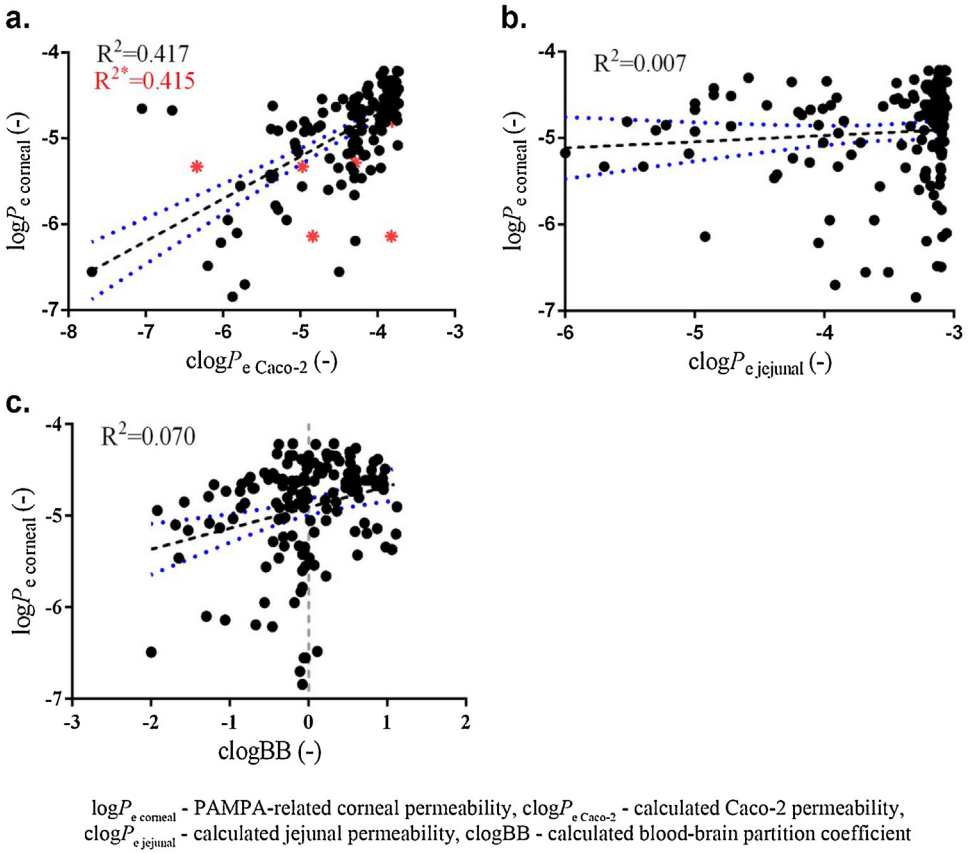


Fig. 3. PAMPA-related corneal permeability as a function of various in silico transport parameters. Correlation between corneal permeability and calculated Caco-2 permeability (pH 7.4, at 350 rpm stirring rate) (a.), calculated jejunal permeability (b.) [47] and calculated blood-brain partition coefficient (c.) [48] of the compounds with squared correlation coefficients (R^2). In the case of Caco-2 permeability (a.) recalculated squared correlation coefficient (R^{2*}) was also introduced including nine P-gp substrates. clog$P_{e\ Caco\text{-}2}$, clog$P_{e\ jejunal}$ and clogBB values were calculated with the ACD/Percepta software [32]. The correlation coefficients were determined using the GraphPad Prism software [45]. The blue dotted lines stand for the 95 % confidence bands.

