molecular pharmaceutics

Article

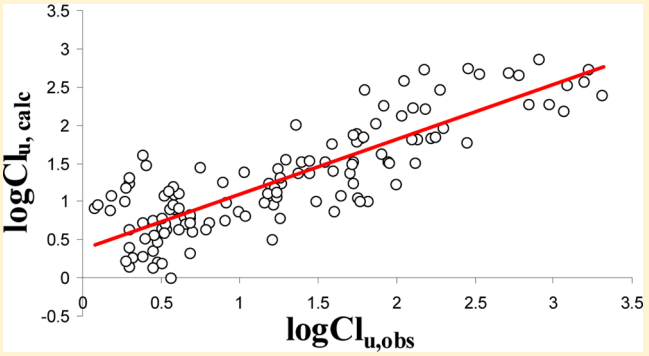
pubs.acs.org/molecularpharmaceutics

Quantitative Structure − Clearance Relationships of Acidic Drugs

Zvetanka Zhivkova* and Irini Doytchinova

Faculty of Pharmacy, Medical University of Sofia, Sofia 1000, Bulgaria

ABSTRACT: Drug clearance is the most important of all pharmacokinetic parameters. It is affected significantly by the binding of drugs to serum proteins. Only the free (unbound) fraction of the drug is able to be cleared. The unbound clearance CL_u is the clearance with reference to unbound drug in plasma C_u. CL_u is independent of the plasma protein binding and depends only on drug chemical structure and properties. In the present study, the relationship between the unbound clearance CL_u and the chemical structures of acidic drugs was modeled by a quantitative structure−clearance relationship (QSCLR) approach. The derived models were used to reveal the main structural features important for CL_u. It was found that the lipophilicity of acidic drugs and the presence of substituents in the aromatic rings, cyano group, and/or nonpolar hydrogen atoms increase the rate of unbound clearance. The presence of sulfonyl groups, quaternary carbon atoms and/or eight-member ring system decreases the unbound clearance of drugs. Additionally, QSCLR models for renal, hepatic, and biliary clearances were derived.



KEYWORDS: *clearance, computational ADME, in silico modeling, QSPR, unbound clearance prediction, acidic drugs, elimination*

INTRODUCTION

The success of a new drug candidate is determined not only by its high efficacy and safety but also by its proper pharmacokinetic (PK) behavior. The deep understanding of the importance of pharmacokinetics led to extensive research and development of numerous approaches for prediction of several parameters characterizing absorption, distribution, metabolism, and excretion (ADME) of druglike compounds. The early prediction of ADME properties is of great importance to minimize time, cost, and labor expenses and to increase the success rate of drug discovery and development. As a result, in 10 years the drug failure due to PK problems has fallen from 39% in 1991, to 10% in 2000.[1]

One of the most reliable and frequently used approaches for prediction of ADME is computational *(in silico)* modeling. It allows derivation of quantitative structure−pharmacokinetic relationships (QSPkR) based on molecular descriptors. An undeniable advantage of this approach is its ability to make predictions at the early stages of drug development, even for virtual compounds. The use of *in silico* ADME modeling for early decisions in drug development is considered in many recently published articles and books[2−9]

Drug clearance (CL) is the proportionality factor that relates the rate of drug elimination to the drug concentration in plasma. For first-order elimination, CL has a constant value and is measured by the plasma volume completely cleared of the drug per unit time (L/h, mL/min). In nonlinear elimination, CL depends on plasma concentration.[10]

Drug clearance considers the entire body as a single drug-containing compartment with a known volume (volume of distribution, V_d) where the drug is eliminated from with a rate constant of elimination *K*. Drug clearance is a very important pharmacokinetic parameter because it is one of the three determinants of the dosing rate:

$$\text{dosing rate} = \frac{\text{clearance} \times \text{therapeutic plasma concentration}}{\text{bioavailability}}$$

where dosing rate is the dose per dosing interval, clearance is expressed for the given dosing interval, therapeutic concentration is the average effective target plasma concentration, and bioavailability is the fraction of extravascularly administered drug reaching the systemic circulation unchanged.[11]

The total body clearance describes the drug elimination from the body without identifying the mechanism of the process.[12] Most drugs are eliminated primarily via liver and kidney and their CL is additive quantity: sum of renal clearance (CL_R) and hepatic clearance (CL_H). Hepatic elimination involves numerous processes: hepatic uptake, biotransformation catalyzed by phase I and phase II enzymes, excretion of unchanged drug and metabolites into bile. Renal clearance also involves several mechanisms: glomerular filtration, active tubular secretion and tubular reabsorption. In addition, the kidneys also have the ability to metabolize drugs. Hepatic uptake, biliary excretion, renal tubular secretion and sometimes tubular reabsorption are mediated by various uptake and efflux transporters. The important role of transporters in drug clearance has been recently recognized and is subject of numerous publications.[13,14]

Received: April 24, 2013
Revised: June 18, 2013
Accepted: July 30, 2013
Published: July 30, 2013

A good number of studies exist in the literature predicting the total body clearance CL from a variety of molecular descriptors. Models are derived by different statistical and machine learning methods as partial least-squares (PLS),[15−17] artificial neural networks (ANN),[18−20] *k* nearest neighbors (*k*NN),[16,20,21] support vector regression (SVR),[20] random forest (RF),[17] classification and regression trees (CART),[17] naïve Bayesian statistics,[22] molecular fingerprints,[22] linear scoring function,[23] etc. Human *in vitro* intrinsic clearance CL_{int} representing the hepatic enzyme activity toward a compound also was modeled.[24,25]

Clearance is affected significantly by the binding of drugs to serum proteins. Only the free (unbound) fraction of the drug is able to be cleared. The unbound clearance CL_u is the clearance with reference to unbound drug in plasma CL_u, given by

$$CL_u = \frac{CL}{f_u}$$

where f_u is the fraction of unbound drug in plasma.[10] For example, if a drug does not bind to plasma proteins ($f_u = 1$), its CL will be equal to CL_u. If 90% of a drug is bound to plasma proteins ($f_u = 0.1$), its unbound clearance CL_u will be 10 times higher than the total body clearance. Thus, CL_u is independent of the plasma protein binding of the drug and depends only on its chemical structure and properties. It could be considered as the clearance of a drug if it would not bind to plasma proteins.

In our previous studies, the QSPkR approach was successfully applied to derive robust models for prediction of steady-state volume of distribution (VD_{ss})[26] and plasma protein binding (PPB).[27] In the present study, the relationship between the unbound clearance CL_u and the chemical structures of acidic drugs was modeled by a Quantitative Structure − CLearance Relationship (QSCLR) approach. The derived models were used to distinguish the main structural features of acidic drugs important for their unbound clearance. For practical uses, we summarized these features into a short checklist of criteria for clearance prediction.

EXPERIMENTAL SECTION

Compound Selection. The acidic drugs used in the present study was extracted from Obach, Lombardo and Waters' database.[28] A drug was considered as acid if its acid fraction (f_A) exceeded 10% when its base fraction (f_B) is zero; or/and its f_A is significantly higher than f_B and reaches 100%. The f_A and f_B values were calculated according to the equations:

$$\text{for acid: } f_A = \frac{1}{1 + 10^{(pK_a - 7.4)}};$$

$$\text{for base: } f_B = 11 + 10^{(7.4 - pK_a)}$$

The pK_a values were calculated by ACD/LogD version 9.08 software (Advanced Chemistry Development Inc., Ontario, Canada). The mol files of the drugs were compiled from DataBank[29] and ChemicalBook.[30] The number of drugs identified as acidic was 132. Artesunate was excluded from the set because of its extraordinary high value for CL (1,070 mL/min/kg). Thus, the final set used in the study consisted of 131 structures.

