Available online at www.sciencedirect.com
SCIENCE DIRECT®
International Journal of Pharmaceutics 319 (2006) 82–97
INTERNATIONAL JOURNAL OF PHARMACEUTICS
www.elsevier.com/locate/ijpharm

QSPR models for the prediction of apparent volume of distribution

Taravat Ghafourian [a,b,*], Mohammad Barzegar-Jalali [a], Siavoush Dastmalchi [c], Tina Khavari-Khorasani [a], Nasim Hakimiha [a], Ali Nokhodchi [b]

[a] *Drug Design and Chemometrics Laboratory, Drug Applied Research Centre and School of Pharmacy, Tabriz University of Medical Sciences, Tabriz, Iran*
[b] *Medway School of Pharmacy, Universities of Kent and Greenwich, Central Avenue, Chatham Maritime, Kent ME4 4TB, United Kingdom*
[c] *Biotechnology Research Center, Tabriz University of Medical Sciences, Tabriz, Iran*

Received 9 November 2005; received in revised form 24 March 2006; accepted 28 March 2006
Available online 7 April 2006

Abstract

An estimate of volume of distribution (V_d) is of paramount importance both in drug choice as well as maintenance and loading dose calculations in therapeutics. It can also be used in the prediction of drug biological half life. This study employs quantitative structure–pharmacokinetic relationship (QSPR) techniques for the prediction of volume of distribution. Values of V_d for 129 drugs were collated from the literature. Structural descriptors consisted of partitioning, quantum mechanical, molecular mechanical, and connectivity parameters calculated by specialized software and pK_a values obtained from ACD labs/log *D* database. Genetic algorithm and stepwise regression analyses were used for variable selection and model development. Models were validated using a leave-many-out procedure. QSPR analyses resulted in a number of significant models for acidic and basic drugs separately, and for all the drugs. Validation studies showed that mean fold error of predictions for the selected models were between 1.79 and 2.17. Although separate QSPR models for acids and bases resulted in lower prediction errors than models for all the drugs, the external validation study showed a limited applicability for the equation obtained for acids. Therefore, the universal model that requires only calculated structural descriptors was recommended. The QSPR model is able to predict the volume of distribution of drugs belonging to different chemical classes with a prediction error similar to that of the other more complicated prediction methods including the commonly practiced interspecies scaling. The structural descriptors in the model can be interpreted based on the known mechanisms of distribution and the molecular structures of the drugs.

Keywords: Volume of distribution; QSPR; Pharmacokinetics; V_d; In silico; Prediction

1. Introduction

The concentration of drug in the plasma or tissues depends on the amount of drug systemically absorbed and the volume in which the drug is distributed as well as the clearance. The apparent volume of distribution in the body, V_d, is a key pharmacokinetic parameter which determines the extent of drug distribution. It is simply a proportionality constant which relates the amount of drug in the body and/or compartments of the body to its plasma concentration. Therefore, despite the fact that V_d has no physical or anatomical meaning, it represents a measure of the relative partitioning of drug between plasma (the central compartment) and the tissues. An estimate of V_d is of paramount importance both in drug choice as well as maintenance and loading dose calculations in therapeutics. Moreover, V_d, in conjunction with clearance, are the two pharmacokinetic parameters determining the drug biological half life. A number of different V_d terms have been defined in the literature. Eq. (1) represents the V_d of the central compartment as the dose taken divided by the plasma concentration of the drug at time zero (C_0) (Shargel and Yu, 1999).

$$V_d = \frac{\text{Dose}}{C_0} \quad (1)$$

Two different terms have been used to describe the volume of distribution for drugs that follow multiple exponential decay. The first, designated $V_{d_{area}}$, is calculated as the ratio

* Corresponding author. Tel.: +44 1634 883846; fax: +44 1634 883927.
E-mail address: t.ghafourian@kent.ac.uk (T. Ghafourian).

0378-5173/$ – see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.ijpharm.2006.03.043

of clearance to the rate of decline of concentration during the elimination phase of the logarithmic concentration versus time curve:

$$V_{d_{area}} = \frac{\text{Dose}}{k\text{AUC}} \quad (2)$$

The second volume term is the volume of distribution at steady state ($V_{d_{ss}}$) which represents the volume in which a drug would appear to be distributed during steady state if the drug existed throughout that volume at the same concentration as that in the measured fluid (plasma or blood). It should be mentioned that when using pharmacokinetics to make drug dosing decisions, the difference between $V_{d_{area}}$ and $V_{d_{ss}}$ is not usually clinically significant (Wilkinson, 2001).

The volume of distribution in man is traditionally predicted from *in vivo* data in preclinical animals with appropriate scaling to man (Smith et al., 2001). This can be based on allometric scaling using the body weight (BW) of the species as is represented by Eq. (3).

$$V_d = a\text{BW}^b \quad (3)$$

where a and b are regression coefficients, and b is ca. 0.9–1.0. In cases where plasma protein binding varies across the species, allometric scaling should be based upon the volume of distribution corrected for the extent of protein binding (V_d of unbound drug). The more successful animal scaling prediction methods include the scaling of fractal volume of distribution (Karalis et al., 2001) and a method based on the incorporation of unbound fraction of drug in tissues of animals as well as human plasma protein binding values for the estimation of V_d in human taking into account physiological parameters such as extracellular fluid and plasma volumes (Obach et al., 1997).

Quantitative structure–pharmacokinetic relationships (QSPRs) offer a convenient alternative to animal scaling (Van de Waterbeemd, 2005). This is particularly of interest in view of the need for high-throughput *in vitro* screening of absorption, distribution, metabolism, and excretion (ADME) in earlier stages of drug development process. Previous QSPR studies have focused on classification into different ratings of volume of distribution (Hirono et al., 1994) as well as regression models for congeneric series of molecules (Gobburu and Shelver, 1995; Turner et al., 2003), and structurally unrelated drugs (Ritschel et al., 1995; Karalis et al., 2002; Lombardo et al., 2004; Ghafourian et al., 2004). In a previous study we developed QSPR models for the prediction of V_d of structurally unrelated drugs (Ghafourian et al., 2004). In this investigation, a larger numbers of drug V_d values have been collated from the literature and a wider range of structural descriptors have been incorporated. The special emphasis of the present investigation is on the interpretability of the models and rigorous leave-many-out validation process. The large number of compounds used in the study as well as the fact that they cover a wide range of chemical and pharmacological classes can add to the significance and consistency of the models. Predictions have been made based on different QSPR models for acidic drugs and basic drugs separately, as well as for all the drugs together. A comparison with previous QSPR, and other prediction methods, is made.

2. Materials and methods

2.1. Dataset

Volume of distribution was collected from the literature for 129 drug entities, belonging to different pharmacological and chemical classes (Ritschel et al., 1995; Moffat et al., 1986; Perry, 2002; Ritschel and Hammer, 1980; Ritschel, 1976; Durnas et al., 1990; Raaflaub and Speiser-Courvoisier, 1974; Lam et al., 1997; Nattell et al., 1987; Schoerlin et al., 1990; Glare and Walsh, 1991; Sonne et al., 1988; Greenblatt, 1981; Fulton and Sorkin, 1995). These included, among others, benzodiazepines, barbiturates, NSAIDs, tricyclic anti-depressants, antibiotics such as betalactams, tetracyclines, and quinolones. Table 1 is a list of the drugs together with the V_d values collected from the literature.

2.2. Structural descriptors

About 250 descriptors calculated by the TSAR 3D version 3.3 for Windows (Accelrys Ltd., USA), ACD Labs/log D Suite release 7.0 (Advanced Chemistry Development Inc., Ontario, Canada), and QSARis version 1.1 (SciVision, Academic Press, San Diego, CA), NEMESIS (software distributed by Oxford Molecular Ltd., Oxford, UK) software packages were used in this study. The descriptors included electronic parameters calculated by VAMP (using the AM1 Hamiltonian) in TSAR, atom and group counts, molecular weight and surface area and volume calculated by TSAR, partitioning parameters of log P, log D at various pH values calculated by the ACD/log D suite, and topological, shape, and three dimensional parameters calculated by QSARis. Values of pK_a were collected from the literature and ACD/pK_a database. Where the experimental values were not available calculated values by the ACD/log D suite were used instead. Fraction of ionized as cations (fiB) and anions (fiA) and the fraction unionized (fu) were calculated according to Eqs. (4)–(6) (note that only the first acidic pK_a and the first basic pK_a were considered):

$$\text{fiB} = \frac{1}{1 + \text{antilog}(7.4 - \text{p}K_a)} \quad (4)$$

$$\text{fiA} = \frac{1}{1 + \text{antilog}(\text{p}K_a - 7.4)} \quad (5)$$

$$\text{fu} = (1 - \text{fiB}) \times (1 - \text{fiA}) \quad (6)$$

Before the calculation of three dimensional descriptors, energy minimization was performed using COSMIC Force Field and AM1 Hamiltonian. Molecular descriptors calculated by the software packages were checked and omitted if more than 98% of the values were the same. Highly intercorrelated descriptors with squared correlation coefficient of >0.998 were also discarded. The analyses were performed using a reduced dataset of 210 descriptors for the 129 compounds.

