International Journal of Pharmaceutics 522 (2017) 222–233

Contents lists available at ScienceDirect

International Journal of Pharmaceutics

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

Full-length article

Quantitative prediction of ionization effect on human skin permeability

Hiromi Baba[a,b,*], Yusuke Ueno[c], Mitsuru Hashida[a,d], Fumiyoshi Yamashita[a,**]

[a] Department of Drug Delivery Research, Graduate School of Pharmaceutical Sciences, Kyoto University, 46-29, Yoshida-shimoadachicho, Sakyo-ku, Kyoto 606-8501 Japan
[b] Drug Discovery Research Laboratories, Kyoto R&D Center, Maruho Co., Ltd., 93 Awata-cho, Chudoji, Shimogyo-ku, Kyoto, Japan
[c] Drug Development Research Laboratories, Kyoto R&D Center, Maruho Co., Ltd., 93 Awata-cho, Chudoji, Shimogyo-ku, Kyoto, Japan
[d] Institute for Integrated Cell-Material Sciences, Kyoto University, 46-29, Yoshida-shimoadachicho, Sakyo-ku, Kyoto 606-8501 Japan

ARTICLE INFO

Article history:
Received 21 January 2017
Received in revised form 5 March 2017
Accepted 6 March 2017
Available online 7 March 2017

Keywords:
Skin permeability
Transdermal absorption
Ionization effect
Support vector regression
Gaussian process
QSPR

ABSTRACT

Although skin permeability of an active ingredient can be severely affected by its ionization in a dose solution, most of the existing prediction models cannot predict such impacts. To provide reliable predictors, we curated a novel large dataset of *in vitro* human skin permeability coefficients for 322 entries comprising chemically diverse permeants whose ionization fractions can be calculated. Subsequently, we generated thousands of computational descriptors, including Log D (octanol–water distribution coefficient at a specific pH), and analyzed the dataset using nonlinear support vector regression (SVR) and Gaussian process regression (GPR) combined with greedy descriptor selection. The SVR model was slightly superior to the GPR model, with externally validated squared correlation coefficient, root mean square error, and mean absolute error values of 0.94, 0.29, and 0.21, respectively. These models indicate that Log D is effective for a comprehensive prediction of ionization effects on skin permeability. In addition, the proposed models satisfied the statistical criteria endorsed in recent model validation studies. These models can evaluate virtually generated compounds at any pH; therefore, they can be used for high-throughput evaluations of numerous active ingredients and optimization of their skin permeability with respect to permeant ionization.

1. Introduction

The skin is a unique and elaborate biological interface to the internal physiological system. Although the skin resists the uncontrolled invasion of xenobiotic compounds, it acts an essential administration route for various pharmaceutical and cosmetic formulations (Prausnitz et al., 2004). To obtain an appropriate rate of absorption of active ingredients from such topical formulations is required for regulating their efficacy and toxic risks.

Various *in vitro*, particularly diffusion studies using excised human, animal, or artificial skin membranes, and *in vivo* techniques have been used extensively to evaluate the skin permeability of chemicals (Godin and Touitou, 2007). However, such experiments require chemical substances of interest and are generally time- and cost-intensive. Furthermore, some potential ethical issues may arise in *in vivo* studies and experiments that use human or animal-derived tissues.

Thus, the past few decades have seen the development of various *in silico* prediction models (Baba et al., 2015a,b; Baert et al., 2007; Khajeh and Modarress, 2014; Lim et al., 2002; Moss et al., 2009; Neely et al., 2009; Neumann et al., 2006; Potts and Guy, 1992) of skin permeability based on the concepts of quantitative structure–property relationships (QSPR), which are generally free of the above-mentioned problems associated with "wet" experiments. Fully computational prediction models are considerably advantageous in terms of time and cost and can evaluate the yet-

* Corresponding author at: Department of Drug Delivery Research, Graduate School of Pharmaceutical Sciences, Kyoto University, Sakyo-ku, Kyoto 606-8501 Japan.
** Corresponding author.
E-mail addresses: baba_dfq@mii.maruho.co.jp (H. Baba), yama@pharm.kyoto-u.ac.jp (F. Yamashita).

http://dx.doi.org/10.1016/j.ijpharm.2017.03.009
0378-5173/

to-be-synthesized or virtually generated permeants. These characteristics make such models practically suitable for evaluating, screening, and optimizing numerous candidate ingredients of topical medicines and cosmetics in terms of skin permeability.

One of the most widely used prediction models of skin permeability is the model of Potts and Guy (1992), which describes a linear relationship of permeability coefficient (k_p) of permeants with their molecular weights (MW) and the octanol-water partition coefficient (Log P). Predictability of skin permeability has hereafter been improved by various *in silico* models based on sophisticated machine learning techniques, such as artificial neural networks (Lim et al., 2002; Neely et al., 2009), regression trees (Baert et al., 2007), *k*-nearest neighbor (Neumann et al., 2006), Gaussian process regression (GPR) (Moss et al., 2009), random forest (Baba et al., 2015a), and support vector regression (SVR) (Baba et al., 2015a), which can describe the nonlinear relationships among permeants' skin permeability and their molecular characteristics.