The total body clearances CL (mL/min/kg) of the drugs and their unbound fractions in plasma f_u were retrieved from Obach's database, and the corresponding unbound clearances CL_u (mL/min/kg) were calculated. For QSCLR purposes, CL_u values were used as log CL_u. Thus, a higher log CL_u value corresponds to a higher unbound clearance. The data set was ranked according to log CL_u in a descending order and was divided into five subsets of 22 and one subset of 21 drugs each by allocating one of every six drugs into a different subset. Every subset was used once as a test set, whereas the remaining five subsets were united into a training set. The training sets were used to derive and cross-validate the QSCLR models, while the test sets were used for external validation.

In order to derive QSCLR models for renal, hepatic, and biliary clearances, the acidic drugs were divided into three groups: drugs excreted mainly unchanged in the urine, drugs eliminated mainly as metabolites, and drugs excreted mainly in the bile. Information about the main routes of elimination of the studied drugs was obtained from various literature sources. A drug was considered as being eliminated via a particular route if more than 60% of the administered dose is cleared by this route.

Variable Selection. A three-step variable selection procedure was applied to derive the most relevant molecular descriptors for clearance prediction. Initially, the chemical structure of drugs was described by a set of 178 descriptors, computed using ACD/LogD and MDL QSAR version 2.2 (MDL Information Systems, Inc., San Leandro, CA). Descriptors with nonzero values for less than three drugs and descriptors correlating to log CL_u with $r < 0.1$ were eliminated. Thus, the number of descriptors was reduced to 112. They included electrotopological (E-state) indices,[31] the number of atoms of a given atom type, molecular connectivity indices,[32] and lipophilicity parameters: log *P* or log $D_{7.4}$.

For every training set, variables were selected by genetic algorithm (GA)[33] followed by a stepwise regression. The GA, as implemented in MDL QSAR, was used with default values for the size of initial population (32), choice of parents (tournament selection), types of crossover (uniform) and mutation (one-point), and fitness function (Friedman's lack-of-fit scoring function with two parameters[34]). The stepwise regression was used in a forward mode with default value for F-to-enter (4.00) and F-to-remove (3.99). Models were assessed by explained variance (r^2), cross-validated coefficient q^2, external-validated coefficient r^2_{pred}, and mean fold error (MFEP) as is explained in the Model Validation section below. Two best-performing models for each training set were selected, and the descriptors emerged in more than one model were compiled into a new descriptor set.

This descriptor set was used further to derive the final QSCLR models.

Model Assessment and Validation. The QSCLR models derived in the present study were assessed by explained variance (r^2), according to the following equation:

$$r^2 = 1 - \frac{\sum_{i=1}^{n} (\log CL_{u,obs,i} - \log CL_{u,calc,i})}{\sum_{i=1}^{n} (\log CL_{u,obs,i} - \log CL_{u,obs,mean})}$$

where $CL_{u,obs,i}$ is the observed CL_u of the *i*th drug, $CL_{u,calc,i}$ is the calculated by the model CL_u of the *i*th drug, and $CL_{u,obs,mean}$ is the mean value of observed CL_u.

Drugs were considered as outliers if the residuals did not obey the normal distribution law. They were removed from the training sets, and the models were rebuilt.

Two validation procedures were applied to access the predictive ability of the QSCLR models: (1) leave-one-out cross-validation (LOO-CV) in the training sets and (2) external

validation by test sets. The models were assessed by the coefficients $q^2_{\text{LOO-CV}}$ and r^2_{pred}, respectively, according to the following equations:

$$q^2_{\text{LOO-CV}} = 1 - \frac{\sum_{i=1}^{n} (\log \text{CL}_{\text{u,obs},i} - \log \text{CL}_{\text{u,pred},i})}{\sum_{i=1}^{n} (\log \text{CL}_{\text{u,obs},i} - \log \text{CL}_{\text{u,obs,mean}})},$$

$$r^2_{\text{pred}} = 1 - \frac{\sum_{i=1}^{n} (\log \text{CL}_{\text{u,obs},i} - \log \text{CL}_{\text{u,pred},i,\text{test}})}{\sum_{i=1}^{n} (\log \text{CL}_{\text{u,obs},i} - \log \text{CL}_{\text{u,obs,mean}})}$$

where $\log \text{CL}_{\text{u,pred},i}$ is predicted by the model $\log \text{CL}_{\text{u}}$ of the ith drug. In the LOO-CV procedure, one drug was excluded from the training set, the model was derived on the basis of the remaining $n - 1$ drugs and used to predict the $\log \text{CL}_{\text{u,pred},i}$ of the excluded ith drug. In the external validation, $\log \text{CL}_{\text{u,pred},i,\text{test}}$ is the $\log \text{CL}_{\text{u}}$ value of a drug from the test set predicted by the model derived on the basis of the training set.

Fold error of prediction for the test sets was calculated according to the equation:

$$\text{FEP} = 10^{|\log \text{CL}_{\text{u,obs},i} - \log \text{CL}_{\text{u,pred},i}|}$$

and the average value of FEPs for the test drugs is given as mean fold error MFEP.

The accuracy of prediction is given as a percent of the total number of drugs predicted with less than 2-fold error.

RESULTS

Selected Compounds. The data set of 131 acidic drugs used in the study was selected according to their f_{A} values as described in the Experimental Section. The values for CL_{u} range from 0.13 mL/min/kg for phenobarbital to 2,100 mL/min/kg for telmisartan. The respective values for $\log \text{CL}_{\text{u}}$ vary between −0.89 and 3.32 with an average value of 1.27 and median value of 1.21. The histogram of $\log \text{CL}_{\text{u}}$ distribution is given in Figure 1.

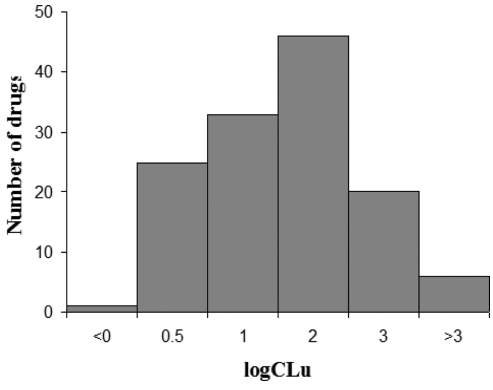


Figure 1. Histogram of the distribution of $\log \text{CL}_{\text{u}}$ values.

The data set of acidic drugs was divided into three subsets according to their dominating mechanism of elimination: unchanged by renal excretion, by metabolism, and by biliary secretion. Specific QSCLR models were derived for each subset.

Selected Variables. The most relevant molecular descriptors for clearance prediction were selected by a three-step variable selection procedure described in the Experimental Section. The final set comprises 19 of the most frequently emerged descriptors divided into two groups:

Descriptors with positive effect on $\log \text{CL}_{\text{u}}$:

- log P or log $D_{7.4}$. These lipophilicity parameters intercorrelate with $r^2 = 0.748$, and they appear in all models.
- *SaasC_acnt*. It counts the number of aromatic substituted carbon atoms in the molecule and ranges from 1 to 18 in 109 of the studied compounds.
- *SHother*. This descriptor gives the sum of hydrogen atom type E-state indices for all nonpolar H-atoms in the molecule. It ranges from 0 (foscarnet) to 55 (micafungin). Correlates highly ($r^2 = 0.879$) with the sum of C-atoms connected with nonpolar H-atoms (atom types aaCH, sCH_3, ssCH_2, sssCH, sdCH, and dCH_2).
- H_{max}. Shows the E-state of the most polar H-atom in the molecule. Ranges from 1.46 (sulfinpyrasone) to 3.16 (suramin).
- *StsC*, *StN*, or *StsC_acnt*. All of them account for the presence of the cyano (−CN) group in the molecule.
- *SsCl*. Encodes the sum of E-state indices for Cl atoms. It is present in 13 molecules ranging from 5.64 to 12.80.
- *SsssNHp_acnt*. Counts the number of quaternary ammonium atoms (>N^+<). It presents in four molecules: nitrofurantoin, entacapone, tolcapone, ceftazidime.
- *SdssC_acnt*. Corresponds to the number of C-atoms with one double and two simple bonds. It is present in 118 molecules ranging from 1 to 9.