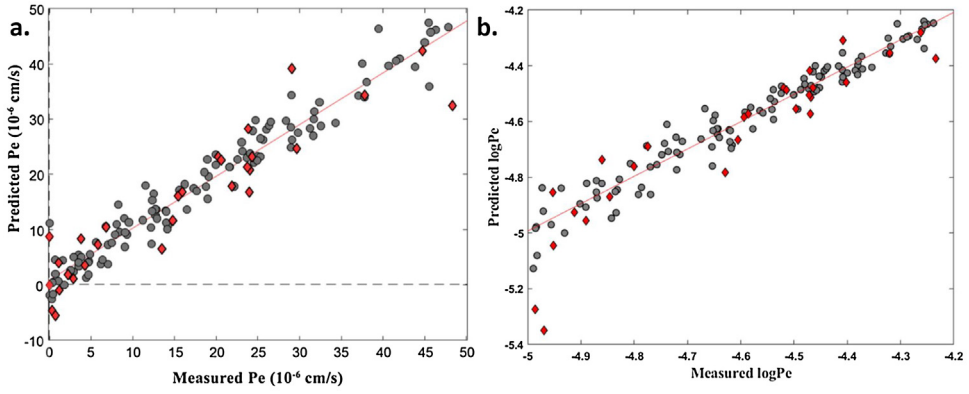


Fig. 4. a b. The PLS calibration curve. a) Predicted corneal-PAMPA values are plotted against the experimental ones. In Fig. 4b) the model is transformed to logarithmic scale. Calibration samples are marked with grey circles, while test validation compounds are marked with red diamonds.

nine compounds were excluded due to their negative membrane retention values (they were out of the 0–100 % range). In total, 83 molecular descriptors were included in the matrix after the variable selection process (VIP scores and genetic algorithm). Classical physicochemical, two-dimensional, three-dimensional descriptors and fingerprint bit positions were included in the selected set of variables. The same DOOD algorithm was used for the training-test split. The training set contained 144 compounds (80 %), while the remaining 36 compounds were used for test validation. Seven-fold cross-validation was used for the internal validation of the model with 20 iterations, and eight PLS components were applied based on the global minimum of the RMSECV curve. The final calibration model with the calibration and test samples together is plotted in Fig. 5. The samples were well-fitted to the calibration curve (even the test ones) and the errors were homogeneous along the range. The R^2 value of the calibration was 0.97, while the Q^2_{CV} and Q^2 prediction were 0.91 and 0.94, respectively. All the performance parameters are summarized in Table 2.

The most important physicochemical properties amongst the 83 selected descriptors (based on their VIP scores) were the following: log aqueous solubility (log*S*), log*P* and PPB. These descriptors had VIP scores above 1.3. The result confirms the thermodynamic

Table 2
Summary of the performance parameters for the final QSPR models*.

Model	Permeability	Membrane retention	Permeability (in log sale)
RMSEC	2.87	4.37	0.04
RMSECV	5.06	8.14	0.08
RMSEP	5.15	7.05	0.11
R^2	0.95	0.97	0.95
Q^2_{CV}	0.87	0.91	0.86
R^2_P	0.86	0.94	0.85
Q^2_{F1}	0.86	0.96	0.75
Q^2_{F2}	0.86	0.93	0.75
Q^2_{F3}	0.85	0.93	0.79
CCC_{Train}	0.98	0.99	0.97
CCC_{CV}	0.93	0.95	0.93
CCC_{Ext}	0.92	0.97	0.89
MAE_{Train}	2.15	3.28	0.04
MAE_{CV}	3.79	4.90	0.06
MAE_{Ext}	3.92	5.74	0.08
r^2_m	0.80	0.89	0.58
$\overline{r^2_m}$	0.76	0.88	0.65
Δr^2_m	0.07	0.01	0.13
$\left\|r^2_0 - r'^2_0\right\|$	0.02	0.00	0.07
k	0.87	0.96	1.09
k'	0.91	0.99	1.01
Y-SCR CV	0.02	0.02	–
Y-SCR P.	0.14	0.04	–
PLS COMP.	5	8	–

* RMSEC – root mean squared error of calibration, RMSECV – root mean squared error of cross-validation, RMSEP – root mean squared error of external validation, R^2 - goodness of fit for the calibration, Q^2_{CV} - goodness of fit for the cross-validation, R^2_P - goodness of fit for the external validation, $Q_{F1}{}^2$ – F1 type goodness of fit for the external validation, $Q_{F2}{}^2$ – F2 type goodness of fit for the external validation, $Q_{F3}{}^2$ – F3 type goodness of fit for the external validation, CCC (Train/CV/Ext) – coefficient of concordance for the training/CV/external test set, MAE (Train/CV/Ext) – mean absolute error for the training/CV/external test set, r^2_m and $\Delta\ r^2_m$ – based on the work of Roy et al. [52], $\left|r^2_0 - r'^2_0\right|$ is the absolute difference of the r^2_0 values without intercept in the case of observed vs. predicted and predicted vs. observed values, k and k' - the slope of the regression lines (with or without passing through the origin), Y-SCR CV - goodness of fit for the cross validation using Y scrambling,Y-SCR P-goodness of fit for the test prediction using Y-scrambling, PLS COMP. – number of PLS components in the model.

