Available online at www.sciencedirect.com
SCIENCE DIRECT®
Bioorganic & Medicinal Chemistry 12 (2004) 4735–4748
Bioorganic & Medicinal Chemistry

# QSPR treatment of rat blood:air, saline:air and olive oil: air partition coefficients using theoretical molecular descriptors

Alan R. Katritzky,[a,*] Minati Kuanar,[a] Dan C. Fara,[a,b] Mati Karelson[b] and William E. Acree, Jr.[c]

[a]*Center for Heterocyclic Compounds, Department of Chemistry, University of Florida, Gainesville, FL 32611-17200, USA*
[b]*Department of Chemistry, University of Tartu, 2 Jakobi Street, Tartu 51014, Estonia*
[c]*Department of Chemistry, University of North Texas, Denton, TX 76203-5070, USA*

Received 29 March 2004; revised 13 May 2004; accepted 25 May 2004
Available online 20 July 2004

**Abstract**—A QSPR treatment has been applied to a data set that consists of 100 diverse organic compounds to relate the logarithmic function of rat blood:air, saline:air and olive oil:air partition coefficients (denoted by $\log K(b{:}a)$, $\log K(s{:}a)$, and $\log K(o{:}a)$, respectively) with theoretical molecular and fragment descriptors. Three QSPR models with squared correlation coefficients of 0.881, 0.926, and 0.922, respectively, were obtained. The verification of the predictive power of these models on a test set of 33 organic chemicals that were not included in the training set gave satisfactory squared correlation coefficients: 0.791 for rat blood:air, 0.794 for saline:air and 0.846 for olive oil:air.


## 1. Introduction

Physiologically based pharmacokinetic (PBPK) models to describe the absorption, distribution, and elimination in animals and humans of volatile organic compounds (VOC) make frequent use of blood:air, saline:air, olive oil:air, and tissue/blood partition coefficients to derive metabolic rate constants, and allometric equations.

The partition coefficient, $PC_{A/B}$, for a given organic compound is defined as the ratio of concentrations achieved at equilibrium between the two different media as expressed mathematically in Eq. 1, where A can be blood, saline, oil etc., and B is air.

$$PC_{A/B} = \frac{\text{concentration in media A}}{\text{concentration in media B at equilibrium}} \quad (1)$$

The solubility of a volatile organic chemical in blood, indicated as the blood:air partition coefficient, is one of the most important physicochemical properties for understanding the pharmacokinetics of organic solvents. The solubility in blood is measured as the sum total of the solubility in the mixture of neutral lipids, polar lipids, and water constituting the blood. Solubilities in *n*-octanol were initially used as a surrogate of the solubility of organic chemicals in neutral lipids. However, *n*-octanol, itself an alcohol, solubilizes hydrophilic organic chemicals (e.g., alcohols, ketones, and acetate esters) to a greater extent than biotic lipids, thus over-predicting the liposolubility of such hydrophilic chemicals.[1]

Abraham et al.[2] used the measured values of $\log L_{water}$ and $\log L_{oil}$ (where $L$ is the Ostwald solubility coefficient or gas–liquid partition coefficient) as descriptors to model the solubilities of a range of nonelectrolyte solutes in biological systems, and obtained good correlation coefficients, $R$, for blood ($R = 0.9951$), plasma ($R = 0.9982$), brain ($R = 0.9945$), lung ($R = 0.9993$), liver ($R = 0.9970$), kidney ($R = 0.9958$), muscle tissue ($R = 0.9877$), and human fat ($R = 0.9988$). Kamlet et al.[3] proposed a general solvation parameter model ($R = 0.9870$) to describe the gas–blood partition coefficients and molar solubilities in blood of 27 liquid aliphatic nonelectrolyte solutes.

*Keywords*: QSPR; Model; Rat blood:air; Saline:air; Olive oil:air; Partition coefficient; Molecular descriptors; CODESSA PRO.
* Corresponding author. Tel.: +1-352-392-0554; fax: +1-352-392-9199; e-mail: katritzky@chem.ufl.edu

0968-0896/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmc.2004.05.028

Paterson and Mackay[4] developed a simple approach for predicting tissue:air, tissue–blood and blood:air partition coefficients of 24 volatile organic compounds from water solubility and vapor pressure data. In 1994, Abraham and Weathersby[5] reexamined the ability to predict the log values of the solubilities for a range of 99 nonelectrolytes in biological phases by using a general solvation equation, as applied to gas-condensed phase processes. The correlation coefficients of the models obtained for solubilities vary from 0.9753 (for kidney) to 0.9985 (for oil).

The partitioning of VOCs between tissues and blood is also determined by the extent of binding of these chemicals to specific components such as plasma proteins and hemoglobin:[6] such physicochemical interactions need to be considered during the development of physicochemically-based methods for the estimation of various partition coefficients.

Model predictions of tissue doses are sensitive to changes in the values of the experimental partition coefficient parameters and the reliability of there values is highly dependent on the experimental accuracy of the analytical procedures used to obtain it. Consequently, reliable analytical methods are required for the experimental determination of partition coefficients.

The vial equilibration method of Gargas et al.[7] is currently used for determining the blood:air and tissue:air partition coefficients for VOCs. This method is a modified version of the vial equilibration technique proposed by Sato and Nakajima,[8] and has several advantages including: (i) elaborate experimental apparatus is not required and (ii) the partition coefficient can be evaluated by measuring headspace concentrations in test vials and appropriate reference vials, without measuring the chemical concentration in the test medium. Gargas et al.[7] underlined the need of determining experimental values of partition coefficients at a minimum of two time points, in order to detect metabolic or other active processes, which are indicated by an unusually large value for the partition coefficient or the absence of any measurable chemical concentration in the headspace.

In 1989, Gargas et al.[7] related rat blood:air partition coefficients to the corresponding saline:air and olive oil:air partition coefficients. These authors reported a bilinear regression with $R^2 = 0.928$ for 55 organic solutes. The significant squared correlation coefficient value reveals a good empirical relationship of blood:air partition coefficients with those of saline:air and olive oil:air partition coefficients.

Meulenberg and Vijverberg[9] examined the bilinear relationship of rat blood:air partition coefficients with saline:air and olive oil:air partition coefficients using 92 compounds ($R^2 = 0.93$). They also developed bilinear correlations for $\log P_{\text{rat tissue:air}}$, $P_{\text{human tissue:air}}$ using as input values $\log P_{\text{olive oil:air}}$ and $\log P_{\text{saline:air}}$ data. Meulenberg et al.[10] recently predicted the rat brain:air partition coefficients for 46 organic solvents (28 out of these 46 having newly measured values of the partition coefficients in vitro) using the previously reported bilinear regression equation.[9] The quality of the regression[10] ($R^2 = 0.92$) is comparable with their previous regression results for rat tissues.[9]

A mechanistic algorithm for predicting tissue:blood and blood:air partition coefficients of organic chemicals has also been formulated by Poulin and Krishnan.[6] Poulin and Krishnan have reported predictive approaches based on the sole consideration of chemical solubility in blood lipids and water;[11] these accurately predicted the rat blood:air PCs of relatively hydrophilic organic chemicals but discrepancies were found for lipophilic chemicals.