However, it should be noted that most existing prediction models, including the Potts and Guy model, do not deal with the impact of permeant ionization in a dose solution on skin permeability; this is because such models either have been established using only the permeability of permeants in their unionized forms, or they have ignored the permeant dissociation that can occur in the actual solution. Generally, the permeability of charged (ionized) permeants is much lower than that of the unionized form. It has been reported that the permeability coefficient of a permeant can be altered more than tens of times, or even one hundred times, by changing the ionization state (Chantasart et al., 2015; Roy and Flynn, 1990; Smith and Irwin, 2000). The ionization of molecules in a solution depends on their dissociation constant (*e.g.*, pKa in the case of acidic compounds) and the pH of the solution; therefore, the pH of the dose solution can greatly impact skin permeability.

Some studies have reported on the impact of permeant ionization on skin permeability, although these have been limited to specific types of permeants or chemical groups, such as salicylic acid (Smith and Irwin, 2000), ß-adrenergic blockers with amino groups (Roy and Flynn, 1990), and non-steroidal anti-inflammatory drugs with carboxylic acids (Chantasart et al., 2015). To the best of our knowledge, there has been only one QSPR model dealing with the ionization effect on human skin permeability. Zhang et al. (2012) compiled a moderate size of the database (118 entries) and built the QSPR model by applying an Abraham's linear free-energy relationship (Abraham and Acree, 2010) with the original terms related to the anion and cation. Their model was cutting-edge, but a complete external validation was not performed. In QSPR modeling, proper external validation is indispensable because the predictive power of the model cannot simply be assessed by the goodness-of-fit or internal cross-validation (Golbraikh and Tropsha, 2002).

In this study, we propose reliable and practical computational predictors for actual human skin permeability affected by the ionization of permeants in their dose solutions. To this end, we take the following four steps: (1) we compile a dataset of the human skin permeability of structurally diverse permeants whose ionization fractions can be estimated—this covers a sufficiently large chemical space spanned for topical drug candidates; (2) we generate computational molecular descriptors as explanatory variables, including factors encoding the ionization effects of permeants; (3) we then apply nonlinear machine learning regression techniques (*i.e.*, SVR and GPR) to the newly generated dataset; and (4) we validate the constructed models internally and externally by the standards that have recently been endorsed.

2. Materials and methods

2.1. Database development

2.1.1. Collection of permeability data

We compiled 322 permeability coefficients (Log k_p) of 203 different permeants measured in *in vitro* diffusion studies on human skin from the literature. During the compilation process, the permeability data from early studies of Scheuplein et al. (1969) were not included in our database because substantial discrepancies between their values and those of other groups have been reported (Johnson et al., 1995). To quantitatively evaluate the effect of permeant ionization on permeability, we collected permeability coefficients for the different ionization states of some individual compounds. To ensure that the ionization fraction of the permeants can be calculated for each entry, we compiled permeability data with the pH values of the aqueous dose solution. Furthermore, we collected permeability of non-electrolytic permeants (*e.g.*, anisole, cortisone, chloroform, and triamcinolone) even without reported pH values because such permeants expect to be unionized in ordinary topical solutions. The dataset with permeants, their human skin permeability coefficients, CAS registry numbers, dose solution pH, and data references is provided as Supplementary material.

To ensure the quality of the dataset for directly analyzing the ionization effects of permeants on human skin permeability, we rigorously screened the data to satisfy the following selection criteria, which were modified from our previous studies (Baba et al., 2015b):

(i) permeability coefficients should be obtained through *in vitro* diffusion studies;
(ii) excised human skin should be used as a diffusion membrane;
(iii) the skin should not be pre-treated with any vehicles except for aqueous solvents to hydrate the skin;
(iv) no physical or chemical penetration enhancing techniques should be employed.

If the permeability coefficients (k_p) were not specified in diffusion studies, they were obtained from the fluxes at steady state (J_{ss}) and the permeant concentration in dose solution (C_d), defined as follows (Couto et al., 2014):

$$k_p = \frac{J_{ss}}{C_d} \quad (1)$$

Where permeability coefficients were provided only in diagrammatical form, we estimated their values from the plots.

2.1.2. Generation of molecular descriptors

To describe the human skin permeability of a permeant as a function of the permeant's molecular structure and its ionization state in the dose solution, we obtained computational molecular descriptors as explanatory variables following the geometrical optimization of the molecular structures of the permeants. The detailed procedures are as follows.

All two-dimensional structures of the permeants were obtained from CAS SciFinder. Three-dimensional structures in neutral form were generated using the LigPrep function in Schrödinger Small-Molecule Drug Discovery Suite package (Schrödinger LLC, 2015), and then optimized by conformational search and energy minimization using the OPLS3 force field (Harder et al., 2016) in MacroModel function of the same software package. The molecular descriptors were calculated using three different software packages, *i.e.*, ADMET Predictor (Simulations Plus Inc., 2015), MOPAC 2012 (Stewart Computational Chemistry, 2015), and Dragon 7.0