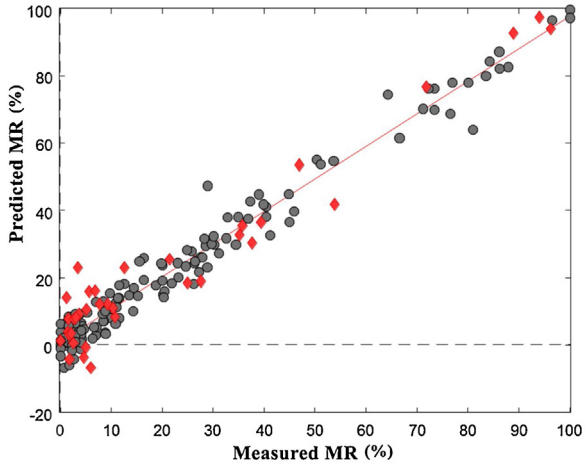


Fig. 5. Prediction of the membrane retention. Predicted values are plotted against the experimental membrane retention values. The test samples are marked with red diamonds and the calibration ones are marked with grey circles. The calibration curve is marked with the red line.

behavior of the MR parameter for the tested compounds in that it relates to their distribution between the donor, acceptor and lipid membrane. The primary relevance of the log*P* and log*S* parameters suggests the dominance of the neutral form of the compounds, which also supports the specificity of the *in silico* description, as the proton dissociation of drug molecules in the membrane is typically depressed. The fact that log*P* is more favorable than $\log D_{pH7.4}$ as a descriptor for membrane retention confirms the results of our preliminary analysis (*See Figure B1., Appendix B*). The weighted appearance of PPB can be explained by the lipophilicity-dependent nature of the free form of the drug, which is consistent with the general kinetic description of drug transport. Additionally, linear regression analysis was carried out in the case of log*S* and PPB against MR and only weak correlations were found with the squared correlation coefficients of 0.232 and 0.189 respectively which also verifies the need for more complex models than the linear regression of one physicochemical parameter at a time (*See Figure B3., Appendix B and Table A2., Appendix A*).

Y-scrambling was also used to check the possibility of overfitting. The model performances strongly decreased with the use of randomized Y values. The Q^2_{CV} was 0.02, while the Q^2 for test prediction was 0.04. The permutation test (with 50 iterations) has found a significant difference between our final model and the randomized one (at $\alpha = 0.05$ level), verifying our model.

The detailed summary of the performance parameters of the final models (permeability and membrane retention) is shown in Table 2. All parameters are good and fulfill the suggested conditions for the acceptance of the model: $R^2 > 0.7$, $Q_{Fn}{}^2 > 0.6$, CCC > 0.85, $r^2_m > 0.5$, $\Delta\ r^2_m < 0.2$ [42,51], and $\left|r^2_0 - r'^2_0\right| < 0.2$. Based on the performances of the models, we are recommending the use of permeability values in linear scale for further calculations.

4. Conclusion

Two quantitative structure-property relationship models were developed for the investigation of corneal permeability and membrane retention. The QSPR models were based on experimentally determined corneal-PAMPA values. Partial least squares regression was applied for the process. The final models contained classical molecular descriptors and ECFP fingerprints in combination. The models included 143 and 180 APIs for P_e and MR prediction, respectively. A complex validation protocol (including internal and external phases) was carried out to provide robust and appropriate predictions for the permeability and membrane retention values. The goodness of the models (R^2) were above 0.90 in both cases and they were not lower than 0.85 for validation either. The permutation tests showed that our models are not overfitted and they are clearly applicable to provide good predictions for the corneal permeability of the APIs. This is, to our knowledge, the first such in silico model reported so far. The results also showed that there was no significant correlation between the experimental corneal-PAMPA values and the existing classical in silico pharmacokinetically relevant drug transport models, such as Caco-2 permeability, jejunal permeability and the blood-brain partition coefficient (logBB). Thus, our models can be good candidates to provide the corneal permeability and corneal membrane retention of compounds in a very short time with appropriate precision.

Authors' contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Declaration of Competing Interest

The authors report no declarations of interest.

Acknowledgements

The authors thank Károly Héberger for the useful discussions and scientific advice. This work was supported by the National Research, Development and Innovation Office of Hungary (under grant number K 119269, PD 134416) and by the Richter Gedeon Excellence PhD Scholarship of Richter Gedeon Foundation.

Appendix F. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.jpba.2021.114218.