The investigation of the partitioning of VOCs into rat erythrocytes suggested that proteins, (chiefly hemoglobin) are the major carriers of some organic chemicals (such as *n*-hexane, toluene, chloroform, methyl *iso*-butyl ketone, diethyl ether) in blood.[12] Beliveau and Krishnan verified the adequacy of reconstituted mixtures of blood components (lipid, water, hemoglobin) for estimating the rat blood:air PCs of VOCs[13] and showed that VOC partitioning into three blood components, namely, water, lipids, and hemoglobin determines, to a large extent, the magnitude of their blood:air PCs. After investigating the concentration dependency of rat blood:air partition coefficients of some volatile organic chemicals, Beliveau and Krishnan suggested that the concentration-dependent nature of blood:air PC need not be considered for PBPK modeling of rat inhalation exposures for concentrations of up to several thousand parts per million of bromoform, chloroform, chlorobenzene, and ethyl benzene (BF, CF, CB, and EB).[14] The comparison of blood:air partition coefficient of lipophilic organic solvents with plasma triglyceride concentrations suggested that the blood triglyceride levels have no significant effect on the blood:air PC of organic solvents.[15]

The factors presented above suggest that QSPR methods could assist the estimation of the various partition coefficients (i.e. blood:air, saline:air, olive oil:air etc.). In particular, Basak et al.[16] used experimental (saline:air, olive oil:air partition coefficients) and theoretical molecular (topostructural, topochemical, 3-dimensional) descriptors to develop various models for the prediction of rat tissue:air partition coefficient for a data set consisting of 46 diverse low molecular-weight volatile chemicals. The authors applied the hierarchical QSAR (HiQSAR) method to assess the relative contribution of each descriptor class. Three regression methods (ridge regression, principal component regression and partial least squares) were compared. Basak et al.[16] concluded that HiQSAR based on theoretical molecular descriptors can be used to predict tissue:air partition coefficients of VOCs, and the theoretically based models are comparable or superior to those developed using experimental properties including saline:air and/or olive oil:air partition coefficients.

Our groups and others have developed QSPR correlation equations successfully for the prediction of aqueous

solubility, vapor pressure and water:air partition coefficients of structurally diverse compounds.[17,18] Many QSAR models were developed using molecular descriptors derived solely from structures.[19–21]

CODESSA PRO software provides a comprehensive set of descriptors calculated: (i) directly from the molecular structural formula (constitutional, topological); (ii) utilizing the molecular 3D geometry for descriptors such as geometrical and CPSA (charged partial surface area), or (iii) from the results of semiempirical quantum chemical (MOPAC) calculations—quantum chemical descriptors, which include orbital energies and coefficients (and their combinations), atomic and bond populations, various components of the MOPAC energy partitioning scheme, polarizabilities up to second order, dipole moments, and calculated thermodynamic functions.

Our principal goal in this study is the search for useful QSPR models for the correlation/prediction of rat blood:air, saline:air, and olive oil:air partition coefficients using theoretical molecular descriptors calculated with CODESSA PRO software. We also used saline:air and olive oil:air partition coefficients as external descriptors (independent variables) in order to predict rat blood:air partition coefficients.

## 2. Data Set

The partition coefficient data ($K$) for different solutes between rat blood and air has been measured experimentally by several investigators. The experimental values of rat blood:air partition coefficients for 133 organic chemical solutes were collected from many different literature sources.[6,7,13,15,22–60] After the examination of the published human and rat blood:air partition coefficient data from several sources, Abraham and Weathersby[5] assigned an uncertainty of about 0.1 logarithmic units to the published values. The human blood:air partition coefficient data considered by Abraham and Weathersby, and the rat blood:air partition coefficient data given in Table 1, were both measured by the vial equilibration method.

Saline:air and olive oil:air partition coefficient data ($K$) for the same set of organic compounds were also

**Table 1.** Data set of observed and predicted logarithmic function of blood:air (*b*:*a*), saline:air (*s*:*a*), and olive oil:air (*o*:*a*) partition coefficients $K$

| # | Compound name | Observed | | | Predicted | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | $\log K(b{:}a)$ | $\log K(s{:}a)$ | $\log K(o{:}a)$ | $\log K(b{:}a)$ by | | | | | $\log K(s{:}a)$ by | $\log K(o{:}a)$ by |
| | | | | | Eq. 2 | Eq. 11 | Eq. 14 | Eq. 15 | Eq. 16 | Eq. 5 | Eq. 6 |
| *Training set* | | | | | | | | | | | |
| 1 | Ethane | −0.98 | −1.60 | 0.21 | −0.58 | −0.96 | −0.61 | −0.73 | −0.49 | −0.65 | 0.23 |
| 2 | Butane | −0.53 | −1.75 | 1.29 | −0.13 | −0.57 | −0.14 | −0.37 | −0.11 | −0.90 | 1.39 |
| 3 | Hexane | 0.23 | −1.74 | 2.18 | 0.34 | −0.18 | 0.07 | 0.08 | 0.25 | −1.38 | 2.29 |
| 4 | Heptane | 0.68 | −1.03 | 2.64 | 0.65 | 0.43 | 0.14 | 0.55 | 0.43 | −1.62 | 2.69 |
| 5 | Octane | 0.88 | −2.23 | 3.16 | 0.86 | −0.03 | 0.20 | 0.36 | 0.61 | −1.87 | 3.07 |
| 6 | 2,3,4-Trimethylpentane | 0.57 | −2.00 | 2.65 | 0.97 | −0.12 | 0.08 | 0.43 | 0.60 | −1.84 | 2.81 |
| 7 | 2,2,4-Trimethylpentane | 0.69 | −2.00 | 2.47 | 0.93 | −0.20 | 0.03 | 0.43 | 0.60 | −1.85 | 2.72 |
| 8 | 2-Methyl-1,3-butadiene | 0.27 | −0.68 | 0.94 | 0.68 | −0.11 | 0.57 | 0.37 | 0.52 | −0.18 | 1.96 |
| 9 | 1,4-Pentadiene | 0.07 | −0.72 | 1.73 | 0.74 | 0.21 | 0.38 | 0.36 | 0.40 | −0.61 | 2.06 |
| 10 | *cis*-Piperylene | 0.43 | −0.52 | 1.99 | 0.64 | 0.44 | 0.65 | 0.43 | 0.54 | −0.13 | 2.06 |
| 11 | *trans*-Piperylene | 0.35 | −0.54 | 1.94 | 0.58 | 0.41 | 0.65 | 0.43 | 0.54 | 0.14 | 2.06 |
| 12 | 2,4-*trans*,*trans*-Hexadiene | 0.78 | −0.57 | 2.42 | 0.69 | 0.60 | 0.85 | 0.64 | 0.77 | −0.20 | 2.50 |
| 13 | 1,3,5-*trans*,*trans*,*trans*-Heptatriene | 0.74 | −0.24 | 3.23 | 1.12 | 1.15 | 1.01 | 0.97 | 0.95 | −0.40 | 3.02 |
| 14 | 1-Hexyne | 0.73 | −0.27 | 2.41 | 1.01 | 0.77 | 0.65 | 0.74 | 0.68 | −0.52 | 2.46 |
| 15 | 3-Hexyne | 0.96 | −0.02 | 2.48 | 0.63 | 0.94 | 0.73 | 0.85 | 0.74 | −0.37 | 2.46 |
| 16 | Cyclopentane | 0.24 | −1.06 | 2.14 | 0.25 | 0.20 | 0.27 | 0.22 | 0.34 | −0.69 | 1.94 |
| 17 | Cyclohexane | 0.18 | −0.96 | 2.49 | 0.42 | 0.41 | 0.37 | 0.48 | 0.52 | −0.93 | 2.37 |
| 18 | Benzene | 1.15 | 0.45 | 2.71 | 1.14 | 1.31 | 1.50 | 1.07 | 1.08 | 0.65 | 2.81 |
| 19 | Toluene | 1.20 | 0.30 | 3.21 | 1.29 | 1.45 | 1.56 | 1.31 | 1.35 | 0.43 | 3.15 |
| 20 | 1,2-Dimethylbenzene | 1.63 | 0.47 | 3.86 | 1.47 | 1.83 | 1.57 | 1.61 | 1.54 | 0.16 | 3.48 |
| 21 | 1,3-Dimethylbenzene | 1.64 | 0.28 | 3.74 | 1.41 | 1.67 | 1.56 | 1.54 | 1.54 | 0.14 | 3.47 |
| 22 | Ethylbenzene | 1.64 | 0.27 | 3.74 | 1.43 | 1.66 | 1.62 | 1.53 | 1.53 | 0.16 | 3.57 |
| 23 | Vinylbenzene | 1.60 | 0.15 | 3.55 | 1.71 | 1.51 | 1.71 | 1.42 | 1.45 | 0.21 | 3.68 |
| 24 | 1-Vinyl-3-methylbenzene | 2.28 | 0.29 | 4.17 | 1.85 | 1.86 | 1.72 | 1.75 | 1.71 | −0.04 | 3.97 |
| 25 | 1-Vinyl-4-methylbenzene | 2.37 | 0.32 | 4.14 | 1.86 | 1.87 | 1.75 | 1.76 | 1.72 | 0.01 | 3.97 |
| 26 | Chloromethane | 0.39 | −0.06 | 0.93 | 0.41 | 0.24 | 0.35 | 0.37 | 0.49 | 0.36 | 0.97 |
| 27 | Dichloromethane | 1.34 | 0.78 | 2.12 | 0.97 | 1.24 | 0.91 | 1.12 | 1.05 | 0.05 | 1.87 |
| 28 | Chloroform | 1.30 | 0.53 | 2.60 | 1.35 | 1.31 | 1.23 | 1.10 | 1.07 | 0.48 | 2.49 |
| 29 | Chloroethane | 0.61 | 0.04 | 1.59 | 0.41 | 0.59 | 0.49 | 0.40 | 0.39 | 0.18 | 1.42 |
| 30 | 1,1-Dichloroethane | 1.08 | 0.39 | 2.27 | 1.00 | 1.09 | 0.88 | 0.96 | 0.90 | 0.19 | 2.14 |
| 31 | 1,2-Dichloroethane | 1.48 | 0.64 | 2.56 | 1.15 | 1.36 | 0.96 | 1.05 | 0.90 | 0.20 | 2.27 |