Table 1
Drugs used in the study together with the volume of distribution (V_d) and plasma protein binding (ppb) values from the literature and pK_a obtained from ACD/log *D* suite

Drug	V_d	ppb (%)	pK_a (acid)	pK_a (base)
Acetanilide	0.16[a]	*	N/A	0.5[r]
Alprazolam	0.86[a,b]	70.0[a]	N/A	2.4[r]
Amitriptyline	10.64[a,c]	94.0[a]	N/A	9.4[r]
Amobarbital	1.05[a,c]	55.0[a]	7.94[r]	N/A
Amoxicillin	0.85[c,d]	17.5[d]	2.6[r]	7.4[r]
Amphetamine	3.91[a,c]	27.5[a]	N/A	9.94[r]
Ampicillin	0.61[c,d]	20.0[d]	2.6[r]	7.1[r]
Bromazepam	1.13[e,f]	55.0[a]	11[r]	2.9[r]
Bupivacaine	1.00[a]	90.0[a]	N/A	8[r]
Bupropion	16.10[a,b]	85.0[a]	N/A	7[r]
Butorphanol	5.00[a]	90.0[a]	N/A	8.6[r]
Caffeine	0.55[a,c]	35.0[a]	14[r]	0.63[r]
Carbamazepine	1.20[a,b]	75.0[a]	14.07[r]	−0.46[r]
Carbenicillin	0.25[d]	50.0[d]	2.6[r]	−1.8[p]
Cefazolin	0.17[c,d]	80.0[d]	2.1[r]	−0.21[p]
Cephalexin	0.34[c,d]	15.0[d]	2.7[r]	7[r]
Cephaloridine	0.23[d]	20.0[d]	2.4[p]	−1.6[p]
Cephalothin	0.42[c,d]	65.0[d]	2.7[p]	−1.6[p]
Cephradine	0.44[c]	20.0[d]	2.63[r]	7.35[r]
Chloral hydrate	0.60[a]	35.0[a]	10[r]	N/A
Chloramphenicol	0.57[d]	63.0[d]	11[p]	−1.7[p]
Chlordiazepoxide	0.36[a,c]	93.5[a]	N/A	4.8[r]
Chloroquine	112.4[d]	55.0[d]	N/A	9.94[r]
Chlorphentermine	2.50[a]	*	N/A	9.6[r]
Chlortetracycline	1.61[c,d]	50.5[d]	6.8[r]	11.01[p]
Clindamycin	1.00[d]	92.0[d]	12.9[p]	7.7[r]
Clobazam	1.12[e]	85.0[a]	8.59[r]	−2.06[r]
Clomipramine	17.00[a]	92.5[a]	N/A	9.46[r]
Clonazepam	3.13[a,c]	85.0[a]	10.5[r]	1.5[r]
Clorazepate	0.92[a,b]	97.0[a]	3.5[r]	N/A
Cloxacillin	0.15[d]	94.0[d]	2.4[p]	−3.5[p]
Cyclacillin	0.40[c]	24.0[d]	2.6[r]	7.2[r]
Demeclocycline	1.79[d]	70.0[d]	3.3[r]	11[p]
Desipramine	30.25[a,b]	80.0[a]	N/A	10.2[r]
Diazepam	1.92[b,c,g]	98.5[a]	N/A	3.3[r]
Dicloxacillin	0.13[d]	96.0[d]	2.4[p]	−3.7[p]
Doxepine	12.70[a,b]	80.0[a]	N/A	8[r]
Doxycycline	0.74[c,d]	56.0[d]	3.4[r]	10.8[p]
Ethambutol	1.34[c,d]	31.5[a]	14.4[p]	9.5[r]
Ethclorvynol	2.50[a]	60.0[a]	12.06[r]	N/A
Ethosuximide	0.67[a,c]	0.0[a]	9.3[r]	N/A
Etidocaine	2.00[a]	94.0[a]	N/A	7.7[r]
Fentanyl	3.60[e]	80.0[a]	N/A	8.4[r]
Flunitrazepam	4.00[a]	78.0[a]	N/A	1.8[r]
Fluoxetine	44.50[a,b]	94.0[a]	N/A	10.06[r]
Gentamycin	0.28[d]	30.0[d]	13.3[p]	10[p]
Glutethimide	3.12[a,c]	54.0[a]	4.52[r]	N/A
Griseofulvin	1.73[c]	80.0[d]	N/A	N/A
Haloperidol	15.84[b,c]	90.0[a]	N/A	8.3[r]
Hetacillin	0.40[d]	20.0[d]	2.45[p]	4.97[p]
Ibuprofen	0.10[a]	99.0[a]	5.2[r]	N/A
Imipramine	16.50[a,b]	90.0[a]	N/A	9.5[r]
Indomethacin	0.95[a,c]	95.0[a]	4.5[r]	N/A
Isoniazid	0.64[c,d]	15.0[d]	10.79[r]	3.52[p]
Kanamycin	0.19[d]	0.0[d]	12.9[p]	9.5[p]
Ketamine	4.00[a]	35.0[a]	N/A	7.5[r]
Lidocaine	1.41[n,c,g,h]	70.0[a]	N/A	7.86[r]
Lincomycin	0.33[d]	72.0[d]	12.91[p]	7.65[r]
Lorazepam	1.20[a,c]	90.0[a]	11.5[r]	1.3[r]
Maprotiline	21.00[a,b]	90.0[a]	N/A	10.02[r]
Meprobamate	0.70[a,c]	20.0[a]	13.36[r]	N/A
Meptazinol	5.47[e]	27.0[a]	N/A	8.7[r]
Methacycline	0.97[d]	75.0[d]	3.5[r]	10.73[p]

Table 1 (*Continued*)

Drug	V_d	ppb (%)	pK_a (acid)	pK_a (base)
Methadone	3.41[a,c]	90.0[a]	N/A	8.25[r]
Methaqualone	6.00[a]	85.0[a]	N/A	2.54[r]
Meticillin	0.36[c,d]	57.1[d]	2.8[r]	−2.1[p]
Metoclopramide	3.00[a]	65.0[a]	N/A	9[r]
Midazolam	1.30[a,b]	95.0[a]	N/A	6.2[r]
Minocycline	0.75[c,d]	65.5[d]	5[r]	2.8[r]
Moclobemide	1.29[i]	50.0[e]	N/A	6.89[r]
Morphine	2.86[n,a,c,j]	25.0[a]	9.26[r]	8.18[r]
Nafcillin	0.29[d]	90.0[d]	2.7[r]	11.91[p]
Naloxone	3.00[a]	40.0[a]	N/A	7.94[r]
Neomycin(a)	0.01[d]	*,d	12.93[p]	9.52[p]
Nitrazepam	2.24[c,e]	86.5[a]	10.8[r]	3.2[r]
Nitrofurantoin	0.32[c]	74.2[d]	7.1[r]	−2.39[p]
Nortriptyline	20.50[a,b]	92.5[a]	N/A	9.73[r]
Oxacillin	0.29[c,d]	92.0[d]	2.8[r]	−3.4[p]
Oxazepam	0.99[a,c,k,l]	95.0[a]	11.1[r]	1.8[r]
Oxyphenbutazone	0.11[a,c]	99.0[a]	4.7[r]	N/A
Oxytetracycline	1.89[d]	25.0[d]	3.27[r]	9.11[r]
Paracetamol	1.21[c,e]	*	9.5[r]	N/A
Paroxetine	23.25[a,b]	95.0[a]	N/A	9.72[r]
PAS	0.23[d]	65.0[d]	3.92[r]	1.78[r]
Penicillin G	0.68[c,d]	65.0[d]	2.75[r]	−1.3[p]
Penicillin V	0.70[c,d]	80.0[d]	2.8[r]	−1.7[p]
Pethidine	5.03[c,e]	45.0[a]	N/A	8.7[r]
Phenacetin	1.31[c]	30.0[a]	N/A	2.2[r]
Phenazone	0.60[c,e]	10.0[a]	N/A	1.4[r]
Phencyclidine	6.00[a]	72.5[a]	N/A	8.5[r]
Phenethicillin K	0.35[d]	82.0[d]	12.54[p]	−1.7[p]
Phenobarbital	0.80[g,c]	50.0[a]	7.4[r]	N/A
Phenylbutazone	0.15[a,c]	99.0[a]	4.5[r]	N/A
Phenytoin	0.60[g,c]	90.0[a]	8.33[r]	N/A
Prazepam	1.50[a]	97.0[a]	N/A	2.7[r]
Primidone	0.60[a]	20.0[a]	12.26[r]	N/A
Propofol	3.50[m]	97.5[a]	11[r]	N/A
Protriptyline	14.76[a,c]	92.0[a]	N/A	10.61[r]
Pyrimethamine	2.19[c]	27.0[d]	N/A	7.34[r]
Quinine sulfate	1.63[d,c]	70.0[d]	12.8[p]	9.28[p]
Rolietracycline	0.58[d]	50.0[d]	4.5[p]	11[p]
Salicylamide	0.15[a,c]	75.0[a]	8.2[r]	N/A
Sertraline	20.00[a]	98.0[a]	N/A	9.47[r]
Spectinomycin	0.12[d]	*	9.25[p]	8.7[r]
Sulbenicillin	0.20[c]	70.0[d]	0.28[p]	−1.97[p]
Sulfadiazine	0.92[d]	42.0[d]	6.56[r]	2[r]
Sulfadimethoxine	0.41[d,c]	94.5[d]	6.21[p]	1.3[p]
Sulfadoxine	0.12[c]	91.5[d]	5.82[r]	1.59[p]
Sulfaethidole	0.18[d]	99.0[d]	5.6[r]	1.9[r]
Sulfamerazine	0.39[d,c]	75.0[d]	5.6[r]	1.6[p]
Sulfameter	0.26[c]	90.1[d]	6.69[p]	1.54[p]
Sulfamethazine	0.61[d]	80.0[d]	7.38[r]	2.36[r]
Sulfamethizole	0.35[c]	62.0[d]	5.45[r]	2.2[r]
Sulfamethopyrazine	0.22[c]	68.0[d]	6.2[r]	1.95[p]
Sulfamethoxazole	0.22[d,c]	63.5[d]	5.8[p]	1.39[p]
Sulfamethoxypyridazine	0.19[c]	80.0[d]	7.19[p]	2.18[p]
Sulfinpyrazone	0.12[c]	92.0[d]	2.8[r]	−0.66[p]
Sulfisomidine	0.32[d]	86.0[d]	7.5[r]	2.36[r]
Sulfisoxazole	0.17[d,c]	85.0[d]	4.83[p]	1.52[p]
Sulfisoxazole acetyl	1.19[d]	85.0[d]	N/A	−0.17[p]
Temazepam	0.98[b,a]	97.0[a]	N/A	1.6[r]
Tetracycline	1.40[d,c]	45.3[d]	3.3[r]	11.02[p]
Theobromine	0.75[a]	*	10.05[r]	N/A
Tinidazole	0.39[c]	*	N/A	1.82[r]
Tramadol	3.00[a]	5.0[a]	N/A	8.3[r]
Triazolam	1.10[b]	78.0[a]	N/A	8.19[r]

Table 1 (*Continued*)

Drug	V_d	ppb (%)	pK_a (acid)	pK_a (base)
Trimethoprim	2.40[d,c]	70.0[d]	N/A	7.12[r]
Valproic acid	0.17[n,a,b]	90.0[a]	5[r]	N/A
Viloxazine	1.00[a]	86.5[a]	N/A	8.1[r]

[a] Data taken from Moffat et al. (1986).
[b] Taken from Perry (2002).
[c] Taken from Ritschel and Hammer (1980).
[d] Taken from Ritschel (1976).
[e] Taken from Durnas et al. (1990).
[f] Taken from Raaflaub and Speiser-Courvoisier (1974).
[g] Taken from Lam et al. (1997).
[h] Taken from Nattell et al. (1987).
[i] Taken from Schoerlin et al. (1990).
[j] Taken from Glare and Walsh (1991).
[k] Taken from Sonne et al. (1988).
[l] Taken from Greenblatt (1981).
[m] Taken from Fulton and Sorkin (1995).
[n] Taken from Ritschel et al. (1995).
[p] Calculated values by ACD/log *D* software.
[r] Values from database.
[*] Data missing.