References

[1] P. Baranowski, B. Karolewicz, M. Gajda, J. Pluta, Ophthalmic drug dosage forms: characterisation and research methods, Sci. World J. (2014) 861–904, http://dx.doi.org/10.1155/2014/861904.
[2] F. Rafiei, H. Tabesh, F. Farzad, Sustained subconjunctival drug delivery systems: current trends and future perspectives, Int. Ophthalmol. 40 (2020) 2385–2401, http://dx.doi.org/10.1007/s10792-020-01391-8.
[3] F.J. Cabrera, D.C. Wang, K. Reddy, G. Acharya, C.S. Shin, Challenges and opportunities for drug delivery to the posterior of the eye, Drug Discov. Today 24 (2019) 1679–1684, http://dx.doi.org/10.1016/j.drudis.2019.05.035.
[4] A. Urtti, Challenges and obstacles of ocular pharmacokinetics and drug delivery, Adv. Drug Deliv. Rev. 58 (2006) 1131–1135, http://dx.doi.org/10.1016/j.addr.2006.07.027.
[5] J. Barar, A.R. Javadzadeh, Y. Omidi, Ocular novel drug delivery: impacts of membranes and barriers, Expert Opin. Drug Deliv. 5 (2008) 567–581, http://dx.doi.org/10.1517/17425247.5.5.567.
[6] S.A. Molokhia, S.C. Thomas, K.J. Garff, K.J. Mandell, B.M. Wirostko, Anterior eye segment drug delivery systems : current treatments and future challenges, J. Ocul. Pharmacol. Ther. 29 (2013) 92–105, http://dx.doi.org/10.1089/jop.2012.0241.
[7] K. Järvinen, T. Järvinen, A. Urtti, Ocular absorption following topical delivery, Adv. Drug Deliv. Rev. 16 (1995) 3–19, http://dx.doi.org/10.1016/0169-409X(95)00010-5.
[8] I. Pepić, J. Lovrić, B. Cetina-Čižmek, S. Reichl, J. Filipović-Grčić, Toward the practical implementation of eye-related bioavailability prediction models, Drug Discov. Today 19 (2014) 31–44, http://dx.doi.org/10.1016/j.drudis.2013.08.002.
[9] E. Moisseiev, A. Loewenstein, Drug delivery to the posterior segment of the eye, Dev. Ophthalmol. 58 (2017) 87–101, http://dx.doi.org/10.1159/000455276.
[10] J. Barar, M. Asadi, S.A. Mortazavi-Tabatabaei, Y. Omidi, Ocular drug delivery; impact of in vitro cell culture models, J. Ophthalmic Vis. Res. 4 (2009) 238–252 http://www.ncbi.nlm.nih.gov/pubmed/23198080%5Cnhttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3498862.
[11] E. Eljarrat-Binstock, J. Pe'er, A.J. Domb, New techniques for drug delivery to the posterior eye segment, Pharm. Res. 27 (2010) 530–543, http://dx.doi.org/10.1007/s11095-009-0042-9.
[12] H. Kidron, K.S. Vellonen, E.M. Del Amo, A. Tissari, A. Urtti, Prediction of the corneal permeability of drug-like compounds, Pharm. Res. 27 (2010) 1398–1407, http://dx.doi.org/10.1007/s11095-010-0132-8.
[13] P. Agarwal, I.D. Rupenthal, In vitro and ex vivo corneal penetration and absorption models, Drug Deliv. Transl. Res. 6 (2016) 634–647, http://dx.doi.org/10.1007/s13346-015-0275-6.
[14] E.A. Offord, N.A. Sharif, K. Mace, Y. Tromvoukis, E.A. Spillare, O. Avanti, W.E. Howe, A.M.A. Pfeifer, Immortalized human corneal epithelial cells for ocular toxicity and inflammation studies, Invest. Ophthalmol. Vis. Sci. 40 (1999) 1091–1101.
[15] R.D. Schoenwald, R.L. Ward, Relationship between steroid permeability across excised rabbit cornea and octanol-water partition coefficients, J. Pharm. Sci. 67 (1978) 786–788, http://dx.doi.org/10.1002/jps.2600670614.
[16] G.L. Mosher, T.J. Mikkelson, Permeability of the n-alkyl p-aminobenzoate esters across the isolated corneal membrane of the rabbit, Int. J. Pharm. 2 (1979) 239–243, http://dx.doi.org/10.1016/0378-5173(79)90023-1.