(*continued on next page*)

**Table 1** (*continued*)

| # | Compound name | Observed | | | Predicted | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | log *K*(*b*:*a*) | log *K*(*s*:*a*) | log *K*(*o*:*a*) | log *K*(*b*:*a*) by | | | | | log *K*(*s*:*a*) by | log *K*(*o*:*a*) by |
| | | | | | Eq. 2 | Eq. 11 | Eq. 14 | Eq. 15 | Eq. 16 | Eq. 5 | Eq. 6 |
| 32 | 1,1,1-Trichloroethane | 0.75 | −0.12 | 2.47 | 1.35 | 0.88 | 1.09 | 1.16 | 1.28 | 0.06 | 2.68 |
| 33 | 1,1,2-Trichloroethane | 1.76 | 1.12 | 3.25 | 1.58 | 1.94 | 1.24 | 1.61 | 1.32 | 0.13 | 2.87 |
| 34 | 1,1,1,2-Tetrachloroethane | 1.62 | 0.55 | 3.43 | 1.84 | 1.69 | 1.44 | 1.75 | 1.67 | 0.05 | 3.35 |
| 35 | Pentachloroethane | 2.02 | 0.37 | 3.83 | 2.22 | 1.76 | 1.61 | 1.98 | 1.94 | −0.11 | 3.84 |
| 36 | Hexachloroethane | 1.80 | −0.18 | 3.70 | 2.51 | 1.39 | 1.70 | 1.95 | 2.03 | −0.34 | 4.24 |
| 37 | 1-Chloropropane | 0.72 | 0.02 | 2.02 | 0.51 | 0.76 | 0.58 | 0.53 | 0.45 | −0.09 | 1.87 |
| 38 | 2-Chloropropane | 0.49 | −0.09 | 1.84 | 0.56 | 0.62 | 0.50 | 0.48 | 0.44 | −0.10 | 1.74 |
| 39 | 1,2-Dichloropropane | 1.27 | 0.44 | 2.63 | 1.18 | 1.27 | 0.89 | 1.10 | 0.92 | −0.17 | 2.55 |
| 40 | 1,1-Dichloroethylene | 0.70 | −0.46 | 1.81 | 0.84 | 0.40 | 0.71 | 0.88 | 1.05 | −0.21 | 2.25 |
| 41 | *cis*-1,2-Dichloroethylene | 1.33 | 0.51 | 2.44 | 0.81 | 1.23 | 0.82 | 1.22 | 1.08 | −0.14 | 2.38 |
| 42 | *trans*-1,2-Dichloroethylene | 0.98 | 0.15 | 2.25 | 0.81 | 0.94 | 0.82 | 1.10 | 1.08 | −0.14 | 2.38 |
| 43 | Trichloroethylene | 1.29 | −0.08 | 2.74 | 1.27 | 1.02 | 1.12 | 1.36 | 1.43 | −0.17 | 2.99 |
| 44 | Tetrachloroethylene | 1.28 | −0.10 | 3.33 | 1.87 | 1.27 | 1.38 | 1.49 | 1.54 | −0.23 | 3.52 |
| 45 | 3-Chloro-1-propylene | 1.24 | 0.31 | 2.04 | 0.88 | 0.94 | 0.61 | 0.83 | 0.70 | −0.12 | 1.97 |
| 46 | Vinyl chloride | 0.31 | −0.37 | 1.39 | 0.54 | 0.26 | 0.33 | 0.46 | 0.54 | −0.19 | 1.51 |
| 47 | Chlorobenzene | 0.77 | 0.45 | 3.34 | 1.56 | 1.59 | 1.84 | 1.56 | 1.63 | 0.60 | 3.52 |
| 48 | Dibromomethane | 1.87 | 1.16 | 2.98 | 2.03 | 1.84 | 1.82 | 1.83 | 1.91 | 1.31 | 2.70 |
| 49 | 1,2-Dibromoethane | 2.08 | 1.24 | 3.11 | 1.68 | 1.94 | 1.48 | 1.96 | 1.74 | 0.43 | 3.02 |
| 50 | Vinylbromide | 0.61 | −0.36 | 1.75 | 1.04 | 0.43 | 0.68 | 0.83 | 0.96 | −0.16 | 2.14 |
| 51 | 1-Bromopropane | 1.07 | 0.16 | 2.43 | 0.78 | 1.03 | 1.03 | 0.92 | 0.92 | 0.22 | 2.39 |
| 52 | 2-Bromopropane | 0.77 | 0.03 | 2.21 | 0.85 | 0.85 | 0.96 | 0.86 | 0.90 | 0.22 | 2.26 |
| 53 | Difluoromethane | 0.20 | 0.12 | 0.68 | 0.05 | 0.23 | −0.33 | −0.04 | −0.21 | −0.22 | 0.31 |
| 54 | 1,1,2,2,3,3-Hexafluoro-propane | −0.48 | −0.49 | 1.01 | −0.35 | 0.03 | −0.12 | −0.53 | −0.71 | −0.58 | 1.09 |
| 55 | 1,1,2,2,3,3,4,4-Octafluoro-butane | −0.36 | −0.80 | 1.48 | −0.63 | 0.06 | −0.30 | −0.66 | −0.89 | −0.99 | 1.19 |
| 56 | Fluorobenzene | 1.06 | 0.36 | 2.82 | 1.07 | 1.31 | 1.52 | 1.05 | 1.10 | 0.66 | 2.84 |
| 57 | 1,2-Difluorobenzene | 0.96 | 0.49 | 2.93 | 0.86 | 1.43 | 1.48 | 1.05 | 1.03 | 0.59 | 2.84 |