2.3. Development of QSPRs

All the structural descriptors were standardized by making the range between 0 and 1. Stepwise regression analysis was used to determine statistically significant relationships between structural parameters and the volume of distribution. The statistical analysis was performed using the MINITAB (release 13.1) statistical software. Because of the large number of descriptors, it is possible that the stepwise regression falls into local minima. Therefore, genetic algorithm was used for the reduction of the number of descriptors prior to further stepwise regression analyses. Thus, feature selection was performed in STATISTICA Neural Networks (version 6) using cross over rate of 1, mutation rates of 0.1 and 0.5, number of populations of 1000, and number of generations of 500. The unit penalty (factor that is multiplied by the number of selected input variables, and added to the selection error leading to reduced number of selected variables) was set to 0.01 or 0.001.

To avoid the risk of chance correlations, loss of interpretability and predictability, the number of parameters in the models was kept as low as possible. Accordingly, the stepwise was cut short when addition of the seventh or eighth parameter did not add to the interpretability and predictability of the models. QSPRs were sought for the whole dataset and also for the acidic and basic drugs separately. A compound was allocated to the acidic group of drugs if the fraction ionized as an acid (anionic fraction, fiA) was higher than the fraction ionized as a base (cationic fraction, fiB) at pH 7.4 and was allocated to the basic group if fiB was higher than fiA. While deletion of outliers often improves the statistics of a QSPR, it was decided to keep all the compounds in the study, unless they affected the coefficients of equations significantly. To test the predictive power of the models, the dataset was divided into the four equal groups. To this end, the data were ranked based on the ascending V_d values and from every four compounds one was assigned to one of the four groups. Each of the groups (containing quarter of the chemicals) were excluded one at a time from the MLR analysis as the test set and the corresponding V_d values were computed using the QSPR obtained for the rest of the chemicals (training set). Fold error of prediction for the test sets was calculated according to Eq. (7) and the mean values were reported.

$$\text{Fold error} = \text{antilog}(|\log V_{d_{obs}} - \log V_{d_{pred}}|) \quad (7)$$

The following statistical details of the models were noted: n, the number of observations; r, the correlation coefficient; s, the standard deviation; F, the Fisher statistic and the P value. The figures in parentheses with the regression coefficients were standard errors of coefficients.

3. Results

The apparent volume of distribution (V_d) and the extent of protein binding for the compounds used in this study are listed in Table 1, together with the relevant references. Also included in the table are pK_a values from ACD labs database or the calculated values. Drugs used in the study covered a wide range of chemical and pharmacological classes with the V_d values ranging from 0.1 to 112.4 L kg^{-1}. Stepwise regression and genetic algorithm led to a number of significant QSPR models from which three models were selected for all drugs, bases and acids based on the R^2 and standard deviation as well as the quality and interpretability of the descriptors in the models. The models have been presented below. One of drugs, neomycin was excluded from the analyses. The reason was that neomycin is a natural antibiotic that can occur in one of the three molecular structures of type A, B and E, with the marketed product containing a mixture of various percentages of these.

3.1. All drugs

Stepwise regression analysis led to a QSPR from which chloroquine was outlier. Chloroquine has an extremely high V_d

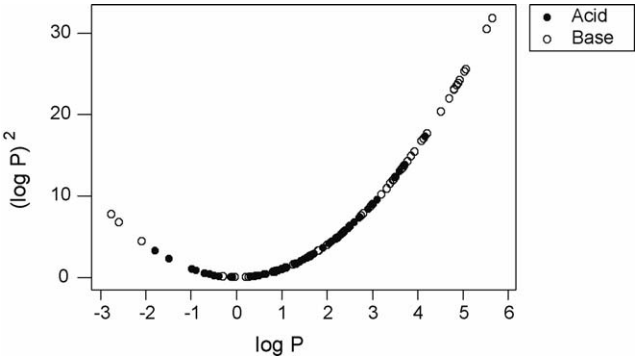

Fig. 1. The plot of squared log P vs. log P for acid and base drugs.

value of 112.4, although its inclusion does not deteriorate the statistics of the QSPR, it has a large influence on the coefficients of the equation. The exclusion of this chemical resulted in the following QSPR:

$$\log V_d = -1.82(\pm0.26) - 0.448(\pm0.07)\text{fiA} + 0.482(\pm0.08)\text{fiB} + 0.579(\pm0.15)(\log P)^2 + 1.30(\pm0.25)\text{SsssCH} + 0.694(\pm0.16)\text{Lipole} + 0.764(\pm0.27)\text{SsF} - 0.533(\pm0.23)\text{ESP}^-,$$

$$n = 125,\ s = 0.317,\ r^2 = 0.747,\ F = 49.4 \quad (8)$$

Missing log P value for cephaloridine and missing Lipole value for sulbenicillin together with the two excluded drugs (neomycin and chloroquine) has reduced the number of chemicals to 125.

In Eq. (8) fiA and fiB are fractions of the drug ionized at pH 7.4 as acids and bases respectively, $(\log P)^2$ is the square of logarithm of partition coefficient calculated by ACD/log D suite, SsssCH is the atom type electrotopological state index for >CH– groups, Lipole is the total lipole of the molecule calculated by TSAR, SsF is atom type electrotopological state index for –F groups, ESP^- is the absolute most negative electrostatic potential on the solvent accessible surface of the molecule. The negative coefficient of fiA and the positive coefficient of fiB in the QSPR show that ionization to anions reduces the V_d but base ionization to cation enhances the volume of distribution. Although the parameters of $(\log P)^2$, SsssCH and lipole (all with positive coefficients) could be regarded as lipophilicity parameters, a closer inspection of the parameters reveals also other structural information within the parameters. One particular debatable issue is the use of squared partition coefficient, rather than the log P. This could be due to lower $(\log P)^2$ values of acidic drugs in comparison with bases. Fig. 1 shows the plot between $(\log P)^2$ and log P where acids and bases have been delineated. It could be seen in the figure that acids have lower log P and $(\log P)^2$ values than most bases but the difference between the $(\log P)^2$ values are more significant. Therefore, it seems reasonable to assume that the preference of $(\log P)^2$ over log P by statistical analysis could be due to the lower V_d values of acidic drugs. Acids are known to be highly protein bound in plasma with reduced V_d values (Karalis et al., 2002). This could also be observed in the scattered plot of log V_d versus the predicted log V_d values by Eq. (8) at Fig. 2 where acids have a smaller range of V_d values. In case of SsssCH, it was observed that drugs such as antibiotics kanamycin and gentamycin (Fig. 3), having a lower number of >CH- groups, possess low values of SsssCH. These are molecules with higher percentages of heteroatoms, oxygen and nitrogen resulting in the high hydrophilicity. Total lipole is a measure of lipophilic distribution in the molecule that is calculated from summed atomic log P values.

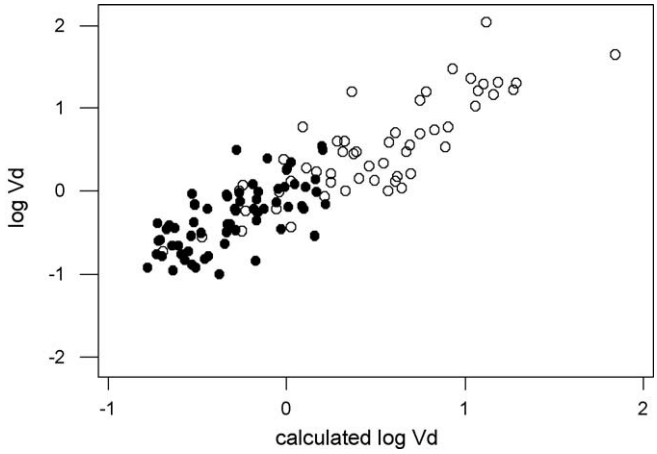

Fig. 2. Plot of observed log V_d vs. the values calculated by Eq. (8); (◯) bases and (●) acids.

Incorporation of plasma protein bound fraction (ppb) resulted in a similar QSPR where ESP^- was no longer significant:

$$\log V_d = -1.38(\pm0.23) - 0.472(\pm0.07)\text{fiA} + 0.460(\pm0.08)\text{fiB} + 0.762(\pm0.17)(\log P)^2 + 1.38(\pm0.24)\text{SsssCH} + 0.645(\pm0.15)\text{Lipole} + 0.881(\pm0.26)\text{SsF} - 0.280(\pm0.12)\text{ppb},$$

$$n = 119,\ s = 0.305,\ r^2 = 0.769,\ F = 52.7 \quad (9)$$

Eq. (9) shows the negative effect of ppb on the volume of distribution of drugs.

3.2. Acids

For acids the best QSPR, in terms of R^2 and s, was obtained from stepwise regression analysis on a reduced set of 88 descriptors obtained from feature selection by genetic algorithm. The resulting QSPR is as follows:

$$\log V_d = -0.655 + 1.01(\pm0.15)N_{6\text{-rings}} - 0.289(\pm0.097)\text{fiA} + 0.558(\pm0.16)^7\chi_{ch} - 1.12(\pm0.43)\text{SsssN} + 0.726(\pm0.32)\log D_{7.4},$$

$$n = 69,\ s = 0.268,\ r^2 = 0.554,\ F = 15.6 \quad (10)$$

log $D_{7.4}$ is not available for cephaloridine, therefore cephaloridine is excluded from Eq. (10).