[17] R.D. Schoenwald, H. -S Huang, Corneal penetration behavior of β-blocking agents I: physicochemical factors, J. Pharm. Sci. 72 (1983) 1266–1272, http://dx.doi.org/10.1002/jps.2600721108.
[18] W. Wang, H. Sasaki, D.-S. Chien, V.H.L. Lee, Lipophilicity influence on conjunctival drug penetration in the pigmented rabbit: A comparison with corneal penetration, Curr. Eye Res. 10 (1991) 571–579, http://dx.doi.org/10.1080/02713680701435391.
[19] F. Yoshida, J.G. Topliss, Unified model for the corneal permeability of related and diverse compounds with respect to their physicochemical properties, J. Pharm. Sci. 85 (1996) 819–823, http://dx.doi.org/10.1021/js960076m.
[20] X.C. Fu, W.Q. Liang, A simple model for the prediction of corneal permeability, Int. J. Pharm. 232 (2002) 193–197, http://dx.doi.org/10.1016/S0378-5173(01)00902-4.
[21] K. Roy, S. Kar, R.N. Das, A Primer on QSAR/QSPR Modeling: Fundamental Concepts, 2015, http://dx.doi.org/10.1007/978-3-319-17281-1.
[22] A. Rácz, G.M. Keserű, Large-scale evaluation of cytochrome P450 2C9 mediated drug interaction potential with machine learning-based consensus modeling, J. Comput. Aided Mol. Des. 34 (2020) 831–839, http://dx.doi.org/10.1007/s10822-020-00308-y.
[23] M. Ghorbanzad'E, M.H. Fatemi, M. Karimpour, P.L. Andersson, Quantitative and qualitative prediction of corneal permeability for drug-like compounds, Talanta 85 (2011) 2686–2694, http://dx.doi.org/10.1016/j.talanta.2011.08.060.
[24] E. Ramsay, E.M. del Amo, E. Toropainen, U. Tengvall-Unadike, V.P. Ranta, A. Urtti, M. Ruponen, Corneal and conjunctival drug permeability: Systematic comparison and pharmacokinetic impact in the eye, Eur. J. Pharm. Sci. 119 (2018) 83–89, http://dx.doi.org/10.1016/j.ejps.2018.03.034.
[25] G. Dargó, A. Vincze, J. Müller, H.J. Kiss, Z.Z. Nagy, G.T. Balogh, Corneal-PAMPA: A novel, non-cell-based assay for prediction of corneal drug permeability, Eur. J. Pharm. Sci. 128 (2019) 232–239, http://dx.doi.org/10.1016/j.ejps.2018.12.012.
[26] A. Avdeef, Permeability-PAMPA, in: Absorpt. Drug Dev., John Wiley & Sons, Inc., Hoboken, NJ, USA, 2012, pp. 319–498, http://dx.doi.org/10.1002/9781118286067, ch7.
[27] ChemAxon Calculator, version: 18.1.0, 2020, Accessed on Sep 15, 2020 https://chemaxon.com/.
[28] Schrödinger Release 2019-4: LigPrep, Schrödinger, LLC, New York, NY, 2019, Accessed on Sep 22, 2020, (2020) https://www.schrodinger.com/ligprep.
[29] G. Madhavi Sastry, M. Adzhigirey, T. Day, R. Annabhimoju, W. Sherman, Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments, J. Comput. Aided Mol. Des. 27 (2013) 221–234, http://dx.doi.org/10.1007/s10822-013-9644-8.
[30] Dragon, version:7.0, 2020, Accessed on Sep 12, 2020 https://chm.kode-solutions.net/.
[31] R. Todeschini, V. Consonni, Handbook of Molecular Descriptors, WILEY-VCH Verlag, Weinheim, Germany, 2000.
[32] ACD/Labs Percepta, version: 2019.1.0, 2020, Accessed: Sep 1, 2020 https://www.acdlabs.com/products/percepta/.
[33] A. Rácz, D. Bajusz, K. Héberger, Intercorrelation Limits in Molecular Descriptor Preselection for QSAR/QSPR, Mol. Inform. 38 (2019) 2–7, http://dx.doi.org/10.1002/minf.201800154.
[34] D. Bajusz, A. Rácz, K. Héberger, Chemical data formats, fingerprints and other molecular descriptors for database analysis and searching, in: S. Chackalamannil, D. Rotella, S.E. Ward (Eds.), Compr. Med. Chem. III., Elsevier, Oxford, 2017, pp. 329–378.