| 58 | 1,4-Difluorobenzene | 0.87 | 0.38 | 2.82 | 0.94 | 1.32 | 1.50 | 1.01 | 1.04 | 0.64 | 2.83 |
| 59 | 1,2,4-Trifluorobenzene | 0.76 | 0.26 | 2.79 | 0.73 | 1.24 | 1.43 | 0.90 | 0.92 | 0.54 | 2.80 |
| 60 | 1,3,5-Trifluorobenzene | 0.49 | −0.14 | 2.53 | 0.93 | 0.90 | 1.35 | 0.76 | 0.88 | 0.41 | 2.79 |
| 61 | Hexafluorobenzene | 0.39 | −0.40 | 2.40 | −0.13 | 0.69 | 1.12 | 0.27 | 0.31 | 0.16 | 2.62 |
| 62 | Methylpentafluorobenzene | 0.73 | −0.37 | 3.17 | 0.26 | 1.05 | 1.19 | 0.79 | 0.88 | −0.02 | 2.96 |
| 63 | Bromochloromethane | 1.62 | 0.94 | 2.56 | 1.25 | 1.53 | 1.29 | 1.74 | 1.77 | 0.71 | 2.36 |
| 64 | Dibromochloromethane | 3.33 | 0.87 | 3.43 | 2.35 | 1.87 | 1.72 | 2.22 | 2.29 | 0.72 | 3.14 |
| 65 | 1-Bromo-2-chloroethane | 1.72 | 0.95 | 2.76 | 1.44 | 1.62 | 1.31 | 1.60 | 1.47 | 0.42 | 2.70 |
| 66 | Chlorofluoromethane | 0.71 | 0.49 | 1.35 | 0.60 | 0.74 | 0.49 | 0.54 | 0.49 | 0.46 | 1.13 |
| 67 | 1,1,1-Trifluoro-2-chloro-ethane | 0.10 | −0.38 | 1.38 | 0.39 | 0.25 | 0.33 | 0.21 | 0.26 | −0.15 | 1.46 |
| 68 | 1,1,1-Trifluoro-2,2-bromo-chloroethane | 0.74 | −0.30 | 2.30 | 1.15 | 0.70 | 0.94 | 1.05 | 1.25 | 0.11 | 2.33 |
| 69 | 2,2-Dichloro-1,1,1-trifluo-roethane | 0.56 | −0.68 | 1.81 | 0.67 | 0.27 | 0.66 | 0.48 | 0.69 | −0.09 | 2.03 |
| 70 | 2-Chloro-1,1,1,2-tetrafluo-roethane | 0.18 | 0.02 | 1.40 | 0.31 | 0.49 | 0.32 | 0.05 | −0.10 | −0.11 | 1.41 |
| 71 | Methanol | 3.52 | 3.55 | 1.82 | 3.45 | 2.70 | 2.61 | 3.43 | 3.54 | 3.77 | 1.58 |
| 72 | Ethanol | 3.37 | 3.38 | 2.04 | 3.15 | 2.70 | 2.87 | 3.15 | 3.22 | 3.66 | 2.17 |
| 73 | 1-Propanol | 3.13 | 3.34 | 2.59 | 3.04 | 2.92 | 2.93 | 3.03 | 2.99 | 3.35 | 2.63 |
| 74 | 2-Propanol | 3.11 | 3.18 | 2.19 | 2.88 | 2.65 | 2.64 | 2.89 | 2.82 | 3.10 | 2.34 |
| 75 | 1-Butanol | 3.19 | 3.17 | 3.05 | 3.17 | 3.03 | 3.14 | 3.05 | 3.01 | 3.19 | 3.17 |
| 76 | 1-Pentanol | 2.92 | 3.05 | 3.45 | 3.12 | 3.13 | 3.13 | 3.02 | 2.92 | 2.85 | 3.52 |
| 77 | 1-Hexanol | 3.21 | 2.89 | 4.04 | 3.32 | 3.30 | 3.28 | 3.12 | 3.04 | 2.68 | 3.99 |
| 78 | 2-Methyl-1-propanol | 2.94 | 3.05 | 2.67 | 2.99 | 2.79 | 2.99 | 3.02 | 3.00 | 3.07 | 3.03 |
| 79 | Diethyl ether | 1.03 | 1.06 | 1.75 | 0.46 | 1.24 | 1.12 | 1.04 | 0.85 | 0.75 | 1.98 |
| 80 | Methyl-*tert*-butyl ether | 1.18 | 1.18 | 2.08 | 0.77 | 1.46 | 1.00 | 1.29 | 0.97 | 0.40 | 2.16 |
| 81 | Methyl acetate | 2.00 | 2.03 | 1.93 | 1.42 | 1.88 | 1.93 | 1.76 | 1.68 | 2.08 | 2.10 |
| 82 | Ethyl acetate | 1.91 | 1.85 | 2.25 | 1.42 | 1.92 | 2.08 | 1.83 | 1.78 | 1.95 | 2.52 |
| 83 | Propyl acetate | 1.88 | 1.72 | 2.70 | 1.60 | 2.04 | 2.19 | 1.96 | 1.91 | 1.77 | 2.92 |
| 84 | Isopropyl acetate | 1.55 | 1.53 | 2.48 | 1.53 | 1.83 | 2.11 | 1.89 | 1.90 | 1.71 | 2.82 |
| 85 | Butyl acetate | 1.95 | 1.51 | 3.21 | 1.79 | 2.14 | 2.27 | 2.08 | 2.06 | 1.55 | 3.30 |
| 86 | Isobutyl acetate | 1.72 | 1.41 | 3.11 | 1.80 | 2.04 | 2.19 | 2.04 | 2.04 | 1.50 | 3.19 |
| 87 | *n*-Pentyl acetate | 1.99 | 1.38 | 3.47 | 1.98 | 2.18 | 2.33 | 2.24 | 2.22 | 1.32 | 3.65 |
| 88 | Acetone | 2.32 | 2.60 | 1.93 | 2.32 | 2.21 | 2.00 | 2.35 | 2.28 | 2.53 | 1.75 |

