GDAŃSKI UNIWERSYTET MEDYCZNY
WYDZIAŁ FARMACEUTYCZNY Z ODDZIAŁEM
MEDYCYNY LABORATORYJNEJ

Łukasz Kubik

MODELOWANIE ZALEŻNOŚCI POMIĘDZY STRUKTURĄ I RETENCJĄ CHROMOTOGRAFICZNĄ ZWIĄZKÓW DYSOCJUJĄCYCH

Zakład Biofarmacji i Farmakokinetyki
Katedry Biofarmacji i Farmakodynamiki
Gdańskiego Uniwersytetu Medycznego

Promotor pracy:
dr hab. Paweł Wiczling

Gdańsk 2016

STRESZCZENIE

Podczas procesu poszukiwania nowych leków jednym z pierwszych kroków jest synteza tysięcy związków chemicznych oraz żmudne odrzucanie tych, które, ze względu na swoje nieodpowiednie właściwości fizykochemiczne, nie mają możliwości żeby stać się lekiem. Te badania przesiewowe związane są ze znaczącymi nakładami finansowymi oraz czasowymi. Znajomość podstawowych parametrów fizykochemicznych, takich jak stała dysocjacji (pK_a) oraz lipofilowość (log$P$), pozwala z dużym prawdopodobieństwem przewidywać zachowanie się cząsteczki w ustroju, dzięki czemu możliwe jest odrzucenie na wczesnym etapie badań przesiewowych związków o niepożądanych cechach. Do tego celu konieczne są metody pozwalające na szybkie i dokładne oznaczanie tych parametrów dla dużej liczby związków chemicznych.

Jedną z technik, która może być wykorzystana w tym celu, jest gradientowa wysokosprawna chromatografia cieczowa w odwróconym układzie faz (RP HPLC) połączona z ilościowymi zależnościami struktura-retencja (QSRR). Do tej pory badania obejmujące QSRR rzadko uwzględniały proces dysocjacji analitów. Ponieważ większość stosowanych obecnie leków należy do grupy związków dysocjujących, konieczne było opracowanie modeli matematycznych, które pozwoliłyby na opisanie zależności pomiędzy strukturą chemiczną analitów, ich retencją chromatograficzną oraz wspomnianymi parametrami fizykochemicznymi. W tym celu, opracowano modele przestrzenne dla ponad stu różnorodnych związków chemicznych i ich form zjonizowanych, przy wykorzystaniu narzędzi modelowania molekularnego. Następnie, dla wszystkich analitów oznaczono czasy retencji, przy użyciu opracowanej w Katedrze Biofarmacji i Farmakodynamiki Gdańskiego Uniwersytetu Medycznego metody gradientowej RP HPLC sprzężonej z analizatorem czasu przelotu (TOF MS). Metoda ta pozwoliła na szybkie i jednoczesne oznaczenie czasów retencji poszczególnych związków obecnych w złożonej mieszaninie oraz estymację podstawowych parametrów chromatograficznych (logk_w – chromatograficzna miara hydrofobowości, pK_a oraz S – współczynnik nachylenia z równania Snydera-Soczewińskiego). Odnotowano wysoką korelację pomiędzy logk_w i pK_a wyznaczonymi chromatograficznie, a wartościami literaturowymi oraz obliczonymi komputerowo przy użyciu oprogramowania ACD/Labs.

Następnie, trzy metody statystyczne (technika Lasso, regresja krokowa i regresja metodą cząstkowych najmniejszych kwadratów) zostały użyte do skorelowania deskryptorów strukturalnych oraz otrzymanych parametrów opisujących retencję analitów w układzie

chromatograficznym. Tak opracowane modele matematyczne pozwoliły przewidywać parametry fizykochemiczne analitów na podstawie struktury chemicznej ich cząsteczki, a w konsekwencji ich retencję w układzie chromatograficznym. Opracowane równania cechowały się zadowalającymi zdolnościami predykcyjnymi oszacowanymi za pomocą sprawdzianu krzyżowego oraz walidacji zewnętrznej. Stworzone modele porównano z równaniami obliczonymi na podstawie prostych deskryptorów pochodzących z oprogramowania ACD/Labs. Wyniki uzyskane dla obu podejść wykazywały nieznaczne różnice, co może świadczyć o braku potrzeby tworzenia (w pewnych sytuacjach) skomplikowanych modeli opartych na trudnych do interpretacji deskryptorach trójwymiarowych, otrzymanych metodami modelowania molekularnego.

ABSTRACT

The search process for new drugs involves synthesis of numerous chemical compounds that are subsequently screened and evaluated in terms of physicochemical properties. These tests are usually related to significant financial and time costs. Knowledge of physicochemical properties, such as dissociation constant (pK_a) and lipophilicity (log$P$), allows to predict the likely behavior of a molecule in the human body. It also allows rejection of compounds with undesirable properties at early stages of drug development process. For this reason, it is important to develop methods for the simultaneous determination of log$P$ and pK_a in a fast and accurate manner.

The technique which can be readily used for this purpose is a gradient Reversed-Phase High-Performance Liquid Chromatography (RP HPLC) associated with the Quantitative Structure-Retention Relationships (QSRR). So far, the QSRR models have rarely considered the process of analytes dissociation. Since the majority of currently used drugs dissociate in the solution, it is crucial to develop mathematical models that take this process into account. It would increase the reliability and accuracy of QSRR relationships in predicting chromatographic retention and physicochemical properties based on chemical structure of the analytes. For this purpose, the three-dimensional models involving neutral molecules and ionic forms were developed applicating molecular modeling methods, for more than a hundred of diverse chemical compounds. Subsequently, the retention times of all forms of the compounds at different pH were determined, using developed in the Department of Biopharmacy and Pharmacokinetics of Medical University of Gdańsk method of the gradient RP HPLC coupled with the Time Of Flight Mass Spectrometer (TOF MS). This technique allowed fast and simultaneous determination of the chromatographic parameters (logk_w – chromatographic measure of hydrophobicity, pK_a and S – slope factor from Snyder-Soczewiński equation) based on the retention times of compounds present in complex mixture. For estimated logk_w and pK_a parameters there was high correlation with the literature and computational data.

Lasso, stepwise regression and partial least squares regression were used to correlate structural descriptors and obtain chromatographic parameters. The developed mathematical models allowed prediction of the physicochemical parameters, based on the compound structure which, in consequence, allowed to determine their retention in the chromatographic system. The developed equations were characterized by a satisfactory prediction performance evaluated by cross and external validation. The models developed were compared with models

built based on simple descriptors from the ACD/Labs software. The results from both approaches showed minor differences, which may indicate that there is no need to develop (in certain situations) complex models based on difficult to interpret, three-dimensional descriptors obtained with molecular modeling methods.