In Eq. (10) the negative coefficient of fiA shows the detrimental effect of acid ionization on the volume of distribution. $N_{6\text{-rings}}$ is the number of six membered aliphatic rings with the

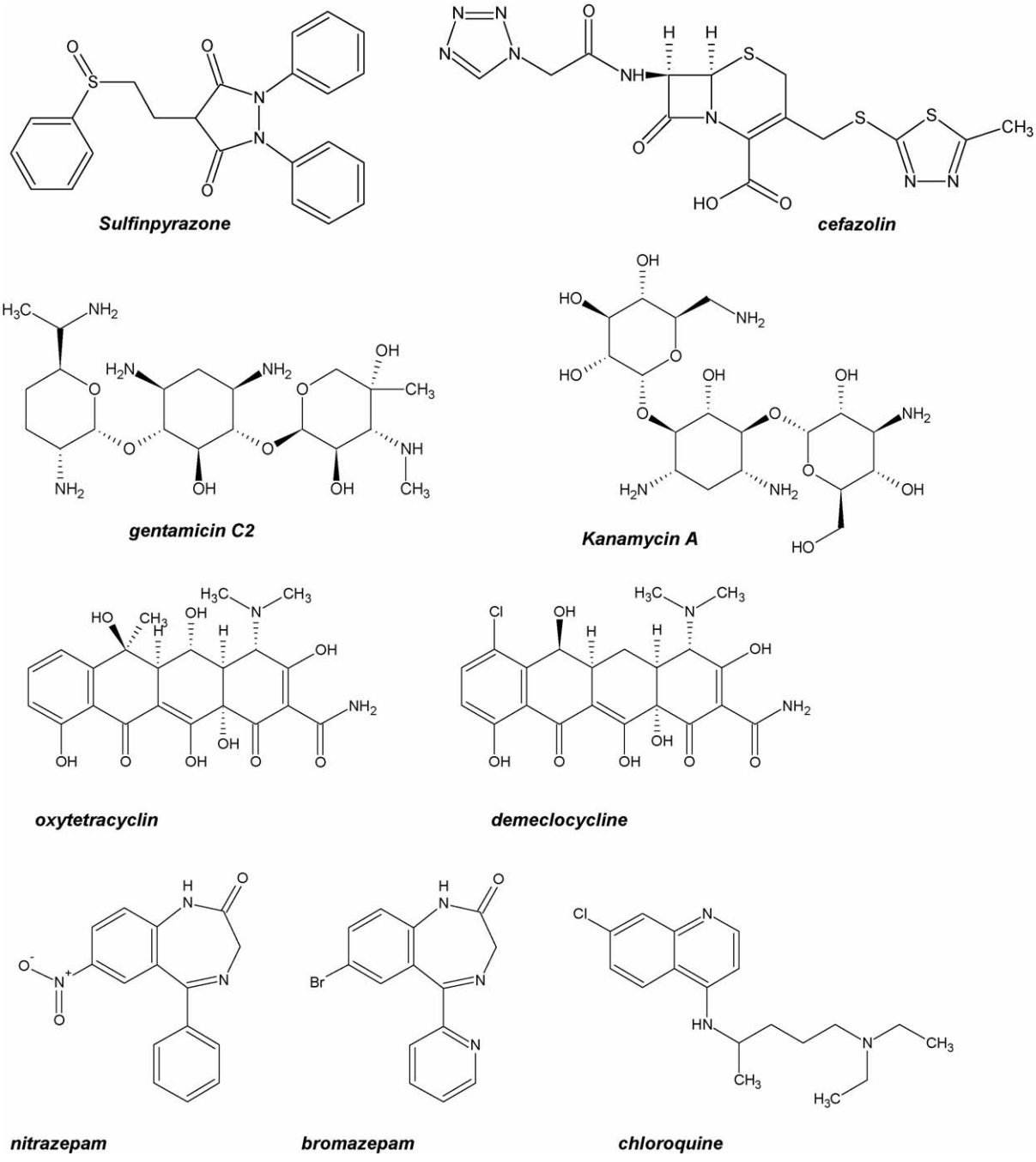

Fig. 3. Molecular structures of some of drugs in the study.

higher values corresponding to tetracyclines. Two examples of these drugs are oxytetracycline and demeclocycline which have a relatively high V_d values compared to other acidic drugs (see Fig. 3 for the molecular structures). $^{7}\chi_{ch}$ is the seventh order chain connectivity index with higher values corresponding to benzodiazepines such as nitrazepam and bromazepam (Fig. 3). SsssN is the atom type electrotopological state index for tertiary amine groups. The negative coefficient of SsssN in the equation shows that presence of these base groups will result in the reduced V_d values. Examples of such drugs are sulfinpyrazone and oxyphenbutazone (Fig. 3). Fig. 4 is plot of log V_d versus the log V_d values predicted by Eq. (10) for acids.

3.3. Bases

The following six-parameter QSPR was the best equation resulted from stepwise regression analyses:

$$\log V_d = -1.32(\pm 0.51) - 2.69(\pm 0.42)\log\left(\frac{fu}{pK_a}\right) - 1.17(\pm 0.31)ESP^{-} + 2.95(\pm 0.70)\mu - 1.27(\pm 0.40)N_{OH} - 2.21(\pm 0.46)ABSQon + 1.94(0.37)\,^{1}\kappa_a,$$
$$n = 55,\ s = 0.290,\ r^2 = 0.825,\ F = 37.6 \quad (11)$$

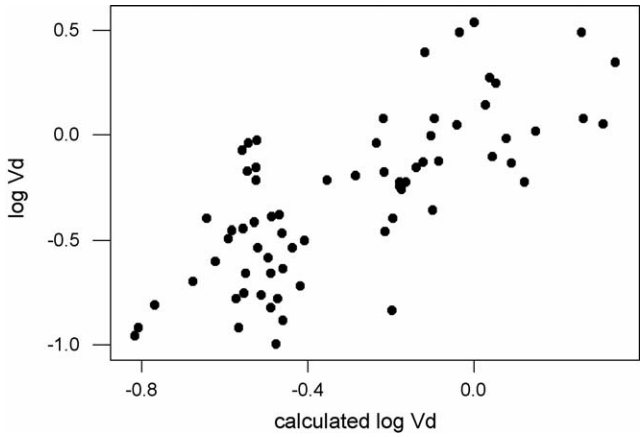

Fig. 4. Plot of observed log V_d vs. log V_d calculated by Eq. (10) for acids.

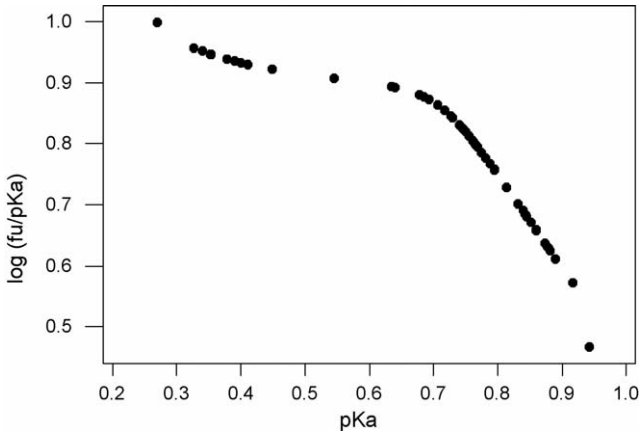

Fig. 6. Plot between log(fu/pK_a) and pK_a for bases.

Griseofulvin and sulfisoxazole acetyl have missing log(fu/pK_a) values, therefore they are not included in Eq. (11).

In Eq. (11) log(fu/pK_a) is the logarithm of unionized fraction of drugs at pH 7.4 divided by the pK_a of the base, μ the dipole moment, N_{OH} the number of hydroxyl groups, ABSQon is sum of the atomic charges on nitrogen and oxygen atoms and $^1\kappa_a$ is the first order kappa alpha shape index. Fig. 5 is the plot of observed against calculated log V_d for bases.

In order to explain the effect of log(fu/pK_a) the relationship between log(fu/pK_a) and pK_a have been presented in Fig. 6. The plot shows that for weak bases with the standardized pK_a values lower than about 0.7, an increase in pK_a values results in a slight reduction of the parameter log(fu/pK_a). On the other hand, for strong bases the correlation has a high negative slope. Therefore, Eq. (11) shows that volume of distribution is directly related to the pK_a values of strong bases, however in case of the weak bases with pK_a values lower than about 7, V_d is not so much affected by the pK_a. The parameters with negative effect are the absolute most negative electrostatic potentials on the molecular surface, number of hydroxyl groups and sum of the atomic charges on nitrogen and oxygen atoms. All these can represent hydrogen bonding abilities of the molecule (Dearden and Ghafourian, 1999), and therefore in Eq. (11) they reveal the negative effect of hydrogen bonding on V_d of bases. The positive coefficient of μ indicates the necessity of a dipolar feature in the molecule, which is possibly required for membrane transport, leading to higher V_d values. The other parameter with a positive effect is $^1\kappa_a$. Although it is considered to be a shape parameter, for this dataset it is highly correlated with size parameters such as $^0\chi$, surface area, and molar volume. In addition, $^1\kappa_a$ and the parameter ABSQon can explain 74% of variation in log P and the presence of both parameters in Eq. (11) could indicate the effect of lipophilicity.

Extent of protein binding (pbb) is not selected by the stepwise regression and forcing the pbb parameter in the regression analysis does not lead to a significant QSPR.

3.4. Validation of the QSPR models

QSARs were obtained for training sets using the descriptors of Eqs. (8)–(11) (see Appendix B) and they were used to predict V_d values for drugs within the four test sets described in the methods section. Mean fold error of prediction (MFE) and R^2_{pred} (correlation coefficient between observed and predicted V_d values) are presented in Table 2. Moreover, two alternative QSPR models for acids and bases (Eqs. (12) and (13) respectively) were also used and validated for the prediction of V_d values. The equations were used both, on their own and in a consensus

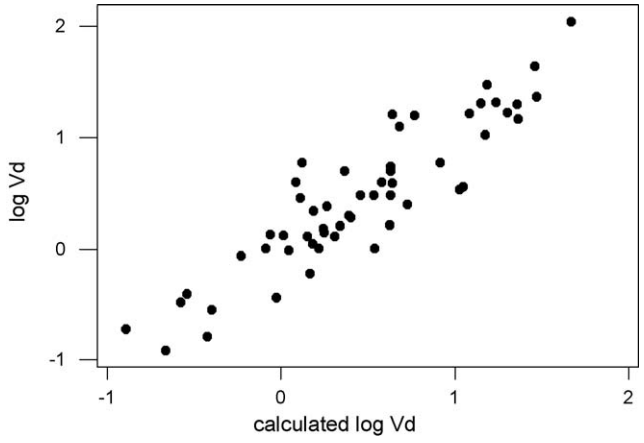

Fig. 5. Plot of observed log V_d vs. log V_d calculated by Eq. (11) for bases.