**Table 1** (*continued*)

| # | Compound name | Observed | | | Predicted | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | log *K*(*b*:*a*) | log *K*(*s*:*a*) | log *K*(*o*:*a*) | log *K*(*b*:*a*) by | | | | | log *K*(*s*:*a*) by | log *K*(*o*:*a*) by |
| | | | | | Eq. 2 | Eq. 11 | Eq. 14 | Eq. 15 | Eq. 16 | Eq. 5 | Eq. 6 |
| 89 | 2-Butanone | 2.22 | 2.40 | 2.42 | 2.14 | 2.31 | 2.18 | 2.29 | 2.23 | 2.39 | 2.23 |
| 90 | 2-Pentanone | 2.10 | 2.22 | 2.80 | 2.23 | 2.37 | 2.29 | 2.32 | 2.28 | 2.23 | 2.61 |
| 91 | 2-Heptanone | 2.35 | 1.98 | 3.83 | 2.61 | 2.68 | 2.49 | 2.53 | 2.46 | 1.79 | 3.45 |
| 92 | 1-Propanethiol | 0.65 | 0.43 | 2.36 | 0.10 | 1.15 | 1.20 | 0.65 | 0.62 | 0.63 | 2.29 |
| 93 | 1-Butanethiol | 0.53 | 0.35 | 2.88 | 0.34 | 1.33 | 1.30 | 0.81 | 0.76 | 0.36 | 2.74 |
| 94 | 1-Hexanethiol | 0.61 | −0.20 | 3.82 | 0.83 | 1.43 | 1.40 | 1.04 | 1.07 | −0.16 | 3.50 |
| 95 | Acrylonitrile | 2.71 | 2.13 | 2.06 | 2.73 | 1.99 | 1.88 | 2.82 | 2.91 | 2.03 | 2.06 |
| 96 | 1-Nitropropane | 2.35 | 2.10 | 3.03 | 2.10 | 2.40 | 2.09 | 2.06 | 1.91 | 1.87 | 2.63 |
| 97 | 2-Nitropropane | 2.26 | 1.99 | 2.81 | 2.14 | 2.24 | 2.03 | 2.01 | 1.89 | 1.85 | 2.54 |
| 98 | 2-Cyanoethylene oxide | 3.22 | 3.38 | 3.06 | 3.45 | 3.15 | 3.17 | 3.72 | 3.84 | 3.54 | 2.86 |
| 99 | Isoflurane | 0.25 | −0.25 | 1.90 | 0.86 | 0.56 | 1.15 | 0.59 | 0.84 | 0.72 | 2.08 |
| 100 | Tricyclo[5.2.1.0 2,6]-decane | 1.79 | −0.68 | 4.11 | 1.72 | 1.28 | 0.99 | 1.67 | 1.64 | −1.25 | 3.88 |
| *Test set* | | | | | | | | | | | |
| 101 | Decane | 1.24 | −2.39 | 4.16 | 1.36 | 0.32 | 0.29 | 0.77 | 0.98 | −2.38 | 3.77 |
| 102 | 1,5-Hexadiene | 0.38 | −0.85 | 2.16 | 0.83 | 0.33 | 0.55 | 0.53 | 0.61 | −0.73 | 2.50 |
| 103 | Cycloheptane | 0.72 | −1.12 | 3.44 | 0.71 | 0.73 | 0.45 | 0.65 | 0.71 | −1.15 | 2.78 |
| 104 | 1,4-Dimethylbenzene | 1.58 | 0.27 | 3.68 | 1.44 | 1.64 | 1.59 | 1.54 | 1.55 | 0.20 | 3.47 |
| 105 | Carbon tetrachloride | 0.66 | −0.46 | 2.57 | 1.61 | 0.73 | 1.40 | 0.97 | 1.24 | 0.32 | 2.98 |
| 106 | 1,1,2,2-Tetrachloroethane | 2.15 | 1.37 | 3.80 | 1.92 | 2.32 | 1.44 | 2.05 | 1.67 | 0.00 | 3.41 |
| 107 | 1,1,2,2-Tetrafluoroethane | −0.12 | −0.18 | 0.67 | −0.06 | 0.06 | −0.05 | −0.35 | −0.50 | −0.24 | 0.85 |
| 108 | Pentafluorobenzene | 0.51 | −0.13 | 2.59 | 0.20 | 0.93 | 1.24 | 0.44 | 0.45 | 0.31 | 2.69 |
| 109 | 3-Methyl-1-butanol | 2.73 | 2.93 | 3.00 | 3.24 | 2.87 | 3.13 | 3.04 | 3.01 | 2.90 | 3.48 |
| 110 | 2-Methoxyethanol | 4.50 | 4.55 | 2.72 | 3.63 | 3.67 | 4.23 | 3.79 | 3.83 | 4.99 | 3.30 |
| 111 | Isopentyl acetate | 1.81 | 1.35 | 3.60 | 2.03 | 2.22 | 2.26 | 2.22 | 2.21 | 1.30 | 3.55 |
| 112 | Methyl isobutyl ketone | 1.90 | 1.90 | 2.97 | 2.20 | 2.26 | 2.17 | 2.27 | 2.19 | 1.74 | 2.91 |
| 113 | 1-Pentanethiol | 0.68 | 0.09 | 3.37 | 0.56 | 1.40 | 1.36 | 0.93 | 0.91 | 0.10 | 3.13 |
| 114 | *t*-Butanol | 2.70 | 2.78 | 2.22 | 2.78 | 2.44 | 2.46 | 2.72 | 2.63 | 2.61 | 2.53 |
| 115 | Ethyl-*t*-butyl ether | 1.06 | 0.92 | 2.28 | 0.93 | 1.40 | 1.10 | 1.41 | 1.14 | 0.19 | 2.55 |
| 116 | 1,1-Dichloro-1-fluoro-ethane | 0.32 | −0.29 | 1.74 | 0.99 | 0.46 | 0.78 | 0.75 | 0.92 | 0.12 | 2.03 |
| 117 | Methyl t-amyl ether | 1.19 | 1.08 | 2.55 | 0.95 | 1.60 | 1.12 | 1.46 | 1.14 | 0.18 | 2.60 |
| 118 | 1,4-Dioxane | 3.27 | 3.32 | 2.83 | 1.29 | 3.01 | 2.39 | 2.37 | 1.93 | 2.42 | 2.60 |
| 119 | Methoxyflurane | 1.40 | 0.62 | 2.98 | 1.18 | 1.53 | 1.66 | 1.57 | 1.68 | 0.90 | 2.84 |
| 120 | Bromobenzene | 2.01 | 1.07 | 4.14 | 2.03 | 2.30 | 2.05 | 2.14 | 2.03 | 0.59 | 3.91 |
| 121 | *tert*-Amyl alcohol | 2.59 | 2.97 | 2.51 | 2.84 | 2.67 | 2.49 | 2.85 | 2.60 | 2.28 | 2.93 |
| 122 | Methane | −1.40 | −1.56 | −0.51 | −0.56 | −1.25 | −1.55 | −0.88 | −0.77 | −1.07 | −1.08 |
| 123 | α-Pinene | 1.23 | −0.92 | 3.46 | 1.71 | 0.86 | 1.18 | 1.58 | 1.77 | −0.77 | 3.73 |
| 124 | Perfluoropropane | −2.90 | −3.17 | −0.68 | −0.63 | −2.25 | −0.41 | −1.81 | −1.22 | −0.81 | 0.79 |
| 125 | Perfluoropentane | −2.56 | −4.64 | −0.18 | −1.28 | −2.87 | −0.92 | −2.28 | −1.48 | −1.75 | 0.83 |
| 126 | Dichlorodifluoromethane | −0.54 | −1.28 | 0.78 | 0.82 | −0.53 | 0.80 | 0.17 | 0.75 | 0.45 | 1.71 |
| 127 | Propene | −0.06 | −1.40 | 0.88 | 0.00 | −0.55 | 0.09 | −0.29 | 0.11 | −0.05 | 0.92 |
| 128 | Ethane | −0.319 | −1.046 | 0.104 | 0.06 | 0.48 | −0.41 | −0.03 | −0.26 | −0.33 | 0.28 |
| 129 | Trichlorofluoromethane | 0.01 | −0.66 | 1.6 | 1.22 | 0.19 | 1.11 | 0.76 | 1.15 | 0.40 | 2.35 |
| 130 | Cyclopropane | −0.12 | −0.70 | 1.11 | 0.07 | −0.05 | −0.07 | −0.05 | 0.02 | −0.30 | 0.89 |
| 131 | 1,2-Dichloro-1,1,2,2-tetra-fluoroethane | −0.91 | −1.70 | 1.25 | 0.52 | −0.56 | 0.60 | −0.19 | 0.31 | −0.07 | 1.89 |
| 132 | Pentane | −0.37 | −1.85 | 1.75 | 0.18 | −0.43 | −0.02 | −0.19 | 0.07 | −1.13 | 1.86 |
| 133 | Enflurane | 0.45 | −0.16 | 1.93 | 0.70 | 0.62 | 1.15 | 0.62 | 0.83 | 0.71 | 2.10 |

collected from the literature[7,9,24,29,31,32,34,36,38,42,43,45,61–67] and used for the QSPR treatment. In some cases, the saline:air partition coefficient values were reported to be identical to water:air values, in the literature compiled by Meulenberg and Vijverberg.[9]

Average values were used when different experimental data were reported for the same compound. For the QSPR treatment, the partitioning data were converted to their logarithms. The corresponding values of ($\log K(b{:}a)$, $\log K(s{:}a)$, and $\log K(o{:}a)$) are given in Table 1. The experimental uncertainty of the rat blood:air partition coefficients in Table 1 is expected to be of about 0.1 logarithmic units.[6]

The full data set (133 compounds) was divided into two subsets: (i) 100 compounds, used as the training set and (ii) 33 compounds, used for testing the predictive power of the models (test set). The 'random selection' technique was used for the generation of subsets. The different classes of compounds present in the training set are all represented in the test set.