Descriptors with negative effect on $\log \text{CL}_{\text{u}}$:

- *xc4*. Encodes the presence, number and type of atoms, connected to four other atoms or groups (star-type structure). It appears in 66 molecules ranging from 0.068 to 0.590 with an extremely high value of 1.732 for suramin. It encompasses ssssC atoms (quaternary carbon atoms), S in sulfonyl groups, and P in phosphonyl groups.
- *xch8* or *xvch8*. They encode the presence of an eight-member ring system and are present in 24 molecules (cefalosporins and ketorolac).
- *SddssS* or *SddssS_acnt*. Both of them account for the presence of sulfonyl groups. This fragment is available in 28 molecules. Slightly correlative with xc4 ($r^2 = 0.595$ and 0.569, respectively).
- *SdO* or *SdO_acnt*. They account for the presence and environment of O-atoms with a double bond.
- *SssssC_acnt*. Counts the number of C-atoms bound to four other atoms or groups (quaternary carbons). It presents in 36 molecules. This information is also encoded in xc4 which is a more general descriptor.

These descriptors were used further in the development of QSCLR models.

Quantitative Structure−Unbound Clearance Relationship Models. The model derived by GA and stepwise regression using the selected molecular descriptors and based on the whole set of 131 acidic drugs is given below as Model 1:

$$\begin{aligned}\log \text{CL}_{\text{u}} = {} & 0.135(\pm 0.018) \log D_{7.4} + 0.125(\pm 0.024) \\ & \text{SaasC\_acnt} + 0.028(\pm 0.008)\text{SHother} + 0.065 \\ & (\pm 0.019)\text{StN} + 0.446(\pm 0.012)H_{\text{max}} - 0.632 \\ & (\pm 0.082)\text{SddssS\_acnt} - 0.399(\pm 0.085) \\ & \text{SsssC\_acnt} - 41.430(\pm 7.285)\text{xvch8} + 0.032 \end{aligned} \quad \text{(Model 1)}$$

$n = 126$, $r^2 = 0.720$, $q^2_{\text{LOO-CV}} = 0.682$, MFEP = 2.84 ± 2.39, accuracy = 52%

dx.doi.org/10.1021/mp400251k | Mol. Pharmaceutics 2013, 10, 3758−3768

outliers: amidotrizoat, phenobarbital, sulbactam, valsartan, warfarin

Five outliers were detected and removed from the data set. The explained variance is 72% and the model predicts log CL_u values with less than 2-fold error for 52% of the drugs. The plot of calculated by Model 1 log CL_u vs observed log CL_u is given in Figure 2.

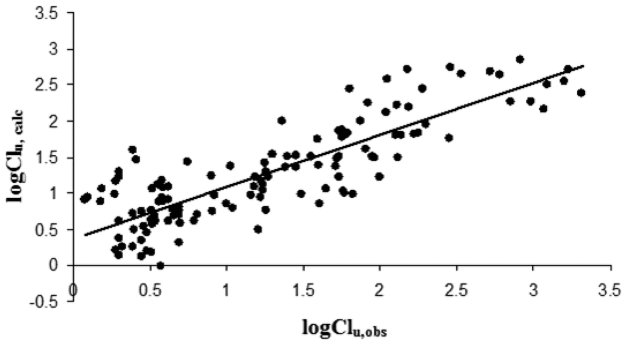


Figure 2. log CL_u values calculated by Model 1 versus observed log CL_u for 131 acidic drugs.

To check the robustness of the derived model, the whole set of 131 drugs was divided into five subsets of 22 and one subset of 21 drugs, as described in Experimental Section. Each subset was used once as a test set whereas the remaining five subsets were used as a training set. QSCLR models were derived on six training sets and externally validated by the corresponding test sets. The best models are presented in Table 1 The model statistics before removing the outliers is given in Appendix 1.

The explained variance of the models varies between 69% and 74%. The $q^2_{LOO\text{-}CV}$ values are from 0.631 to 0.691 (mean 0.670), and external validation r^2_{pred} – from 0.605 to 0.824 (mean 0.737). The accuracy ranges from 32% to 55%. Only two additional descriptors appear in the models. These are StsC_acnt and SsssNHp_acnt. StsC_acnt is similar to StN and accounts for the presence of cyano group, while SsssNHp_acnt corresponds to the number of quaternary ammonium atoms in the molecule.

Quantitative Structure–Renal Clearance Relationship Model. A set of 47 acidic drugs eliminated primarily or entirely by renal excretion was selected from the initial training set. The model derived by GA and stepwise regression on the initial set of molecular descriptors is given below as Model 2:

$$\log CL_u = 0.081(\pm0.019)\log D_{7.4} + 0.769(\pm0.162)\,\mathrm{SaaO\_acnt} - 0.090(\pm0.007)\mathrm{SdsN} + 1.015 \quad \text{(Model 2)}$$

$$n = 44,\ r^2 = 0.622,\ q^2_{LOO\text{-}CV} = 0.511,\ \mathrm{MFEP} = 1.69 \pm 0.57,\ \text{accuracy} = 70\%$$

outliers: betamiprone, milrinone, sulfadiazine

Quantitative Structure–Hepatic Clearance Relationship Model. A set of 50 acidic drugs eliminated primarily or entirely as metabolites was selected from the initial training set. The model derived by GA and stepwise regression on the initial set of molecular descriptors is given below as Model 3:

$$\log CL = 0.239(\pm0.034)\mathrm{SaasC\_acnt} + 0.161(\pm0.05)\,\mathrm{SdsCH} - 0.806(\pm0.212)\mathrm{SddssSS\_acnt} - 1.598(\pm0.022)xc4 + 1.256 \quad \text{(Model 3)}$$

$$n = 48,\ r^2 = 0.665,\ q^2_{LOO\text{-}CV} = 0.597,\ \mathrm{MFEP} = 3.21 \pm 2.50,\ \text{accuracy} = 42\%$$

outliers: glimepyride, warfarin

Quantitative Structure–Biliary Clearance Relationship Model. A set of 16 acidic drugs excreted in the bile was selected from the initial training set. The model derived by GA

Table 1. QSCLR Models Derived on Six Training Sets and Externally Validated by the Corresponding Test Sets