Table 2
Mean fold prediction error and the squared correlation coefficient between observed and predicted V_d values using QSPR models (8)–(13), and the consensus prediction

Prediction model	R^2_{pred}	MFE			Accuracy[a]
		All	Acids	Bases	
Eq. (8)	0.699	2.11	2.19	2.11	66.7
Eq. (9)	0.716	2.08	2.01	2.17	62.2
Eq. (10)	0.488	–	1.79	–	71.0
Eq. (11)	0.785	–	–	1.93	64.3
Eq. (12)	0.470	–	1.84	–	71.0
Eq. (13)	0.694	–	–	2.18	58.9
Consensus	0.795	1.83	1.79	1.89	72.2

[a] Percent of the total number of drugs that are predicted with less than two fold error.

prediction procedure, where the log V_d values predicted by Eqs. (8)–(13) were averaged (Table 2).

$$\log V_d = -0.572(\pm 0.22) + 0.598(\pm 0.12)N_{6\text{-rings}} - 1.43(\pm 0.44)\text{SsssN} - 0.487(\pm 0.09)\text{fiA} + 0.542(\pm 0.26)\text{Pol} - 0.419(\pm 0.19)\text{SaaN} + 0.572(\pm 0.29)\frac{\text{MV}}{\text{TA}},$$
$$n = 68,\ s = 0.266,\ r^2 = 0.575,\ F = 13.8 \quad (12)$$

$$\log V_d = 1.21(\pm 0.61) - 2.75(\pm 0.49)\log\left(\frac{\text{fu}}{\text{p}K_a}\right) + 0.776(\pm 0.33)\text{ESP}^- + 2.59(\pm 0.76)\mu - 1.09(\pm 0.33)N_{OH} + 1.31(\pm 0.32)\log P,$$
$$n = 55,\ s = 0.333,\ r^2 = 0.763,\ F = 31.5 \quad (13)$$

4. Discussion

Volume of distribution is an important pharmacokinetic property that needs to be determined during drug development process. It is normally estimated using animal scaling which is associated with certain levels of error (Mahmood, 1998). QSPR technique can offer an alternative method for the estimation, especially in early stages of drug development. One particular limitation of QSPR could be the narrow range of applicability which arises from the limited chemical space covered by the training set. In this study, a broad range of drugs have been employed to derive statistically significant QSPRs (see Fig. 3 for some examples). We have mined the literature extensively for volume of distribution, and assembled a dataset of 129 compounds comprising neutral, basic and acidic drugs. Moreover, a rigorous validation procedure involving four cycles of leave-25%-out validation, where each chemical was out once resulted in a better knowledge about the prediction capabilities of the models. Table 2 shows that it is possible to predict V_d of all drugs without separating acids and bases with the mean prediction errors of 2.11 and 2.08 folds for Eqs. (8) and (9) respectively. In other words, knowledge of extent of plasma protein binding only slightly improved the prediction of V_d. Also note that ppb was not a significant parameter in the QSPR analyses of separate groups of acids and bases. This is despite the common practice that views the unbound (intrinsic) V_d (volume of distribution divided by the non-protein bound fraction) as a better QSPR endpoint than the V_d (Ritschel et al., 1995; Ritschel and Hammer, 1980; Blakey et al., 1997). The concept of unbound V_d originates from the fact that distribution of a compound in the human body is related to the extent of binding in tissues versus the extent of binding in plasma (Rowland and Tozer, 1995). Thus, plasma protein binding can limit the concentration of drug available for distribution *in vivo* (Urien et al., 2001). The lack of a linear relationship between V_d and plasma protein binding is conceivable as the observed V_d value results from a summation of affinity profiles to various body tissues and plasma proteins. Therefore, several drugs with equal affinities to plasma proteins can have different affinities to various body tissues resulting in differing values of V_d (Urien et al., 2001). It must also be noted that although model (8) lacks plasma protein binding parameter, it has benefited from parameters that distinguish acids and bases (namely fiA, fiB and $(\log P)^2$) and the low V_d values of acids in comparison with bases could be related to the high ppb of acids like NSAIDs.

In a previous QSPR study we compared the unbound V_d and the V_d in terms of the prediction power of the resulting model and concluded that V_d, although leading to an inferior QSPR in terms of fit, resulted in QSPRs with higher predictive power (Ghafourian et al., 2004). In fact, the higher correlation coefficient observed for the QSPRs using the unbound V_d as the dependent variable could be due to a broader range of the unbound V_d in comparison with the V_d. The use of unbound V_d in QSPR has also been criticized by Davis et al. (2000) on the basis that protein binding itself is related to the structural parameters, such as partition coefficient, which are commonly used in the QSAR models.

Table 2 shows that prediction powers of the separate QSPRs for acids and bases (Eqs. (10) and (11)) are greater than the universal QSPRs. This is especially remarkable in case of acids for which R^2 and R^2_{pred} of the models (Eqs. (10) and (12)) are much lower than that of the models for bases or all drugs. This is probably due to the narrow range of V_d values for acids. Furthermore, the average predicted V_d values by all the models (Eqs. (8)–(13)), i.e. consensus modeling, have the smallest MFE level as well as the highest percent of drugs predicted within two-fold error range. A further analysis of the fold errors indicates a high estimation error of V_d values for some drugs using any of the presented models. These drugs do not belong to a particular chemical group or to a specific V_d range. The drugs associated with high prediction errors include ibuprofen, glutethimide, acetanilide, bupivacaine, methaqualone and bupropione, with minimum fold errors in the range of 3–4.8. Table 3 is the observed and predicted V_d values by different models, with the predicted values with minimum deviation for each drug highlighted in bold.

In order to further test the validity of our proposed models, the V_d values were predicted for some drugs the V_d data for which was collected from the literature (Cutler et al., 1978; Barbeau and Belanger, 1982; Klinge et al., 1982; Lombardo et al., 2004). These drugs were not included in any of the original analyses. Table 4 is the results of this external validation together with the observed V_d data. The table shows that QSAR model (8) predicts the external validation set very well with average fold error of 1.68. On the other hand, equations for acids ((10) and (12)) perform particularly badly in prediction of V_d for steroid group (prednisone, dexamethasone, hydrocortisone and prednisolone). This could be due to a limited chemical space of the drugs used for the development of this equation which results in a small applicability domain. The prediction of these equations is good for the other drugs with mean fold error values of 1.60 and 1.92 for Eqs. (10) and (12), respectively. Mean fold error of prediction for Eq. (9) is 1.86.

Table 3
The V_d values predicted by models (8)–(12) for acidic and basic drugs in test sets; the test sets in which the drugs were predicted have been outlined

Drug	$V_{d_{obs}}$	$V_{d_{pred}}$						Acid or base	Test set code
		(8)	(9)	(10)	(11)	(12)	(13)		
Amobarbital	1.05	**0.958**	0.874	1.328		1.518		Acid	b
Amoxicillin	0.85	0.301	**0.425**	0.277		0.248		Acid	b
Ampicillin	0.613	0.345	**0.545**	0.306		0.269		Acid	z
Bromazepam	1.13	0.873	**1.132**	2.149		0.876		Acid	z
Caffeine	0.55	0.434	0.631	**0.598**		0.398		Acid	z
Carbamazepine	1.2	0.905	1.172	1.846		1.328		Acid	b
Carbenicillin	0.25	0.178	0.193	**0.253**		0.340		Acid	t
Cefazolin	0.165	0.470	0.406	0.345		**0.119**		Acid	t
Cephalexin	0.34	0.378	0.557	**0.332**		0.453		Acid	z
Cephaloridine	0.23							Acid	z
Cephalothin	0.415	0.251	0.275	0.330		**0.474**		Acid	b
Cephradine	0.44	0.508	0.704	0.839		0.795		Acid	x
Chloral hydrate	0.6	0.652	**0.587**	0.635		0.548		Acid	z
Chloramphenicol	0.57	0.647	**0.539**	0.718		1.213		Acid	t
Clobazam	1.12	**0.962**	0.838	0.672		0.633		Acid	x
Clonazepam	3.127	2.105	1.609	1.782		**2.334**		Acid	b
Clorazepate	0.915	0.505	0.461	0.588		**0.957**		Acid	z
Cloxacillin	0.15	0.289	0.261	**0.307**		0.261		Acid	b
Cyclacillin	0.4	0.346	0.512	0.664		**0.434**		Acid	z
Demeclocycline	1.79	0.954	0.725	1.220		**1.294**		Acid	x
Dicloxacillin	0.13	0.358	**0.288**	0.368		0.296		Acid	t
Doxycycline	0.739	**0.765**	0.874	1.189		1.263		Acid	b
Ethclorvynol	2.5	**0.875**	0.697	0.722		0.733		Acid	z
Ethosuximide	0.665	0.875	1.169	**0.677**		0.962		Acid	b
Glutethimide	3.115	0.441	0.417	**0.686**		0.569		Acid	b
Hetacillin	0.4	**0.322**	0.617	0.173		0.150		Acid	x
Ibuprofen	0.1	0.557	0.609	**0.376**		0.483		Acid	x
Indomethacin	0.945	**0.584**	0.449	0.261		0.418		Acid	b
Isoniazid	0.64	0.775	1.056	0.560		**0.626**		Acid	t
Lorazepam	1.203	0.624	0.653	0.744		**1.400**		Acid	x
Meprobamate	0.7	1.303	1.921	**0.783**		0.886		Acid	t
Methacycline	0.97	**0.729**	0.508	1.303		1.386		Acid	x
Meticillin	0.36	0.211	0.256	**0.258**		–		Acid	x
Minocycline	0.745	0.405	0.440	**0.664**		0.618		Acid	x
Nafcillin	0.29	1.467	1.559	0.394		**0.376**		Acid	z
Nitrazepam	2.235	1.224	0.926	**2.313**		1.108		Acid	z
Nitrofurantoin	0.32	0.391	0.434	**0.263**		0.254		Acid	b
Oxacillin	0.29	0.372	**0.293**	0.278		0.239		Acid	x
Oxazepam	0.992	2.017	1.703	**0.852**		1.514		Acid	t
Oxyphenbutazone	0.11	0.364	0.233	**0.157**		0.169		Acid	z
Oxytetracycline	1.89	0.672	1.096	0.984		**1.223**		Acid	z
Paracetamol	1.205	0.629		0.545		**0.657**		Acid	z
PAS	0.23	0.384	0.707	0.389		**0.321**		Acid	x
Penicillin G	0.675	0.315	**0.323**	0.261		0.293		Acid	x
Penicillin V	0.7	**0.396**	0.304	0.306		0.290		Acid	z
Phenethicillin K	0.348	1.276	1.076	0.712		**0.391**		Acid	t
Phenobarbital	0.795	**0.778**	0.710	1.135		0.991		Acid	t
Phenylbutazone	0.154	0.484	0.373	0.139		**0.138**		Acid	x
Phenytoin	0.6	0.825	0.654	**0.628**		1.150		Acid	x
Primidone	0.6	**0.921**	1.382	1.425		0.986		Acid	t
Propofol	3.5	**2.415**	2.200	0.920		1.064		Acid	x
Salicylamide	0.145	0.981	0.736	**0.657**		0.640		Acid	b
Sulbenicillin	0.2			**0.199**		0.312		Acid	x
Sulfadiazine	0.92	0.245	0.275	**0.285**		0.205		Acid	t
Sulfadimethoxine	0.407	0.252	0.159	**0.323**		0.218		Acid	t
Sulfadoxine	0.12	0.238	**0.173**	0.269		0.290		Acid	z
Sulfaethidole	0.176	**0.185**	0.192	0.280		0.275		Acid	z
Sulfamerazine	0.385	0.234	0.200	**0.291**		0.190		Acid	t
Sulfameter	0.26	0.246	0.204	0.294		**0.264**		Acid	b
Sulfamethazine	0.61	**0.463**	0.350	0.455		0.401		Acid	x
Sulfamethizole	0.35	0.170	0.199	**0.239**		0.211		Acid	b
Sulfamethopyrazine	0.22	0.291	**0.224**	0.321		0.206		Acid	t