## 3. Methodology

The two-dimensional molecular structures of the 133 organic compounds were drawn using ISIS/Draw 2.4.[68] The pre-optimization and three-dimensional conversions of these structures were performed using the molecular mechanics force field method (MM+) included in Hyperchem 7.0.[69] The CODESSA PRO program using the AM1 parameterization within the semi-empirical quantum-chemical program CMOPAC, which is based on MOPAC, was used for the final minimum energy optimizations. The following keywords were used: AM1 VECTORS BONDS PI POLAR PRECISE ENPART EF. A gradient norm 0.01 kcal/Å was used to calculate electronic, geometric, and energetic parameters for the isolated molecules. AM1 force calculations were carried out using keywords AM1 FORCE PRECISE THERMO ROT = 1 to produce thermodynamic parameters. After optimization, the structures were transferred to the CODESSA PRO program to calculate constitutional, topological, geometrical, thermodynamic, quantum chemical, and electrostatic descriptors.[70]

The best multilinear regression (BMLR) procedure was used to find the best correlation models from selected noncollinear descriptors.[71] The BMLR selects the best two-parameter regression equation, the best three-parameter regression equation etc., based on the highest $R^2$ value in the stepwise regression procedure. During the BMLR procedure the descriptor scales are normalized, centered automatically and the final result is given in natural scales. This result has the best representation of the property in the given descriptors pool.

A major decision required in developing successive QSPRs is when to stop adding descriptors to the model during the stepwise regression procedure. The lack of adequate control leads to over-correlated equations, which contain an excess of descriptors and are difficult to interpret in terms of physical interactions. A simple procedure to control the model expansion is the so-called 'break point' in improvement of the statistical quality of the model.[19] By analysis of the plot of the number of descriptors involved in the obtained models versus squared correlation coefficient values (and cross-validated squared correlation coefficient) corresponding to those models, the statistical improvement of the model gets higher (steeper ascent of the relationship) until a point ('break point') is achieved after which the improvement is negligible (low ascent of the relationship) (see Figs. 1–3). The model corresponding to the break point is then considered to be the best/optimum model.

The QSPR models obtained were validated (i) by the leave-one-out method; (ii) using an external data set of 33 compounds, and (iii) by internal correlation whereby 1/3 of the compounds are predicted with the model fitted with 2/3 of the compounds.

## 4. Results and discussion

To build a useful QSPR tool for the prediction of rat blood:air, saline:air and olive oil:air partition coefficients, we studied various types of correlations using the training set. In all cases the three partition coefficients were used in the form of their logarithms. All models are presented as linear regression equations and their statistical characteristics ($N$ is the size of data set, $n$ is the number of parameters in the model, $R^2$ is the squared correlation coefficient, $R^2_{cv}$ is squared cross-validation correlation coefficient, $F$ is Fisher's criterion, and $s^2$ is the squared standard deviation) are given. The models obtained were used to predict the partition coefficients values for the external dataset (test set).

### 4.1. Correlations using theoretical molecular descriptors

First, we correlated the logarithms of each of the partition coefficients, $\log K(b{:}a)$, $\log K(s{:}a)$, $\log K(o{:}a)$, with theoretical molecular descriptors calculated by CODESSA PRO. To find the model with the optimum number of descriptors, we built multiple parameter

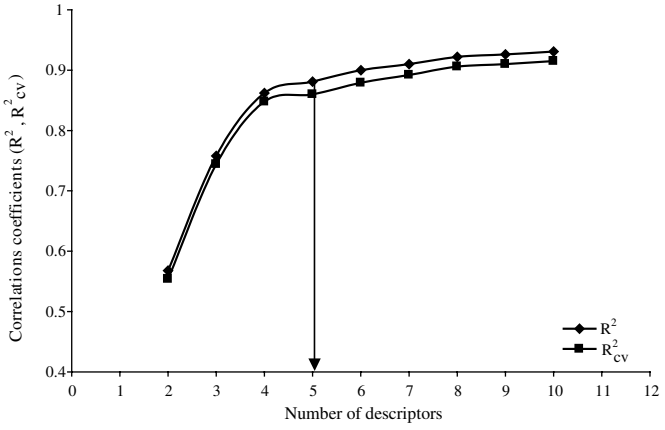


**Figure 1.** Correlation coefficients ($R^2$, $R^2_{cv}$) versus number of descriptors for $\log K(b{:}a)$.

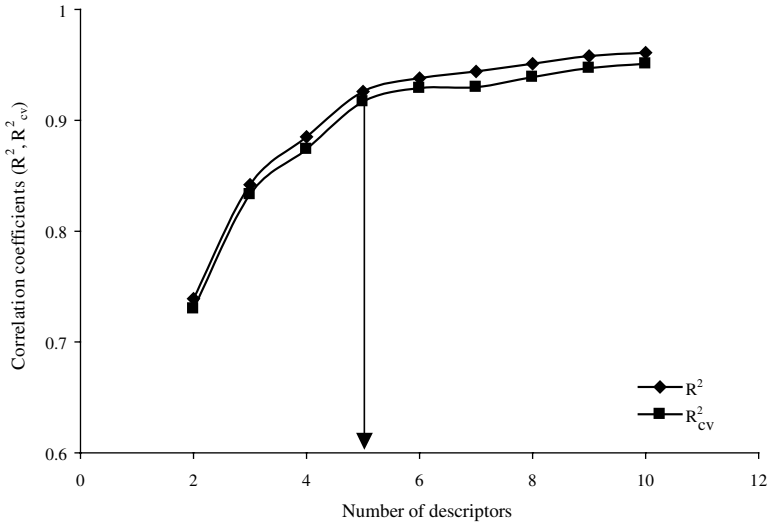


**Figure 2.** Correlation coefficients ($R^2$, $R^2_{cv}$) versus number of descriptors for $\log K(s{:}a)$.

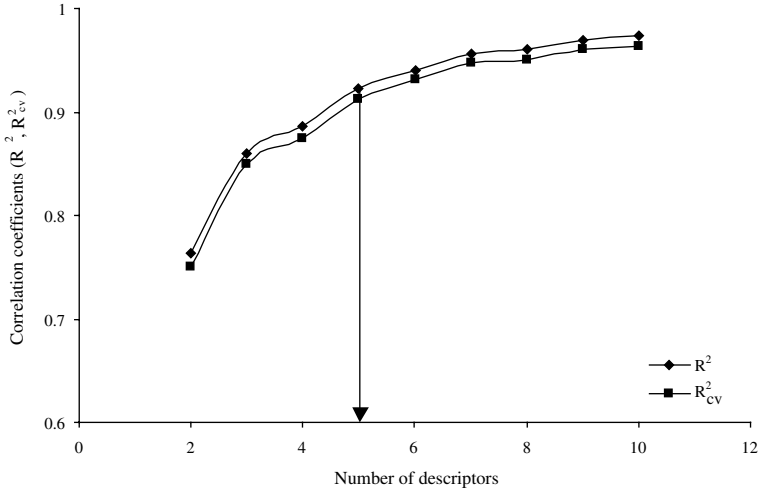


**Figure 3.** Correlation coefficients ($R^2$, $R^2_{cv}$) versus number of descriptors for $\log K(o{:}a)$.

correlations with up to ten descriptors for each partition coefficient. The 'break point' algorithm described in the previous section then suggested that the 'best' model in all three cases contains five parameters (see Figs. 1–3).