training set	model	r^2	$q^2_{LOO\text{-}CV}$	r^2_{pred}	MFEP	accuracy (%)
1	log CL_u = 0.109(±0.018) log $D_{7.4}$ + 0.138(±0.024)SaasC_acnt + 0.031(±0.008)SHother + 0.522(±0.119)StsC_acnt − 0.642(±0.089)SddssS_acnt − 0.356(±0.105)SssssC_acnt − 35.221 (±7.691)xvch8 + 1.014	0.693	0.643	0.763	2.65 ± 1.38	32
	outliers from the training set: amidotrizoate, bromfenac, fluvastatin, sulbactam, valsartan, warfarin; *outliers from the test set*: phenobarbital					
2	log CL_u = 0.132(±0.019) log $D_{7.4}$ + 0.113(±0.027)SaasC_acnt + 0.030(±0.008)SHother + 0.652(±0.272)StsC_acnt + 0.438(±0.013)H_{max} + 0.686(±0.113)SsssNHp_acnt − 0.692(±0.119) SddssS_acnt − 0.423(±0.092)SssssC_acnt − 44.722(±8.024)xvch8 + 0.094	0.741	0.691	0.699	2.88 ± 1.66	38
	outliers from the training set: amidotrizoate, phenobarbital, sulbactam, warfarin; *outliers from the test set*: enalaprilat, valsartan					
3	log CL_u = 0.118(±0.017) log $D_{7.4}$ + 0.124(±0.025)SaasC_acnt + 0.033(±0.008)SHother + 0.061(±0.019)StN − 0.600(±0.086)SddssS_acnt − 0.399(±0.085)SssssC_acnt − 32.671(±6.684) xvch8 + 1.069	0.723	0.681	0.605	3.79 ± 2.59	41
	outliers from the training set: bromfenac, phenobarbital, valsartan, warfarin; *outliers from the tese set*: sulbactam					
4	log CL_u = 0.129(±0.024) log $D_{7.4}$ + 0.127(±0.028)SaasC_acnt + 0.027(±0.010)SHother + 0.552(±0.089) StsC_acnt + 0.427(±0.012)H_{max} − 0.597(±0.096)SddssS_acnt − 0.355(±0.085) SssssC_acnt − 41.433(±7.285)xvch8 + 0.087	0.690	0.631	0.743	2.59 ± 1.95	50
	outliers from the training set: amidotrizoate, phenobarbital, valsartan, warfarin; *outliers from the test set*: verlukast, micafungin					
5	log CL_u = 0.127(±0.019) log $D_{7.4}$ + 0.135(±0.028)SaasC_acnt + 0.019(±0.008)SHother + 0.060(±0.018) StN + 0.503(±0.112)H_{max} − 0.605(±0.085)SddssS_acnt − 0.381(±0.090) SssssC_acnt − 44.942(±7.619)xvch8 + 0.015	0.726	0.685	0.729	2.84 ± 2.51	55
	outliers from the training set: amidotrizoate, phenobarbital, valsartan, warfarin; *outliers from the test set*: iothalamic acid, isoxicam					
6	log CL_u = 0.136(±0.019) log $D_{7.4}$ + 0.128(±0.028)SaasC_acnt + 0.028(±0.008)SHother + 0.608(±0.175) StsC_acnt + 0.536(±0.012)H_{max} − 0.680(±0.084)SddssS_acnt − 0.392(±0.092) SssssC_acnt − 44.123(±7.816)xvch8 + 0.170	0.738	0.686	0.824	2.27 ± 1.67	55
	outliers from the training set: enalaprilat, iothalamic acid, phenobarbital, sulbactam, valsartan; *outliers from the test set*: amidotrizoate, valproic acid, warfarin					

dx.doi.org/10.1021/mp400251k | *Mol. Pharmaceutics* 2013, 10, 3758–3768

and stepwise regression on the initial set of molecular descriptors is given below as Model 4:

$$\log CL_u = 0.200(\pm 0.033)\text{SaasC\_acnt} - 0.191(\pm 0.034)\text{SaasC\_acnt} - 0.389(\pm 101)\text{SaaNH} + 1.813 \quad \text{(Model 4)}$$

$n = 16$, $r^2 = 0.892$, $q^2_{\text{LOO-CV}} = 0.817$,

MFEP = 1.81 ± 0.69, accuracy = 63%

no outliers

Checklist for Clearance Prediction. The descriptors involved in Model 1 reveal the main structural features responsible for the clearance of free drugs in plasma. The lipophilicity of drugs (expressed as log $D_{7.4}$), the number of substituents in the aromatic rings (descriptor SaasC_acnt), the presence of cyano group (descriptor StN), and/or the number of nonpolar and polar hydrogen atoms (descriptors SHother and H_{max}, respectively) increase the rate of clearance. The presence of sulfonyl groups (descriptor SddssS_acnt), the number of quaternary carbon atoms (descriptor SssssC_acnt), and/or the presence of eight-member ring system (descriptor xvch8) decrease the clearance of drugs. The descriptors from Model 1 were summarized into a short checklist of criteria responsible for clearance prediction (Table 2).

Table 2. Checklist of Criteria for Unbound Clearance CL_u Prediction

no.	presence of	CL_u increases	CL_u decreases
1	log $D_{7.4} \geq 0$	√	
2	number of substituents in aromatic ring ≥4	√	
3	cyano (CN) group ≥1	√	
4	number of C atoms connected with nonpolar H atoms ≥8	√	
5	sulfonyl group ≥1		√
6	quaternary C atoms ≥1		√
7	eight-member ring system ≥1		√

In order to define the cutoff for each descriptor discriminating between drugs with high, moderate, and low rates of clearance, the checklist was applied to the data set of studied acidic drugs (Appendix 2). The drugs were divided into three subsets: drugs with low rates of clearance ($CL_u \leq 3$ mL/min/kg, log $CL_u \leq 0.48$), drugs with moderate rates ($3 < CL_u < 100$ mL/min/kg, $0.48 < \log CL_u < 2$), and drugs with high rates ($CL_u \geq 100$ mL/min/kg, log $CL_u \geq 2$).

The subset of low-rate cleared drugs comprises 26 molecules. Here predominate drugs with negatively contributing structural features. Thirty one percent of the drugs contain sulfonyl group(s); 31%—quaternary carbon atom(s), and 23%—eight-member ring systems. Small number of molecules possesses structural features with a positive impact on CL_u: only 19% contain more than four substitients in an aromatic ring; for 27% the number of C atoms carrying nonpolar H-atoms is higher than eight, and 19% have a positive value of log $D_{7.4}$.

The subset of drugs with moderate rate of clearance consists of 78 structures. These molecules meet approximately equal numbers of positive and negative criteria: 18% of them have positive values of log $D_{7.4}$; 23% contain four or more substituents in aromatic rings; 68%—eight or more C-atoms bound to nonpolar H-atoms; 22% have sulfonyl group; 35%—quaternary carbon atom; and 23%—eight-member ring system. Two of the five cyano group-containing acidic drugs are located in this subset.

The subset of high-clearance drugs consists of 27 molecules with a considerable preponderance of positively contributing structural features. Seventy eight percent of the drugs have positive log $D_{7.4}$ values, 78% contain at least four substituents in aromatic rings, 67% have eight or more C-atoms connected to nonpolar H-atoms. Three of the five cyano group-containing acidic drugs are located in this subset. There are only three drugs with a sulfonyl group, one with quaternary carbon atom, and none with eight-member ring system.

The distribution of the drugs according to the difference between the numbers of positively and negatively contributing features is shown in Figure 3. This difference could be used to distinguish between low- and high-clearance drugs. The difference between the numbers of positive and negative criteria equals 2 or 3 for 78% of the high-clearance drugs. Therefore, we suggest a difference ≥2 to be used as a requirement for high value of the unbound clearance. In opposite, for 77% of the low-clearance drugs, the difference in the numbers of positive and negative criteria is ≤0.

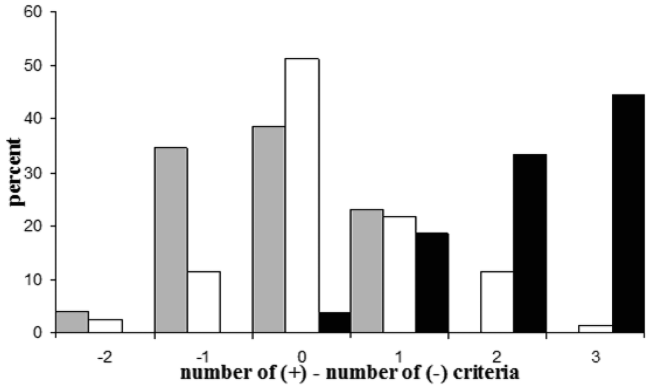


Figure 3. Distribution of the drugs (given as %) according to the difference between the number of positive and negative criteria. Gray: low-clearance drugs. White: moderate-clearance drugs. Black: high-clearance drugs.

DISCUSSION

Quantitative structure—clearance relationships (QSCLRs) for acidic drugs were derived in the present study. The rate of clearance was expressed as logarithm of clearance of the unbound fraction of drug in plasma log CL_u. The chemical structure of 131 acidic drugs was described by 178 molecular descriptors. A three-step variable selection procedure (GA, stepwise regression, selection of most frequently emerged descriptors) was applied to identify the most relevant descriptors. They were used further in the development of a QSCLR model. The model was validated by leave-one-out cross-validation and external validation by six test groups. The close values of the cross-validated and noncross-validated correlation coefficients are indicative for the robustness of the derived model. The descriptors involved in the model were summarized in a short checklist of criteria for clearance prediction. Additionally, QSCLR models for renal, hepatic and biliary clearances were derived.