Table 3 (*Continued*)

Drug	$V_{d_{obs}}$	$V_{d_{pred}}$						Acid or base	Test set code
		(8)	(9)	(10)	(11)	(12)	(13)		
Sulfamethoxazole	0.22	**0.231**	0.240	0.282		0.306		Acid	z
Sulfamethoxypyridazine	0.19	**0.234**	0.250	0.389		0.324		Acid	t
Sulfinpyrazone	0.12	0.410	0.302	**0.171**		0.308		Acid	t
Sulfisomidine	0.316	0.354	**0.297**	0.406		0.404		Acid	t
Sulfisoxazole	0.165	0.221	**0.194**	0.242		0.251		Acid	b
Tetracycline	1.395	**1.371**	1.704	1.086		1.246		Acid	t
Theobromine	0.75	0.479		**0.724**		0.512		Acid	z
Valproic acid	0.172	0.378	**0.327**	0.346		0.419		Acid	x
Acetanilide	0.161	0.749			**0.425**		0.800	Base	z
Alprazolam	0.86	1.074	1.836		0.559		**0.971**	Base	x
Amitriptyline	10.643	13.682	**11.502**		15.366		14.461	Base	t
Amphetamine	3.91	2.920	3.560		**2.810**		5.364	Base	t
Bupivacaine	1	4.146	3.997		**3.601**		3.632	Base	x
Bupropion	16.1	2.665	2.408		**4.067**		3.297	Base	t
Butorphanol	5	3.416	**4.627**		3.305		6.872	Base	t
Chlordiazepoxide	0.363	1.077	0.920		0.835		**0.611**	Base	z
Chloroquine	112.4	9.210	10.597		**40.764**		29.384	Base	b
Chlorphentermine	2.5	**4.587**			5.359		7.211	Base	x
Chlortetracycline	1.61	**1.769**	1.889		1.429		1.430	Base	z
Clindamycin	1	0.404	0.479		**0.731**		0.375	Base	z
Clomipramine	17	28.628	29.235		**24.948**		28.838	Base	x
Desipramine	30.25	9.637	8.029		12.979		**14.624**	Base	z
Diazepam	1.92	**1.461**	1.335		3.075		1.352	Base	t
Doxepine	12.7	**5.933**	5.199		4.602		4.248	Base	b
Ethambutol	1.335	**1.754**	3.196		0.505		0.438	Base	z
Etidocaine	2	2.381	3.010		2.437		**2.139**	Base	b
Fentanyl	3.6	**3.705**	4.662		11.918		5.133	Base	z
Flunitrazepam	4	1.336	1.386		**5.181**		1.693	Base	x
Fluoxetine	44.5	186.16	183.01		**33.689**		17.218	Base	t
Gentamycin	0.28	0.213	0.696		**0.295**		0.414	Base	b
Griseofulvin	1.73	1.215	**1.515**					Base	b
Haloperidol	15.84	4.182	**5.562**		6.109		2.121	Base	z
Imipramine	16.5	11.797	10.156		11.047		**12.013**	Base	b
Kanamycin	0.19	0.095	0.483		0.100		**0.150**	Base	b
Ketamine	4	1.175	**2.044**		0.894		0.884	Base	z
Lidocaine	1.408	2.238	2.661		**1.675**		1.730	Base	b
Lincomycin	0.33	0.607	0.435		**0.169**		0.229	Base	x
Maprotiline	21	13.928	12.400		16.508		**16.958**	Base	b
Meptazinol	5.47	**5.220**	5.946		4.344		4.776	Base	b
Methadone	3.41	7.121	6.982		10.779		**5.678**	Base	b
Methaqualone	6	0.873	1.064		1.295		**1.740**	Base	z
Metoclopramide	3	1.305	**2.451**		1.986		2.055	Base	z
Midazolam	1.3	3.822	4.645		2.199		**1.956**	Base	b
Moclobemide	1.29	1.905	2.027		**1.241**		0.872	Base	t
Morphine	2.855	1.934	**2.192**		0.851		1.126	Base	t
Naloxone	3	1.641	1.926		3.091		**3.393**	Base	t
Neomycin	0.009							Base	b
Nortriptyline	20.5	18.724	**20.018**		13.405		22.654	Base	t
Paroxetine	23.25	13.489	9.191		40.325		**30.541**	Base	x
Pethidine	5.03	3.368	**3.979**		2.090		2.282	Base	b
Phenacetin	1.31	0.820	1.141		0.982		**1.449**	Base	x
Phenazone	0.6	**0.852**	0.903		1.767		1.367	Base	b
Phencyclidine	6	12.173	**12.310**		8.674		12.183	Base	x
Prazepam	1.5	4.193	6.195		**1.874**		4.116	Base	x
Protriptyline	14.763	**20.107**	19.830		29.284		39.200	Base	x
Pyrimethamine	2.19	2.875	3.840		1.242		**2.388**	Base	x
Quinine sulfate	1.63	4.562	**4.536**		3.863		5.145	Base	t
Rolietracycline	0.58	**0.549**	0.585					Base	b
Sertraline	20	17.744	11.756		24.849		**18.324**	Base	z
Spectinomycin	0.12	0.924			**0.129**		0.822	Base	x
Sulfisoxazole acetyl	1.19	0.644	0.530		2.265		**0.743**	Base	t
Temazepam	0.975	0.846	0.789		1.287		**1.017**	Base	z
Tinidazole	0.39	**0.691**	–		0.204		1.457	Base	b

Table 3 (*Continued*)

Drug	$V_{d_{obs}}$	$V_{d_{pred}}$						Acid or base	Test set code
		(8)	(9)	(10)	(11)	(12)	(13)		
Tramadol	3	3.736	5.250		3.200		**2.977**	Base	x
Triazolam	1.1	2.812	4.312		**1.603**		1.639	Base	b
Trimethoprim	2.4	1.053	0.887		**1.783**		1.321	Base	t
Viloxazine	1	3.499	2.024		1.429		**1.303**	Base	t

Table 4
Results of V_d prediction for drugs that were not included in the original analyses together with the observed V_d from literature; values in the parentheses are fold error of prediction

Drug	$V_{d_{obs}}$	pK_a	$V_{d_{pred}}$						Acid or base
			(8)	(9)	(10)	(11)	(12)	(13)	
Flucytosine	0.68[a]	10.87	0.568 (1.20)	1.396 (2.05)	**0.801** (1.18)		1.022 (1.50)		Acid
Nalidixic acid	0.51[b]	3.45	0.247 (2.06)		**0.317** (1.61)		0.261 (1.95)		Acid
Methenamine	0.56[c]	6.30	1.058 (1.89)			**0.348** (1.61)		1.641 (2.93)	Base
Phenobarbital	0.88[d]	7.36	0.661 (1.33)	0.621 (1.42)	1.091 (1.24)		**1.060** (1.20)		Acid
Prednisone	0.97[d]	12.39	**1.460** (1.51)		7.246 (7.47)		6.554 (6.76)		Acid
Dexamethasone	1.14[d]	12.14	**2.107** (1.85)		7.532 (6.61)		7.004 (6.14)		Acid
Hydrocortisone	0.44[d]	12.48	**0.839** (1.91)		7.116 (16.2)		6.731 (15.3)		Acid
Prednisolone	0.52[d]	12.47	**1.206** (2.32)	1.362 (2.62)	7.171 (13.8)		6.495 (12.5)		Acid
Salicylic acid	0.17[d]	2.98	0.260 (1.53)		0.262 (1.54)		**0.239** (1.41)		Acid
Theophylline	0.57[d]	8.54	**0.687** (1.21)	0.769 (1.35)	0.235 (2.42)		0.160 (3.56)		Acid

[a] V_d value from Cutler et al. (1978).
[b] Average of the V_d values reported in Barbeau and Belanger (1982) for nalidixic acid.
[c] V_d value from Klinge et al. (1982).
[d] Values reported in Lombardo et al. (2004).

The mean fold error of prediction values obtained from internal (Table 2) and external validation (Table 4) of QSPR models (8) and (9) for all drugs, and model (11) for bases are within the range of prediction errors using the interspecies scaling (reported to be in the range 1.56–2.78 folds (Obach et al., 1997)). For another prediction method based on the prediction of the fraction unbound in tissues using two experimentally determined physicochemical parameters, the fraction of compound ionized at pH 7.4, and the fraction of free drug in plasma, Lombardo et al. (2004) reported a training set MFE of 2.08 for V_d. The model worked well in the leave-class-out prediction with the MFE for classes being between 1.26 and 2.51. In comparison, the performance of model (8) is particularly good as it does not require any experimental measurement.

In the present QSPR approach we decided to use a simple multiple regression approach. Although increasing the complexity of the model, e.g. the use of non-linear techniques and/or a higher number of the descriptors, might improve the statistics of the models, the improvement will be associated with either complication or loss of the interpretability. The structural descriptors used in the QSPR models can be explained based on the previous understanding of the mechanisms involved and the chemical structures of the drugs. These also agree with our previous QSPR findings as follows. Models (8)–(11) all contain lipophilicity terms that are required for membrane transport. These are $(\log P)^2$ and lipole in Eq. (8) and (9), $\log D_{7.4}$ in Eq. (10) and $^1\kappa_a$ in Eq. (11). The other important feature of the models is the incorporation of ionized fractions of acids and bases at physiologic pH value. This is consistent in all the models and indicates that base ionization to cations increases the V_d while acid dissociation reduces the V_d. This can be deduced from the positive coefficients of fiB in Eqs. (8) and (9), the negative coefficient of $\log(fu/pK_a)$ in Eq. (11), and the negative coefficients of fiA in Eqs. (8)–(10). This could be due partly to the high protein binding of most acidic drugs. Karalis et al. (2002) observed that in a class of compounds with higher volume of distribution, the acid/base ratio was lower, whereas protein-binding extent was highest in the class with the lowest volume of distribution.