These five-parameter models are Eqs. 2, 5, 6. The correlations between the predicted and observed values for the logarithm of rat blood:air, saline:air, and olive oil:air partition coefficients are plotted in Figs. 4–6, respectively. These figures also show the predictive power of each model. The plot of predicted versus observed values for the test set using the five-parameter model are given together with the corresponding value of the squared correlation coefficient.

The names of the descriptors involved in the models, and the corresponding symbols used are shown in Table 2.

**4.1.1. Correlations of rat blood:air partition coefficients.** The best five-parameter model obtained by correlating $\log K(b{:}a)$ with theoretical molecular descriptors is shown as Eq. 2.

$$\begin{aligned}\log K(b:a) &= (-1.300 \pm 0.161) + (38.481 \pm 1.845)D_7 \\ &\quad - (0.611 \pm 0.037)D_{15} + (0.003 \pm 0.0002)D_4 \\ &\quad - (12.654 \pm 1.572)D_3 + (0.036 \pm 0.005)D_1 \\ N &= 100, \quad n = 5, \quad R^2 = 0.881, \quad R^2_{cv} = 0.860, \\ F &= 139.72, \quad s^2 = 0.117\end{aligned} \tag{2}$$

In Eq. 2, the most significant descriptor $D_7$ (area weighted surface charge of hydrogen bonding donor atoms) reflects the hydrogen bond donation ability of a molecule.[72,73] Descriptor $D_7$ is defined by Eq. 3, where $s_D$

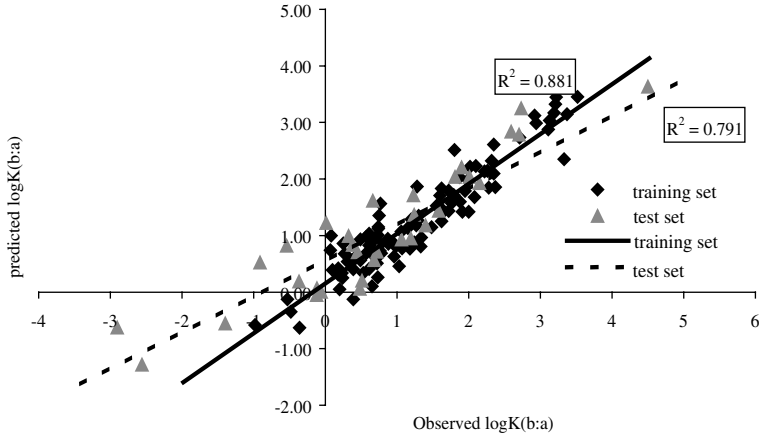


**Figure 4.** Predicted versus observed $\log K(b{:}a)$ for training and test set.

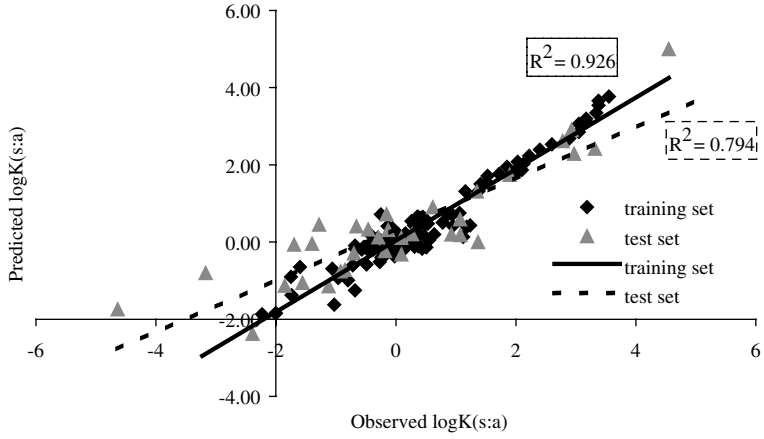


**Figure 5.** Predicted versus observed $\log K(s{:}a)$ for training and test set.

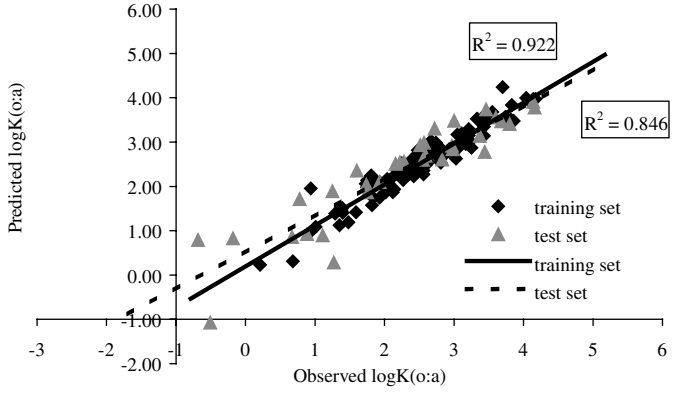


**Figure 6.** Predicted versus observed $\log K(o{:}a)$ for training and test set.

is the solvent-accessible surface area of H-bonding donor H atoms, selected by threshold charge, $q_D$ is the partial charge on H-bonding donor H atoms, selected by threshold charge, and $S_{tot}$ is the total solvent-accessible molecular surface area. The positive contribution of this descriptor (38.48) indicates the strong influence of hydrogen bond donor ability on the blood:air partition coefficient.

$$\mathrm{HDCA2} = \sum_{D} \frac{q_D\sqrt{s_D}}{\sqrt{S_{tot}}} \quad D \in H_{\mathrm{H\text{–}donor}} \tag{3}$$

**Table 2.** Descriptors involved in the QSPR models

| Descriptors | | |
|---|---|---|
| # | Name | Symbol |
| 1 | Average atom weight | $D_1$ |
| 2 | Average bonding information content (order zero) | $D_2$ |
| 3 | FNSA-3 Fractional PNSA (PNSA-3/TMSA) | $D_3$ |
| 4 | Gravitation index (all bonds) | $D_4$ |
| 5 | HA dependent HDCA-1 | $D_5$ |
| 6 | HA dependent HDCA-2 | $D_6$ |
| 7 | HA dependent HDCA-2/SQRT (TMSA) | $D_7$ |
| 8 | HA dependent HDCA-2/TMSA | $D_8$ |
| 9 | HOMO–LUMO energy gap | $D_9$ |
| 10 | Internal heat capacity (300 K) | $D_{10}$ |
| 11 | $\log K(o{:}a)$ experimental | $D_{11}$ |
| 12 | $\log K(o{:}a)$ predicted by Eq. 6 | $D_{12}$ |
| 13 | $\log K(s{:}a)$ experimental | $D_{13}$ |
| 14 | $\log K(s{:}a)$ predicted by Eq. 5 | $D_{14}$ |
| 15 | Number of F atoms | $D_{15}$ |
| 16 | Number of O atoms | $D_{16}$ |
| 17 | Number of single bonds | $D_{17}$ |
| 18 | Randic index (order 1) | $D_{18}$ |
| 19 | Relative number of double bonds | $D_{19}$ |
| 20 | Translational entropy (300 K) | $D_{20}$ |