The total body clearance of a drug is a result of many specific and complex mechanisms. Modeling of drug clearance is further

complicated by the ability of drugs to bind to plasma proteins in different extent. Plasma protein binding (PPB) is a limiting factor for the clearance as only the unbound fraction of drug is filtered through glomerules and is able to pass into hepatocytes. This problem could be eliminated by using the unbound plasma clearance ($CL_u = CL/f_u$) instead of total plasma clearance CL. Thus, CL_u depends mainly on the drug structural features and physicochemical properties.

In the present study, we found that the lipophilicity of drugs at pH 7.4 is one of the main properties responsible for their clearance. The value of log $D_{7.4}$ for the studied compounds ranges between −11 and 5.19. About 55% of the drugs with positive log $D_{7.4}$ values are cleared with rate higher than 100 mL/min/kg. The positive influence of lipophilicity on drug clearance was detected previously.[17,19]

The presence of more substituents in the aromatic rings increases the drug clearance. Molecules with cyano group also are cleared rapidly. We found recently that this group increases plasma protein binding.[27] In our data set, there are five drugs with cyano groups. These are cefmetazole, cilomilast, entacapone, levosimendan, and milrinone. The last four of them are highly bound to plasma proteins with $f_u < 0.05$. Their total body clearances are moderate varying between 0.5 and 12 mL/min/kg, but because of the low f_u values, the unbound clearances CL_u are very high ($CL_u > 80$ mL/min/kg). This is a good example of how high plasma protein binding can obscure for lower drug clearance.

The presence of aromatic and aliphatic C-atoms connected with nonpolar H-atoms is favorable for the clearance. These atoms contribute to lipophilicity although the descriptor SHother does not correlate directly to log $D_{7.4}$. H_{max} reflects the ionizability of a molecule; higher H_{max} corresponds to stronger ionized acid. We found recently that H_{max} contributes negatively to the volume of distribution (VD_{ss}), i.e. the ionized acids have low VD_{ss}'s.[26] Here we find that stronger acids are cleared faster. This is in a good agreement with the general consideration that the ionized acids have high rates of renal clearance as they do not reabsorb and most of them are actively secreted in tubules.

We found recently that the presence of sulfur atoms from the type (−S−), as well as SO-groups affect negatively the VD_{ss}'s.[26] Here, it is evident that the presence of sulfonyl groups reduces the clearance of acidic drugs. Unfavorable for the clearance are also the presence of quaternary carbon atoms and 8-member ring system in the molecule. Cephalosporin antibiotics contain such ring system. Cephalosporins are weak acids with a low pK_a, poor lipid solubility, small VD_{ss}'s, short half-lives and variable protein binding. Their total body clearances CL are close and fall in the range between 1 and 4 mL/min/kg.[28] However, because of the variable protein binding, they have quite diverse unbound clearances CL_u.

As the descriptors used in the present study have clear and unambiguous physical sense, they were used to define a checklist of several structural criteria for clearance prediction (Table 2).

Amidotrizoate, phenobarbital, sulbactam, valsartan, and warfarin are outliers from Model 1. Data for pK_a, log $D_{7.4}$, f_u, $CL_{u,obs}$ and $C_{u,calc.}$ of the outliers are presented in Table 3.

The calculated value of CL_u for amidotrizoate (diatrizoic acid) exceeds considerably the observed value. Amidotrizoate is a strong acid, not bound to plasma proteins. It is eliminated exclusively via renal excretion and is used as a marker for glomerular filtration.[35] Amidotrizoate has one positive and none negative structural features (Appendix 2). However, it contains three iodine atoms attached to an aromatic ring. Probably, the iodine atoms contribute negatively to the clearance, and their absence in the model results in overestimation of amidotrizoate's CL_u. This suggestion is supported by the fact that similar large difference between predicted and experimental value of CL_u is observed for iothalamic acid, a structural isomer of amidotrizoate (38.90 vs 2.45 mL/min/kg).

Table 3. Data for pK_a, log $D_{7.4}$, f_u, $CL_{u,obs}$ and $C_{u,calc}$ of the Outliers from Model 1

outlier	pK_a	log $D_{7.4}$	f_u	$CL_{u,obs}$ mL/min/kg	$CL_{u,pred}$ mL/min/kg
amidotrizoate	0.92	1.53	1.0	1.70	50
phenobarbital	7.93	1.47	0.49	0.137	12.5
sulbactam	2.52	−5.11	0.62	8.233	0.47
valsartan	3.69	0.01	0.04	12.25	254
warfarin	4.5	0.61	0.015	3.667	141.25

Phenobarbital is the drug with the lowest value of CL_u in the data set. The calculated CL_u is highly overestimated. Phenobarbital meets one positive and one negative structural criterion. The drug is eliminated primarily by metabolism (*N*-glycosylation, hydroxylation to *p*-hydroxyphenobarbital, glucuronidation, and sulfate conjugation). The fraction excreted unchanged in urine varies considerably (9−33%) and depends crucially on urine flow and pH because of the high rate of tubular reabsorption.[36]

The calculated value of CL_u for sulbactam is considerably lower than the observed one. Its molecule contains two negative criteria and no one positive. Although sulbactam is structurally similar to penicillins, it has much lower molecular mass and does not meet the positive criteria like the number of substituents in aromatic rings and the number of C-atoms connected to nonpolar H-atoms. Sulbactam is eliminated primarily by renal excretion. Its renal clearance accounts for 76% of the total clearance and is dominated by tubular secretion.[37]

The CL_u's of valsartan and warfarin are also overestimated. Valsartan meets three positive and no one negative criteria. The drug is eliminated primarily unchanged in the bile, and various organic anion transporters are involved in its clearance.[38] Warfarin molecule contains two positive criteria and no one negative. The drug exists as a pair of enantiomers. They both are eliminated hepatically and less than 1% of the dose is excreted unchanged in urine. Metabolism is stereoselective with S-warfarin being metabolized by CYP2C9, while R-warfarin undergoes biotransformation mediated by a variety of enzymes including CYP1A1, CYP1A2, CYP2C19, CYP3A4, catehol-*O*-methyltransferase and ketoreductase.[39] Wide interindividual variations in warfarin pharmacokinetics are reported most likely due to the genetic polymorphism of CYP2C9. The complex elimination via various pathways with different structural requrements might be the reason for the large difference between the observed and predicted values of CL_u for most of the outliers.

The QSCLR model for renal clearance prediction Model 2 was constructed for a data set of 47 drugs cleared primarily or entirely by renal excretion. The model contains the descriptor log $D_{7.4}$ and two additional ones: SaaO_acnt and SdsN. The descriptor SaaO_acnt counts the atoms from type aaO (O in aromatic ring). It presents in three molecules and contributes positively to CL_u. The descriptor SdsN encodes the presence and arrangement of =N− atoms. This fragment exists in nine molecules and affects negatively CL_u.