Hydrogen bonding parameters are also present in these models. The need for hydrogen bonding parameter in order to model membrane transport is very well established both as a correction to lipophilicity-based models (Feher et al., 2000) and in solvatochromic approaches (Abraham et al., 1999). As explained in the results section, other structural parameters present in the QSPR models point to more precise molecular features such as shape of the molecules and presence of specific functional groups.

5. Conclusion

Apparent volume of distribution for drug entities belonging to different chemical classes was studied using a QSPR approach. Some of the suggested QSPR models resulted in encouragingly low prediction errors. The errors were within the range of more complicated prediction methods such as interspecies scaling and a method requiring experimentally determined parameters, e.g. extent of plasma protein binding. Furthermore, the structural descriptors used in the models can be interpreted based on the

mechanisms involved in distribution process and the molecular structures of the drugs. These are fractions ionized to anions or cations, partition (or apparent distribution) coefficient, hydrogen binding parameters, and some atom type electrotopological state indexes indicating the presence of certain atomic groups in the molecular structures. The model could find use in novel drug design and high throughput screening laboratories.

Acknowledgments

We are grateful to the Research Council of Tabriz Medical Sciences University for financial support of parts of this study. T.G. thanks Dr. Mark Cronin from Liverpool John Moores University for providing some of the descriptors used in the study.

Appendix A. Correlation matrix for descriptors of Eqs. (8)–(13)

Descriptors of Eqs. (8) and (9) for all the drugs:

	fiA	fiB	SsssCH	SsF	ESP^-	$(\log P)^2$	Lipole
fiB	−0.265						
SsssCH	−0.164	−0.252					
SsF	−0.156	0.116	0.075				
ESP^-	−0.214	0.088	0.184	0.022			
$(\log P)^2$	−0.348	0.428	0.219	0.143	0.244		
Lipole	0.197	0.354	−0.215	0.102	−0.143	0.181	
ppb	−0.002	−0.09	0.301	0.138	−0.002	0.423	0.082

Descriptors of Eqs. (10) and (12) for acids:

	$N_{6\text{-ring}}$	fiA	$^7\chi_{ch}$	SsssN	Pol	SaaN
fiA	0.212					
$^7\chi_{ch}$	−0.277	0.004				
SsssN	0.327	0.414	0.209			
$\log D_{7.4}$	−0.583	−0.568	−0.041	−0.462		
Pol	0.349	0.465	0.070	0.442		
SaaN	−0.176	0.171	−0.162	−0.189	0.033	
MV/TA	−0.123	0.162	0.256	0.042	0.205	−0.128

Descriptors of Eqs. (11) and (13) for bases:

	$\log(fu/pK_a)$	ESP^-	Dipole	N_{OH}	ABSQon	$^1\kappa_a$
ESP^-	−0.229					
Dipole	0.156	−0.195				
N_{OH}	−0.287	−0.228	0.309			
ABSQon	−0.12	−0.308	0.498	0.856		
$^1\kappa_a$	−0.31	−0.422	0.343	0.723	0.729	
$\log P$	−0.157	0.243	−0.379	−0.632	−0.771	−0.334

Appendix B. Equations for training sets

Structural descriptors from Eq. (8):

Test set *b* excluded :
$$\log V_d = -1.85 - 0.497\text{fiA} + 0.480\ \text{fiB} + 0.528(\log P)^2 + 1.26\text{SsssCH} + 0.724\text{Lipole} + 0.744\text{SsF} + 0.660\text{ESP}^-, \quad n = 110,\ s = 0.310,\ r^2 = 77.0\%,\ F = 48.9$$

Test set *t* excluded :
$$\log V_d = -1.76 - 0.488\ \text{fiA} + 0.401\ \text{fiB} + 0.592(\log P)^2 + 1.23\text{SsssCH} + 0.845\text{Lipole} + 0.713\text{SsF} + 0.550\text{ESP}^-, \quad n = 107,\ s = 0.322,\ r^2 = 0.750,\ F = 42.4$$

Test set *x* excluded :
$$\log V_d = -1.61 - 0.479\ \text{fiA} + 0.514\ \text{fiB} + 0.639(\log P)^2 + 1.17\text{SsssCH} + 0.662\text{Lipole} + 0.559\text{SsF} + 0.397\text{ESP}^-, \quad n = 111,\ s = 0.314,\ r^2 = 0.740,\ F = 42.0$$

Test set z excluded : $\log V_d = -1.88 - 0.447\text{fiA} + 0.427\text{fiB} + 0.523(\log P)^2 + 1.26\text{SsssCH} + 0.779\text{Lipole} + 0.666\text{SsF} + 0.668\text{ESP}^-, \quad n = 112,\ s = 0.301,\ r^2 = 0.757,\ F = 46.3$

Structural descriptors from Eq. (9):

Test set b excluded : $\log V_d = -1.36 - 0.528\text{fiA} + 0.477\text{fiB} + 0.674(\log P)^2 + 1.34\text{SsssCH} + 0.640\text{Lipole} + 0.860\text{SsF} - 0.00227\text{ppb}, \quad n = 104,\ s = 0.301,\ r^2 = 0.786,\ F = 50.4$

Test set t excluded : $\log V_d = -1.32 - 0.536\text{fiA} + 0.376\text{fiB} + 0.758(\log P)^2 + 1.33\text{SsssCH} + 0.835\text{Lipole} + 0.821\text{SsF} - 0.00302\text{ppb}, \quad n = 101,\ s = 0.307,\ r^2 = 0.774,\ F = 45.5$

Test set x excluded : $\log V_d = -1.33 - 0.489\text{fiA} + 0.439\text{fiB} + 0.873(\log P)^2 + 1.34\text{SsssCH} + 0.669\text{Lipole} + 0.644\text{SsF} - 0.00306\text{ppb}, \quad n = 107,\ s = 0.307,\ r^2 = 0.758,\ F = 44.2$

Test set z excluded : $\log V_d = -1.41 - 0.457\text{fiA} + 0.437\text{fiB} + 0.725(\log P)^2 + 1.41\text{SsssCH} + 0.689\text{Lipole} + 0.833\text{SsF} - 0.00322\text{ppb}, \quad n = 107,\ s = 0.295,\ r^2 = 0.771,\ F = 47.6$

Structural descriptors from Eq. (10) for acids:

Test set b excluded : $\log V_d = -0.454 + 0.951N_{\text{6-rings}} - 0.405\text{fiA} + 0.551^7\chi_{\text{ch}} - 0.980\text{SsssN} + 0.485\ \log D_{7.4}, \quad n = 54,\ s = 0.247,\ r^2 = 0.612,\ F = 15.1$

Test set t excluded : $\log V_d = -0.625 + 0.995N_{\text{6-rings}} - 0.329\text{fiA} + 0.580^7\chi_{\text{ch}} - 0.946\text{SsssN} + 0.725\ \log D_{7.4}, \quad n = 51,\ s = 0.281,\ r^2 = 0.568,\ F = 11.9$

Test set x excluded : $\log V_d = -0.655 + 1.04N_{\text{6-rings}} - 0.207\text{fiA} + 0.485^7\chi_{\text{ch}} - 1.58\text{SsssN} + 0.698\ \log D_{7.4}, \quad n = 51,\ s = 0.273,\ r^2 = 0.548,\ F = 10.9$

Test set z excluded : $\log V_d = -0.810 + 1.02N_{\text{6-rings}} - 0.228\text{fiA} + 0.615^7\chi_{\text{ch}} - 1.12\text{SsssN} + 0.894\ \log D_{7.4}, \quad n = 51,\ s = 0.278,\ r^2 = 0.519,\ F = 9.7$

Structural descriptors from Eq. (11) for bases:

Test set b excluded : $\log V_d = 1.19 - 2.41\ \log\left(\frac{\text{fu}}{\text{p}K_a}\right) + 1.05\text{ESP}^- + 3.07\text{Dipole} - 0.990N_{\text{OH}} - 2.65\text{ABSQon} + 1.97^1\kappa_a, \quad n = 40,\ s = 0.297,\ r^2 = 0.799,\ F = 21.9$

Test set t excluded : $\log V_d = 1.30 - 2.73\ \log\left(\frac{\text{fu}}{\text{p}K_a}\right) + 1.20\text{ESP}^- + 2.13\text{Dipole} - 1063N_{\text{OH}} - 1.75\text{ABSQon} + 2.02^1\kappa_a, \quad n = 42,\ s = 0.300,\ r^2 = 0.837,\ F = 30.0$

Test set x excluded : $\log V_d = 1.67 - 3.05\ \log\left(\frac{\text{fu}}{\text{p}K_a}\right) + 1.15\text{ESP}^- + 2.98\text{Dipole} - 1.40N_{\text{OH}} - 2.12\text{ABSQon} + 1.87^1\kappa_a, \quad n = 41,\ s = 0.312,\ r^2 = 0.809,\ F = 24.0$

Test set z excluded : $\log V_d = 1.03 - 2.55\ \log\left(\frac{\text{fu}}{\text{p}K_a}\right) + 1.36\text{ESP}^- + 3.12\text{Dipole} - 1.25N_{\text{OH}} - 2.20\text{ABSQon} + 2.00^1\kappa_a, \quad n = 42,\ s = 0.261,\ r^2 = 0.864,\ F = 37.0$

Structural descriptors from Eq. (12) for acids:

$$\text{Test set } b \text{ excluded}: \quad \log V_d = -0.519 + 0.634N_{\text{6-rings}} - 0.567\text{fiA} - 1.19\text{SsssN} + 0.491\text{Pol} - 0.341\text{SaaN} + 0.549\frac{\text{MV}}{\text{TA}},$$
$$n = 53,\ s = 0.245,\ r^2 = 0.633,\ F = 13.2$$

$$\text{Test set } t \text{ excluded}: \quad \log V_d = -0.495 + 0.504N_{\text{6-rings}} - 0.568\text{fiA} - 1.33\text{SsssN} + 0.779\text{Pol} - 0.627\text{SaaN} + 0.395\frac{\text{MV}}{\text{TA}},$$
$$n = 50,\ s = 0.279,\ r^2 = 0.591,\ F = 10.4$$

$$\text{Test set } x \text{ excluded}: \quad \log V_d = -0.476 + 0.657N_{\text{6-rings}} - 0.356\text{fiA} - 1.96\text{SsssN} + 0.348\text{Pol} - 0.554\text{SaaN} + 0.548\frac{\text{MV}}{\text{TA}},$$
$$n = 51,\ s = 0.262,\ r^2 = 0.594,\ F = 10.7$$

$$\text{Test set } z \text{ excluded}: \quad \log V_d = -0.662 + 0.579N_{\text{6-rings}} - 0.410\text{fiA} - 1.55\text{SsssN} + 0.582\text{Pol} - 0.338\text{SaaN} + 0.579\frac{\text{MV}}{\text{TA}},$$
$$n = 50,\ s = 0.280,\ r^2 = 0.534,\ F = 8.2$$