[35] P. Geladi, B.R. Kowalski, Partial least-squares regression: a tutorial, Anal. Chim. Acta 185 (1986) 1–17.
[36] C.M. Andersen, R. Bro, Variable selection in regression-a tutorial, J. Chemom. 24 (2010) 728–737, http://dx.doi.org/10.1002/cem.1360.
[37] PLS Toolbox, Eigenvector research, Inc., 2020, Accessed on Oct 5, 2020 https://eigenvector.com/software/pls-toolbox/.
[38] P. Gramatica, Principles of QSAR models validation: internal and external, QSAR Comb. Sci. 26 (2007) 694–701, http://dx.doi.org/10.1002/qsar.200610151.
[39] I.M. Olsson, J. Gottfries, S. Wold, D-optimal onion designs in statistical molecular design, Chemometr. Intell. Lab. Syst. 73 (2004) 37–46, http://dx.doi.org/10.1016/j.chemolab.2004.04.001.
[40] T. Hastie, R. Tibshirani, J. Friedman, Cross-validation, in: Elem. Stastistical Learn. Data Mining, Interf. Predict., 2nd ed., Springer New York LLC, 2009, pp. 241–249.
[41] V. Consonni, D. Ballabio, R. Todeschini, Evaluation of model predictive ability by external validation techniques, J. Chemom. 24 (2010) 194–201, http://dx.doi.org/10.1002/cem.1290.
[42] N. Chirico, P. Gramatica, Real external predictivity of QSAR models. Part 2. New intercomparable thresholds for different validation criteria and the need for scatter plot inspection, J. Chem. Inf. Model. 52 (2012) 2044–2058, http://dx.doi.org/10.1021/ci300084j.
[43] A. Vincze, G. Dargó, G.T. Balogh, Cornea-PAMPA as an orthogonal in vitro physicochemical model of corneal permeability, Period. Polytech. Chem. Eng. 64 (2020) 384–390, http://dx.doi.org/10.3311/PPch.15601.
[44] MarvinSketch, version: 20.10, 2020, Accessed on Sep 15, 2020 https://chemaxon.com/products/marvin/.
[45] GraphPad Prism, version: 7.03, 2020, Accessed: Oct 1, 2020 https://www.graphpad.com/scientific-software/prism/.
[46] H. Pham-The, I. González-Álvarez, M. Bermejo, T. Garrigues, H. Le-Thi-Thu, M.Á. Cabrera-Pérez, The use of rule-based and QSPR approaches in ADME profiling: a case study on caco-2 permeability, Mol. Inform. 32 (2013) 459–479, http://dx.doi.org/10.1002/minf.201200166.
[47] D.P. Reynolds, K. Lanevskij, P. Japertas, R. Didziapetris, A. Petrauskas, Ionization-specific analysis of human intestinal absorption, J. Pharm. Sci. 98 (2009) 4039–4054, http://dx.doi.org/10.1002/jps.21730.
[48] K. Lanevskij, J. Dapkunas, L. Juska, P. Japertas, R. Didziapetris, QSAR analysis of blood-brain distribution: the influence of plasma and brain tissue binding, J. Pharm. Sci. 100 (2011) 2147–2160, http://dx.doi.org/10.1002/jps.22442.
[49] A. Avdeef, Relationship between permeability and partition coefficient, in: Absorpt. Drug Dev., 2nd ed., 2012, pp. 333–335, New Jersey.
[50] E.C. Sherer, A. Verras, M. Madeira, W.K. Hagmann, R.P. Sheridan, D. Roberts, K. Bleasby, W.D. Cornell, QSAR prediction of passive permeability in the LLC-PK1 cell line: Trends in molecular properties and cross-prediction of Caco-2

permeabilities, Mol. Inform. 31 (2012) 231–245, http://dx.doi.org/10.1002/minf.201100157.
[51] P. Gramatica, A. Sangion, A historical excursus on the statistical validation parameters for QSAR models: a clarification concerning metrics and terminology, J. Chem. Inf. Model. 56 (2016) 1127–1131, http://dx.doi.org/10.1021/acs.jcim.6b00088.
[52] K. Roy, On some aspects of validation of predictive quantitative structure-activity relationship models, Expert Opin. Drug Discov. 2 (2007) 1567–1577, http://dx.doi.org/10.1517/17460441.2.12.1567.