The drugs in this data set have relatively low values of CL (0.0057 − 9.2 mL/min/kg, mean 2.77 ± 1.85). For most of them, CL exceeds the glomerular filtration rate (GFR) which is about 1.7 mL/min/kg. This implies a considerable contribution

dx.doi.org/10.1021/mp400251k | Mol. Pharmaceutics 2013, 10, 3758−3768

of tubular secretion to renal excretion. Tubular reabsorption seems to be negligible since most of the drugs are strong acids completely ionized at physiological pH. The drugs are moderately bound to plasma proteins (f_u = 0.003 − 1, mean 0.54 ± 0.30) which results in relatively low values for CL_u (1.25−177 mL/min/kg). According to Model 2, log $D_{7.4}$ affects positively log CL_u. All drugs in the data set have negative values of log $D_{7.4}$ (except of milrinone). The negative value of log $D_{7.4}$ is considered as a prerequisite for high renal excretion since it prevents tubular reabsorption.[40] Recently, it was found that drugs with log $D_{7.4}$ between −2 and 0 have maximal values of renal clearance, because these values span the region where the drugs have sufficient lipophility to interact with transport proteins but are not substantially reabsorbed.[41]

There are three outliers from Model 2: betamiprone, milrinone and sulfadiazine. Betamiprone has CL_u = 54.12 mL/min/kg, while the calculated by Model 2 CL_u = 6.58 mL/min/kg. According to the original reference[42] cited in Obach's database,[28] betamiprone has CL_u = 33.47 mL/min/kg—a value closer to the predicted one. The calculated value for milrinone CL_u (10.76 mL/min/kg) is considerably lower than the observed one (177 mL/min/kg). Milrinone is a weak acid with pK_a = 7.74, positive value of log $D_{7.4}$ and extensive PPB. These features make it quite different from the other structures in the data set and might explain the incorrect prediction. Sulfadiazine has observed CL_u = 1.25 mL/min/kg and calculated CL_u = 9.16 mL/min/kg. The renal excretion is the major elimination pathway for sulfadiazine but 20−40% of the dose is eliminated as acetylated metabolites. Oxidative metabolism to hydroxylamine and 4-hydrohy sulfadiazine has also been reported recently.[43]

Table A1. Model Statistics before Removing the Outliers

model	n	r^2	$q^2_{LOO\text{-}CV}$
Model 1	131	0.618	0.565
training set 1	109	0.617	0.562
training set 2	110	0.616	0.574
training set 3	109	0.626	0.577
training set 4	109	0.592	0.513
training set 5	109	0.621	0.560
training set 6	109	0.638	0.581
Model 2	47	0.552	0.358
Model 3	50	0.539	0.457

The model for hepatic clearance Model 3 was derived for a data set of 50 molecules. The model contains the already well-known descriptors SaasC_acnt (number of substitutents in aromatic rings) and SddssS_acnt (number of sulfonyl groups) and two additional ones: SdsCH and *xc4*. The descriptor SdsCH encodes the presence and electronic state of ═CH− groups, and *xc4* indicates the presence of a star type structure which may be a quaternary carbon, sulfonyl or phosphonate group. SaasC_acnt and SdsCH have positive contribution to hepatic clearance, while SddssS_acnt and *xc4* decrease it.

The drugs in this data set have relatively high values of CL (0.03−26 mL/min/kg, mean 4.00 ± 5.33) and extensive PPB (f_u = 0.11 ± 0.19) which results in high values of CL_u (1.21−2.025 mL/min/kg). There are two outliers from Model 3: glimepyride and warfarin. Warfarin is also an outlier from Model 1. Glimepyride has CL_u = 100 mL/min/kg. Model 3 predicts value for CL_u 4.00 mL/min/kg. Glimepyride is completely metabolized by CYP2C9 in the liver and there data for considerable differences in disposition due to genetic polymorphism of CYP2C9.[44]

The model for biliary clearance, Model 4, was developed for a data set of 16 drugs explicitly excreted in the bile. According to this model, the biliary excretion increases with the number of substituents in aromatic rings and decreases in the presence of aromatic NH and ═C< atoms in the molecule. This model should be considered by caution because of the small number of compounds in the training set.

CONCLUSIONS

Statistically significant QSCLRs were constructed for the unbound clearance of 131 acidic drugs. The predictive ability of the models was validated by internal and external validations. The unbound clearance of acidic drugs depends on the presence of several well identified structural fragments. The lipophilicity, the presence of substituents in the aromatic rings, cyano groups, and/or nonpolar hydrogen atoms affects positively the drug clearance. The presence of sulfonyl groups, quaternary carbons and/or 8-member ring system disfavors the clearance of drugs.

APPENDIX 1

APPENDIX 2

Table A2. Seven-Item Checklist of Criteria for Unbound Clearance Applied to the Dataset of Acidic Drugs

		positive criteria				negative criteria					
name	CL_u, mL/min/kg	log $D_{7.4}$ > 0	CH_n ≥8	substituted arC > 4	CN	quaternary C atom	8-member ring	sulphonyl group	sum (+) criteria	sum (−) criteria	sum (+) − sum (−) criteria
Drugs with Low Rates of Clearance (CL_u ≤ 3 mL/min/kg, log CL_u ≤ 0.48)											
phenobarbital	0.13	1				1			1	1	0
pentobarbital	1.20	1				1			1	1	0
sulfadiazine	1.26							1	0	1	0
chlorpropamide	1.51	1						1	1	1	0
sulfamethoxazole	1.55			1				1	1	1	0
amidotrizoate	1.70			1					1	0	1
moxalactam	1.86		1			1	1		1	2	−1
probenicid	1.91	1	1					1	2	1	1
suramin	1.91		1	1				1	2	1	1
flucytosine	2.00								0	0	0

dx.doi.org/10.1021/mp400251k | Mol. Pharmaceutics 2013, 10, 3758−3768

Table A2. continued

name	CL_u, mL/min/kg	positive criteria				negative criteria			sum (+) criteria	sum (−) criteria	sum (+) − sum (−) criteria
		log $D_{7.4}$ > 0	CH_n ≥8	substituted arC > 4	CN	quaternary C atom	8-member ring	sulphonyl group			
Drugs with Low Rates of Clearance (CL_u ≤ 3 mL/min/kg, log CL_u ≤ 0.48)											
fosfomycin	2.00								0	0	0
isoxicam	2.00			1				1	1	1	0
risedronate	2.00					1			0	1	−1
valproic acid	2.00	1							1	0	1
carumonam	2.09							1	0	1	−1
cefetamet	2.45						1		0	1	−1
foscarnet	2.45								0	0	0
iothalamic_acid	2.45			1					1	0	1
cidofovir	2.51								0	0	0
enalaprilat	2.57		1						1	0	1
cefepime	2.82		1				1		1	1	0
cefotetan	2.82					1	1		0	2	−2
zoledronate	2.82					1			0	1	−1
ceftizoxime	2.88						1		0	1	−1
ceftazidime	3.00		1			1	1		1	2	−1
mezlocillin	3.00		1			1		1	1	2	−1
Drugs with Moderate Rates of Clearance (100 > CL_u > 3 mL/min/kg, 2 > log CL_u > 0.48)											
cefixime	3.24						1		0	1	−1
sulbenicillin	3.24		1			1		1	1	2	−1
ticarcillin	3.24		1			1			1	1	0
ampicillin	3.31		1			1			1	1	0
tenofovir	3.31								0	0	0
clavulanic acid	3.39								0	0	0
cephradine	3.47		1				1		1	1	0
tenoxicam	3.55							1	0	1	−1
carbenicillin	3.63		1			1			1	1	0
warfarin	3.63	1	1						2	0	2
aztreonam	3.72					1		1	0	2	−2
cilastatin	3.80		1			1			1	1	0
sulfisoxazole	3.80			1				1	1	1	0
amoxicillin	3.89		1			1			1	1	0
azlocillin	3.89		1			1			1	1	0
cephaloridine	4.17		1				1		1	1	0
fludarabine	4.17								0	0	0
tolbutamide	4.17	1	1					1	2	1	1
cefotaxime	4.47						1		0	1	−1
cephalexin	4.57		1				1		1	1	0
cefazolin	4.90						1		0	1	−1
cefoxitin	4.90		1			1	1		1	2	−1
cefprozil	4.90		1				1		1	1	0
cephapirn	4.90		1				1		1	1	0
ceftriaxone	5.01						1		0	1	−1
methotrexate	5.62		1	1					2	0	2
cefuroxime	6.17		1				1		1	1	0
cefadroxil	6.46		1				1		1	1	0
piperacillin	7.94		1			1			1	1	0
cephalothin	8.13		1				1		1	1	0
cefatrizine	8.32		1				1		1	1	0
sulbactam	8.32					1		1	0	2	−2
cefmetazole	10.00				1	1	1		1	2	−1
perindoprilat	10.72		1						1	0	1
epristeride	10.96	1	1			1			2	1	1
valsartan	12.30	1	1	1					3	0	3
cefamandole	14.45		1				1		1	1	0
phenoxymethylpenicillin	15.14		1			1			1	1	0
methicillin	15.49		1			1			1	1	0
acetazolamide	16.22							1	0	1	−1