Structural descriptors from Eq. (13) for bases:

$$\text{Test set } b \text{ excluded}: \quad \log V_d = 1.20 - 2.53\log\left(\frac{\text{fu}}{\text{p}K_a}\right) + 0.689\text{ESP}^- + 2.64\text{Dipole} - 1.17N_{\text{OH}} + 1.14\log P,$$
$$n = 40,\ s = 0.357,\ r^2 = 0.701,\ F = 15.9$$

$$\text{Test set } t \text{ excluded}: \quad \log V_d = 1.31 - 2.88\log\left(\frac{\text{fu}}{\text{p}K_a}\right) + 0.716\text{ESP}^- + 2.42\text{Dipole} - 1.03N_{\text{OH}} + 1.35\log P,$$
$$n = 42,\ s = 0.347,\ r^2 = 0.775,\ F = 24.8$$

$$\text{Test set } x \text{ excluded}: \quad \log V_d = 1.31 - 2.90\log\left(\frac{\text{fu}}{\text{p}K_a}\right) + 0.690\text{ESP}^- + 2.25\text{Dipole} - 0.967N_{\text{OH}} + 1.52\log P,$$
$$n = 41,\ s = 0.329,\ r^2 = 0.782,\ F = 25.1$$

$$\text{Test set } z \text{ excluded}: \quad \log V_d = 0.730 - 2.50\log\left(\frac{\text{fu}}{\text{p}K_a}\right) + 1.10\text{ESP}^- + 2.87\text{Dipole} - 1.08N_{\text{OH}} + 1.33\log P,$$
$$n = 42,\ s = 0.305,\ r^2 = 0.809,\ F = 30.5$$

References

Abraham, M.H., Chadha, H.S., Martins, F., Mitchell, R.C., Bradbury, M.W., Gratton, J.A., 1999. Hydrogen bonding—Part 46. A review of the correlation and prediction of transport properties by an LFER method: physicochemical properties, brain penetration and skin permeability. Pestic. Sci. 55, 78–88.

Barbeau, G., Belanger, P.M., 1982. Pharmacokinetics of nalidixic acid in old and young volunteers. J. Clin. Pharmacol. 22, 490–496.

Blakey, G.E., Nestorov, I.A., Arundel, P.A., Aarons, L.J., Rowland, M., 1997. Quantitative structure–pharmacokinetics relationships: I. Development of a whole body physiologically based model to characterize changes in pharmacokinetics across a homologous series of barbiturates in the rat. J. Pharmacokinet. Biopharm. 25, 277–312.

Cutler, R.E., Blair, A.D., Kelly, M.R., 1978. Flucytosine kinetics in subjects with normal and impaired renal function. Clin. Pharmacol. Ther. 24, 333–342.

Davis, A.M., Webborn, P.J.H., Salt, D.W., 2000. Robust assessment of statistical significance in the use of unbound/intrinsic pharmacokinetic parameters in quantitative structure–pharmacokinetic relationships with lipophilicity. Drug Metab. Dispos. 28, 103–106.

Dearden, J.C., Ghafourian, T., 1999. Hydrogen bonding parameters for QSAR: comparison of indicator variables, counts, molecular orbital and other parameters. J. Chem. Inf. Comp. Sci. 39, 231–235.

Durnas, C., Mingloi, C., Cusack, B., 1990. Hepatic drug metabolism and aging. Clin. Pharmacokinet. 19, 359–389.

Feher, M., Sourial, E., Schmidt, J.M., 2000. A simple model for the prediction of blood-brain partitioning. Int. J. Pharm. 201, 239–247.

Fulton, B., Sorkin, E.M., 1995. Propofol an overview of its pharmacology and a review of its clinical efficacy in intensive care sedation. Drugs 50, 587–767.

Ghafourian, T., Barzegar-Jalali, M., Hakimiha, N., Cronin, M.T.D., 2004. Quantitative structure-pharmacokinetics relationships study of the apparent volume of distribution of CNS drugs. J. Pharm. Pharmacol. 56, 339–350.

Glare, P.A., Walsh, T.D., 1991. Clinical pharmacokinetics of morphine. Ther. Drug Monit. 13, 1–23.

Gobburu, J.V.S., Shelver, W.H., 1995. Quantitative structure pharmacokinetic relationships (QSPR) of beta blockers derived using neural networks. J. Pharm. Sci. 84, 862–865.

Greenblatt, D.J., 1981. Clinical pharmacokinetics of oxazepam and lorazepam. Clin. Pharmacokinet. 6, 89–105.

Hirono, S., Nakagome, I., Hirano, H., Yoshii, F., Moriguchi, I., 1994. Non-congeneric structure–pharmacokinetic property correlation studies using fuzzy adaptive least squares: volume of distribution. Biol. Pharm. Bull. 17, 686–690.

Karalis, V., Claret, L., Iliadis, A., Macheras, P., 2001. Fractal volume of drug distribution: it scales proportionally to body mass. Pharm. Res. 18, 1056–1060.

Karalis, V., Tsantili-Kakoulidou, A., Macheras, P., 2002. Multivariate statistics of disposition pharmacokinetic parameters for structurally unrelated drugs used in therapeutics. Pharm. Res. 19, 1827–1834.

Klinge, E., Mannisto, P., Mantyla, R., 1982. Pharmacokinetics of methenamine in healthy volunteers. J. Antimicrob. Chemother. 9, 209–216.

Lam, Y.W., Banerji, S., Hatfield, C., Talbert, L., 1997. Principle of drug administration in renal insufficiency. Clin. Pharmacokinet. 32, 30–57.

Lombardo, F., Obach, R.S., Shalaeva, M.Y., Gao, F., 2004. Prediction of human volume of distribution values for neutral and basic drugs. 2: Extended data set and leave-class-out statistics. J. Med. Chem. 47, 1242–1250.

Mahmood, I., 1998. Interspecies scaling: predicting volumes, mean residence time and elimination half-life. Some suggestions. J. Pharm. Pharmacol. 50, 493–499.

Moffat, A.C., Jackson, J.V., Moss, M.S., Winddop, B., 1986. Clarke's Isolation of Drug in Pharmaceuticals, Body Fluids and Post-Mortem Material, second ed. The Pharmaceutical Press, London.

Nattell, S., Gagne, G., Pineau, M., 1987. The pharmacokinetics of lidocaine and β-adrenoreceptor antagonists in patients with acute myocardial infarction. Clin. Pharmacokinet. 13, 293–316.

Obach, R.S., Baxter, J.G., Liston, T.E., Silber, B.M., Jones, B.C., MacIntyre, F., Rance, D.J., Wastall, P., 1997. The prediction of human pharmacokinetic parameters from preclinical and in vitro metabolism data. J. Pharm. Exp. Ther. 283, 48–58.

Perry, P.J., 2002. Psychotropic agents. In: Schoenwald, R.D. (Ed.), Pharmacokinetics in Drug Discovery and Development. CRC Press, Boca Raton, FL, pp. 175–197.

Raaflaub, V.J., Speiser-Courvoisier, J., 1974. Zur pharmakokinetik von bromazepam beim menschen. Arzneimittelforschung 24, 1841–1844.

Ritschel, W.A., 1976. Handbook of Basic Pharmacokinetics. Drug Inteligence Publication, Hamilton.

Ritschel, W.A., Hammer, G.V., 1980. Prediction of the volume of distribution from in vitro data and use for estimating the absolute extent of absorption. Int. J. Clin. Pharmacol. Ther. Toxicol. 18, 298–316.

Ritschel, W.A., Akileswaran, R., Hussain, A.S., 1995. Application of neural networks for the prediction of human pharmacokinetic parameters. Methods Find. Exp. Clin. Pharmacol. 17, 629–643.

Rowland, M., Tozer, T.N., 1995. Clinical Pharmacokinetics. Concepts and Applications, third ed. Lippincott, Williams and Wilkins, Philadelphia.

Schoerlin, M., Horber, F.F., Ferry, F.J., Mayersohn, M., 1990. Disposition kinetics of moclobemide, a new MAO-A inhibitor in subjects with impaired renal function. J. Clin. Pharmacol. 30, 272–284.

Shargel, L., Yu, A.B.C., 1999. Applied Biopharmaceutics and Pharmacokinetics. Appleton & Lange, Stamford, pp. 267–269.

Smith, D.A., Van de Waterbeemd, H., Walker, D.K., 2001. Pharmacokinetics and metabolism in drug design. Wiley-VCH, Weinheim, Germany, pp. 123–132.

Sonne, J., Loft, S., Dosing, M., Vollmerlarsen, A., Olesen, K.L., Victor, M., Andreasen, F., Andreasen, P.B., 1988. Bioavailability and pharmacokinetics of oxazepam. Eur. J. Clin. Pharmacol. 35, 385–389.

Turner, J.V., Maddalena, D.J., Cutler, D.J., Agatonovic-Kustrin, S., 2003. Multiple pharmacokinetic parameter prediction for a series of cephalosporins. J. Pharm. Sci. 92, 552–559.

Urien, S., Tillement, J.-P., Barré, J., 2001. The significance of plasma-protein binding in drug research. In: Testa, B., van de Waterbeemd, H., Folkers, G., Guy, R. (Eds.), Pharmacokinetic Optimization in Drug Research. Wiley-VCH, Weinheim, pp. 189–198.

Van de Waterbeemd, H., 2005. Which in vitro screens guide the prediction of oral absorption and volume of distribution? Basic Clin. Pharmacol. Toxicol. 96, 162–166.

Wilkinson, G.R., 2001. Pharmacokinetics: the dynamics of drug absorption, distribution, and elimination. In: Hardman, J.G., Limbird, L.E., Gillman, A.G. (Eds.), Goodman & Gillman's The Pharmacological Basis of Therapeutics, 10th ed. McGraw-Hill, New York, pp. 20–22.