dx.doi.org/10.1021/mp400251k | *Mol. Pharmaceutics* 2013, 10, 3758−3768

Table A2. continued

name	CL_u, mL/min/kg	positive criteria				negative criteria			sum (+) criteria	sum (−) criteria	sum (+) − sum (−) criteria
		log $D_{7.4}$ > 0	CH_n ≥8	substituted arC > 4	CN	quaternary C atom	8-member ring	sulphonyl group			
Drugs with Moderate Rates of Clearance (100 > CL_u > 3 mL/min/kg, 2 > log CL_u > 0.48)											
captopril	16.60								0	0	0
phenethicillin	16.98		1			1			1	1	0
penicillin_G	17.38		1			1			1	1	0
propylthiouracil	17.38	1							1	0	1
acetylsalicylic acid	17.78								0	0	0
acetylcysteine	18.20								0	0	0
doxifluridine	18.20								0	0	0
cefoperazone	18.62		1				1		1	1	0
sulfinpyrazone	19.95		1						1	0	1
irbesartan	22.91	1	1	1		1			3	1	2
5-aminosalicilic acid	23.99								0	0	0
nafcillin	25.12		1			1			1	1	0
glipicide	28.18	1	1	1				1	3	1	2
pravastatin	28.18		1						1	0	1
fosfluconazole	30.90		1			1			1	1	0
tesaglitazar	35.48		1	1				1	2	1	1
glyburide	38.90	1	1	1				1	3	1	2
meloxicam	39.81			1				1	1	1	0
5-fluorouracil	40.74								0	0	0
methohexital	44.67	1	1			1			2	1	1
ketorolac	51.29		1				1		1	1	0
torsemide	52.48	1	1	1				1	3	1	2
betamipron	53.70								0	0	0
fosinoprilat	53.70		1						1	0	1
piretanide	53.70		1	1				1	2	1	1
bosentan	56.23		1	1		1		1	2	2	0
flucloxacillin	56.23		1	1		1			2	1	1
nitrofurantoin	57.54								0	0	0
thiopental	58.88	1				1			1	1	0
dicloxacillin	60.26		1	1		1			2	1	1
nateglinide	61.66	1	1						2	0	2
diflunisal	63.10	1		1					2	0	2
folinic acid	66.07		1						1	0	1
micafungin	74.13		1	1				1	2	1	1
bumetanide	81.28		1	1				1	2	1	1
cilomilast	83.18	1	1		1	1			3	1	2
oxacillin	89.13		1	1		1			2	1	1
rosuvastatin	91.20		1	1				1	2	1	1
Drugs with High Rates of Clearance (CL_u ≥ 100 mL/min/kg, log CL_u ≥ 2)											
glimepyride	100.00	1	1					1	2	1	1
pantoprazole	109.65	1	1	1					3	0	3
eprosartan	112.20	1	1	1					3	0	3
suprofen	125.89		1						1	0	1
indomethacin	128.82		1	1					2	0	2
furosemide	131.83			1				1	1	1	0
ibuprofen	138.04	1	1						2	0	2
atovaquone	151.36	1	1	1					3	0	3
dexloxiglumide	154.88	1	1						2	0	2
lovastatin	165.96	1	1						2	0	2
milrinone	177.83	1			1				2	0	2
levosimendan	190.55				1				1	0	1
ketoprofen	199.53		1						1	0	1
chlorambucil	281.84	1	1						2	0	2
cerivastatin	288.40	1	1	1					3	0	3
montelucast	338.84	1	1	1		1			3	1	2
repaglinide	524.81	1	1	1					3	0	3
entacapone	602.56	1		1	1				3	0	3

dx.doi.org/10.1021/mp400251k | *Mol. Pharmaceutics* 2013, 10, 3758−3768

Table A2. continued

name	CL_u, mL/min/kg	positive criteria: log $D_{7.4}$ > 0	positive criteria: CH_n ≥8	positive criteria: substituted arC > 4	positive criteria: CN	negative criteria: quaternary C atom	negative criteria: 8-member ring	negative criteria: sulphonyl group	sum (+) criteria	sum (−) criteria	sum (+) − sum (−) criteria
Drugs with High Rates of Clearance (CL_u ≥ 100 mL/min/kg, log CL_u ≥ 2)											
diclofenac	707.95	1	1	1					3	0	3
losartan	812.83	1	1	1					3	0	3
tezosentan	954.99		1	1				1	2	1	1
bromfenac	1174	1	1	1					3	0	3
quercetin	1230	1		1					2	0	2
tolcapone	1585	1		1					2	0	2
verlucast	1698	1	1	1					3	0	3
fluvastatin	2042	1	1	1					3	0	3
telmisartan	2089	1	1	1					3	0	3

AUTHOR INFORMATION

Corresponding Author

*Telephone: +359 2 9236514. Fax: +359 2 9879874. E-mail: zzhivkova@pharmfac.acad.bg.

Notes

The authors declare no competing financial interest.

REFERENCES

(1) Kola, I.; Landis, J. Can the pharmaceutical industry reduce attrition rates? *Nature Reviews/Drug Discovery* **2004**, *3*, 711−715.

(2) Boobis, A.; Gundert-Remy, U.; Kremers, P.; Macheras, P.; Pelkonen, O. In silico prediction of ADME and pharmacokinetics. Report of an expert meeting organized by COST B15. *Eur. J. Pharm. Sci.* **2002**, *17*, 183−193.

(3) Butina, D.; Segall, M. D.; Frankcombe, K. Predicting ADME properties in silico: methods and models. *DDT* **2002**, *7*, S83−S88.

(4) Chohan, K. K.; Paine, S. W.; Waters, N. J. Advancements in predictive in silico models for ADME. *Curr. Chem. Biol.* **2008**, *2*, 215−228.

(5) Ekins, S.; Waller, C. I.; Swaan, P. W.; Cruciani, G.; Wrighon, S. A.; Wikel, J. H. Progress in prediction human ADME parameters in silico. *J. Pharmacol. Toxicol. Methods* **2000**, *44*, 251−272.

(6) Mager, D. E. Quantitative structure-pharmacokinetic/pharmacodynamic relationships. *Adv. Drug. Deliv. Rev.* **2006**, *58*, 1326−1356.

(7) Van de Waterbeemd, H.; Gifford, E. ADMET in silico modeling: Towards prediction paradise? *Nat. Rev./Drug Discovery* **2000**, *2*, 192−204.

(8) Wang, J.; Hou, T. Recent advances on in silico ADME modeling. . In *Annual reports in Computational Chemistry*; Wheelerr, R. A., Ed.; Elsevier: Amsterdam, San Diego, 2009; Vol. *5*, pp 102−127.

(9) Yamashita, F.; Hashida, M. In silico approaches for predicting ADME properties of drugs. *Drug. Metab. Pharmacokinet.* **2004**, *19*, 327−338.

(10) Rowland, M.; Tozer, T. N. *Clinical Pharmacokinetics and Pharmacodynamics. Concepts and Applications*, 4th ed.; Wolters Kluwer/Lippincott Williams & Wilkins: Philadelphia, 2011.

(11) Toutain, P. L.; Bousquet-Mélou, A. Plasma clearance. *J. Vet. Pharmacol. Therap.* **2004**, *27*, 415−425.

(12) Shargel, L.; Wu-Pong, S.; Yu, A. B. C. *Applied Biopharmaceutics and Pharmacokinetics*; 5th ed.; McGraw-Hill: Boston, 2005.

(13) Shitara, Y.; Harie, T.; Sugiarma, Y. Transporters as a determinant of drug clearance and tissue distribution. *Eur. J. Pharm. Sci.* **2006**, *27*, 425−446.

(14) Ho, R. H.; Kim, R. B. Transporters and drug therapy: Implications to drug disposition and disease. *Clin. Pharmacol. Ther.* **2005**, *78*, 260−277.

(15) Karalis, V.; Tsaltili-Kakoulidou, A.; Macheras, P. Quantitative structure − pharmacokinetic relationships for disposition parameters of cephalosporines. *Eur. J. Pharm. Sci.* **2003**, *20*, 115−123.

(16) Ng, C.; Xiao, Y. D.; Putnam, W.; Lum, B.; Tropsha, A. Quantitative structure − pharmacokinetic parameters relationships (QSPKR) analysis of antimicrobial agents in humans using simulated annealing k-nearest-neighbor and partial least-square analysis methods. *J. Pharm. Sci.* **2004**, *93*, 2535−2544.

(17) Paine, S. W.; Barton, P.; Bird, J.; Denton, R.; Menochet, K.; Smith, A.; Tomkinson, N. P.; Chohan, K. K. A rapid computational filter for predicting the rate of human renal clearance. *J. Mol. Graphics Model.* **2010**, *29*, 529−537.

(18) Turner, J. V.; Maddalena, D. J.; Culter, D. J.; Agatonovic-Kustrin, S. Multiple pharmacokinetic parameter prediction for a series of cephalosporins. *J. Pharm. Sci.* **2003**, *92*, 552−559.

(19) Turner, J. V.; Maddalena, D. J.; Culter, D. J. Pharmacokinetic parameter prediction from drug structure using artificial neural networks. *Int. J. Pharm.* **2004**, *270*, 209−219.

(20) Yap, C. W.; Li, Z. R.; Chen, Y. Z. Quantitative structure − pharmacokinetic relationships for drug clearance by using statistical learning methods. *J. Mol. Graphics Model.* **2006**, *24*, 383−395.

(21) Yu, M. J. Predicting total clearance in humans from chemical structure. *J. Chem. Inf. Model.* **2010**, *50*, 1284−1294.

(22) Cucurull-Sanchez, L. Successful identification of key chemical structure modifications that lead to improved ADME profiles. *J. Comput.-Aided Mol. Des.* **2010**, *24*, 449−458.

(23) Demir-Kavuk, O.; Bentzien, J.; Muegge, I.; Knapp, E.-W. DemQSAR: predicting human volume of distribution and clearance of drugs. *J. Comput.-Aided Mol. Des.* **2011**, *25*, 1121−1133.

(24) Ekins, S.; Obach, R. S. Three-dimensional quantitative structure activity relationship computational approaches for prediction of human in vitro intrinsic clearance. *J. Pharmacol. Exp. Ther.* **2000**, *295*, 463−473.

(25) Nikolic, K.; Agababa, D. Prediction of hepatic microsomal intrinsic clearance and human clearance values for drugs. *J. Mol. Graphics Model.* **2009**, *28*, 254−252.

(26) Zhivkova, Z.; Doytchinova, I. Prediction of steady-state volume of distribution of acidic drugs by quantitative structure-pharmacokinetics relationships. *J. Pharm. Sci.* **2012**, *101*, 1253−1266.

(27) Zhivkova, Z.; Doytchinova, I. Quantitative stricture-plasma protein binding relationships of acidic drugs. *J. Pharm. Sci.* **2012**, *101*, 4627−4641.

(28) Obach, R. S.; Lombardo, F.; Waters, N. J. Trend analysis of a database of intravenous pharmacokinetic parameters in humans for 670 drug compounds. *Drug Metab. Dispos.* **2008**, *36* (7), 1385−1405.

(29) http://www.drugbank.ca. Last accessed April 23, 2013.

(30) htpp://www.chemicalbook.com. Last accessed April 23, 2013.

(31) Hall, L. H.; Kier, L. B. Electrotopological state. Structure modeling for QSAR and database analysis. In *Topological indices and*

related descriptors in QSAR and QSPR; Devillers, J., Balaban, A., Eds.; Gordon and Breach: London, 1999; pp 491−562.
(32) Hall, L. H.; Kier, L. B. Molecular connectivity chi indices for database analysis and structure-property modeling. In *Topological Indices and Related Descriptors in QSAR and QSPR*; Devillers, J., Balaban, A., Eds.; Gordon and Breach: London, 1999; pp 307−360.
(33) Leardi, R. Genetic algorithms in chemistry. *J. Chromatogr., A.* **2007**, *1158*, 226−233.
(34) Friedman, J. *Multivariate Adaptive Regression Splines*; Technical report No. 102. Laboratory for Computational Statistics, Department of Statistics, Stanford University: Stanford, CA, 1988.
(35) Back, S. E.; Krutzen, E.; Nilsson-Ehle, P. Contrast media as markers for glomerular filtration: a pharmacokinetic comparison of four agents. *Scand. J. Clin. Lab. Invest.* **1988**, *48*, 247−253.
(36) Anderson, G. D. Phenobarbital and other barbiturates. Chemistry, biotransformation, and pharmacokinetics. In *Antiepileptic Drugs*; Levy, R., Mattson, R. H., Meldrum, B. S., Perucca, E., Eds.; Lippincott Williams & Wilkins: Philadelphia, 2002; pp 496−503.
(37) Foulds, G.; Stankevich, J. P.; Marshall, D. C.; O'Brian, M. M.; Hayes, S. L.; Weidler, D. J.; McMahon, F. G. Pharmacokinetics of sulbactam in humans. *Antimicrob. Agents Chemother.* **1983**, *23*, 692−699.
(38) Yamashita, W.; Maeda, K.; Hirouchi, M.; Adachi, Y.; Hu, Z.; Sugiyama, Y. Involvment of transporters in the hepatic uptake and biliary excretion of valsartan, a selective antagonist of the angiotensin II AT1-receptor, in humans. *Drug. Metab. Dispos.* **2006**, *34*, 1247−1254.
(39) Haines, S. T.; Dowling, T.; Harvey, R. D. Warfarin. In *Applied Pharmacokinetics and Pharmacodynamics*, 4th ed.; Burton, M. E., Shaw, L. M., Schentag, J. J., Evans, W. E., Eds.; Lippincott Williams & Wilkins: Philadelphia; 2006; pp 715−751.
(40) Fagerholm, U. Prediction of human pharmacokinetics − renal metabolic, and excretion clearance. *J. Pharm. Pharmacol.* **2007**, *59*, 1463−1471.
(41) Smith, D. A.; Allerton, C.; Kalgutkar, A.; Van der Waterbeemd, H.; Walker, D. K. In *Pharmacokinetics and Metabolism in Drug Design*, 3rd ed.; Mannhold, R., Kubinyi, H., Folkers, G., Eds.; Wiley-VCH Verlag & Co.: Weinheim, Germany, 2012; pp 103−110.
(42) Kurihara, A.; Naganuma, H.; Hisaoka, M.; Tokiwa, H.; Kawahara, Y. Prediction of human pharmacokinetics of panipenem-betamipron, a new carbapenem, from animal data. *Antimicrob. Agents Chemother.* **1992**, *36*, 1810−1816.
(43) Winter, H. L.; Unadkat, J. D. Identification of cytochrome P450 and arylamine N-acetyltransferase isoforms involved in sulfadiazine metabolism. *Drug. Metab. Dispos.* **2005**, *33*, 969−976.
(44) Niemi, M.; Cascorby, I.; Timm, R.; Kroemer, H. K.; Neuvonen, P. J.; Kivisto, K. T. Glyburide and glimepiride pharmacokinetics in subjects with different CYP2C9 genotypes. *Clin. Pharm. Ther.* **2002**, *42*, 323−332.