The second descriptor in the model is the number of fluorine atoms, $D_{15}$, which has a negative contribution to the dependent variable ($\log K(b{:}a)$). The third descriptor in the model is the gravitation index for all bonded atoms $i$ and $j$ in the molecule, $D_4$, defined by Eq. 4, where $m_i$ and $m_j$ are the atomic masses of the bonded atoms and $r_{ij}$ denote the respective bond lengths and $N_b$ is the number of chemical bonds in the molecule. The $D_4$ term is positive for all compounds and reflects the influence of the effective mass distribution in the molecule and molecular dispersion forces in the bulk liquid media on the blood:air distribution.[74]

$$G_b = \sum_{i<j}^{N_b} \frac{m_i m_j}{r_{ij}^2} \tag{4}$$

These two descriptors, $D_4$ and $D_7$, evidently mirror adequately the intermolecular attractive forces in the condensed medium; $D_4$ is connected with the dispersion and cavity formation effects in liquids, whereas $D_7$ accounts for the hydrogen-bonding ability of the compounds.

The fourth descriptor of the model is the fractional atomic charge weighted partial negative surface area, $D_3$.[72,73] As all values of this descriptor are negative, the overall contribution of the $D_3$ term to $\log K(b{:}a)$ is positive. Like $D_7$, the descriptor $D_3$ reflects the electrostatic and hydrogen-bonding related stabilization of the dissolved compounds by the solvent (however, $D_3$ and $D_7$ are not intercorrelated as $R^2 = 0.003$). The fifth descriptor is the average atomic weight, $D_1$, which also has a positive contribution to the model. The regression model was found to be statistically significant ($R^2 = 0.881$). External validation of the model also gave a good correlation coefficient ($R^2 = 0.791$). The applicability of Eq. 2 for the test set is illustrated in Figure 4.

**4.1.2. Correlations of saline:air partition coefficients.** The best five-parameter model obtained for $\log K(s{:}a)$ with theoretical molecular descriptors is given as Eq. 5.

$$\begin{aligned}\log K(s:a) &= (4.451 \pm 0.313) + (1.353 \pm 0.068)D_{16} \\ &\quad + (2.539 \pm 0.132)D_6 - (0.296 \pm 0.024)D_9 \\ &\quad - (0.056 \pm 0.005)D_{10} - (4.687 \pm 0.617)D_{19} \\ N &= 100, \quad n = 5, \quad R^2 = 0.926, \quad R^2_{cv} = 0.917, \\ F &= 233.99, \quad s^2 = 0.130\end{aligned} \tag{5}$$

The descriptors involved in this model are (i) the number of oxygen atoms, $D_{16}$; (ii) the HA dependent HDCA-2 (calculated using AM1 partial charges), $D_6$; (iii) the internal heat capacity (at 300 K), $D_{10}$; (iv) the HOMO–LUMO energy gap, $D_9$; (v) the relative number of double bonds, $D_{19}$. The two most important descriptors in this regression model are $D_{16}$ and $D_6$, as indicated by the $t$-test values (19.8 and 19.2). $D_6$ is known to represent adequately the specific hydrogen-bonding interactions in pure liquids.[74,75] The inclusion of the number of oxygen atoms in the correlation model may be a result of the inadequacy of AM1 calculated charges to describe fully accurately the electrostatic and hydrogen bonding interactions. The HOMO–LUMO energy gap, $D_9$, relates to the dispersion energy of polar solutes in solutions.[76,77] The thermodynamic parameter; internal heat capacity, $D_{10}$, represents the compactness of the molecule and suggests that highly compact molecules are preferentially distributed in air over the solution. Inclusion of the relative number of double bonds in a molecule, $D_{19}$, statistically improves the regression model.

The quality of correlations between predicted and observed $\log K(s{:}a)$, both for training set ($R^2 = 0.926$) and for test set ($R^2 = 0.794$), indicates that saline:air partition coefficients relate well to theoretical molecular descriptors calculated by CODESSA PRO (see Fig. 5).

**4.1.3. Correlations of olive oil:air partition coefficients.** The best five-parameter model obtained for $\log K(o{:}a)$ with theoretical molecular descriptors is Eq. 6.

$$\begin{aligned}\log K(o:a) &= (-14.325 \pm 1.012) - (0.493 \pm 0.021)D_{15} \\ &\quad + (0.711 \pm 0.039)D_{18} + (0.392 \pm 0.026)D_{20} \\ &\quad + (0.491 \pm 0.039)D_5 - (0.044 \pm 0.006)D_{17} \\ N &= 100, \quad n = 5, \quad R^2 = 0.922, \quad R^2_{cv} = 0.912, \\ F &= 222.65, \quad s^2 = 0.054\end{aligned} \tag{6}$$

The regression model for the olive oil:air partition coefficient given in Eq. 6 correlates well with descriptors derived solely from molecular structure, as indicated by $R^2 = 0.922$. External validation performed on the test set shows the good predictive power of this five-parameter model for $\log K(o{:}a)$ ($R^2 = 0.846$) (see Fig. 6).

The most important descriptors involved in the regression equation are the number of fluorine atoms, $D_{15}$, and Randi index (order 1), $D_{18}$, which is defined by Eq. 7 and is a measure of the compactness of the molecule.[78,79] In this equation, $d_i$ and $d_j$ are the edges of adjacent atoms $i$ and $j$ and the summation is performed over all pairs of edges $i$ and $j$ in the molecule.

$$^1\chi = \sum_{\text{path}} (d_i d_j)^{-1/2} \tag{7}$$

The third descriptor, translational entropy (at 300 K), $D_{20}$, is a bulk parameter depending on the shape and size of the molecule and shows that a larger molecule with flexible conformation is more likely to be found in oil rather than air.[80]

The next relevant descriptor is the HA dependent HDCA-1 calculated using Zefirov's partial charges,[81] $D_5$, that is related to the hydrogen bonding donor ability of the molecule.[72,73] The mathematical expression of this descriptor is given by Eq. 8, where $s_D$ is the solvent-accessible surface area of H-bonding donor H atoms, selected by threshold charge.

$$\text{HDCA1} = \sum_D s_D \quad D \in H_{\text{H–donor}} \tag{8}$$

The number of single bonds in the molecule, $D_{17}$, as a representative variable for all molecules, brings a negative contribution to the model.

### 4.2. Correlations of rat blood:air partition coefficient using saline:air and olive oil:air coefficients as external descriptors

In this section, the relationships between $\log K(b{:}a)$ and the other two partition coefficients are discussed, using both experimental values of $\log K(s{:}a)$ and $\log K(o{:}a)$ ($D_{13}$ and $D_{11}$ and also predicted values of $\log K(s{:}a)$ and $\log K(o{:}a)$ ($D_{14}$ and $D_{12}$, from Eqs. 5 and 6). The models obtained are given in Table 3 (see Table 2 for the designation of descriptors).

The Eqs. 11 and 14 express satisfactory empirical relationships between the blood:air partition coefficients and a combination of saline:air and oil:air partition coefficients using both the observed (Eq. 11, $R^2_{11} = 0.848$) and the predicted (Eq. 14, $R^2_{14} = 0.786$) data. We may also conclude that (i) the correlation coefficient for experimental values is 0.06 better then that for predicted values, and (ii) the dominant contribution to the models is that of the saline:air partition coefficients ($D_{13}$ or $D_{14}$).

### 4.3. Correlations of rat blood:air partition coefficient using the entire pool of descriptors: theoretical and external (experimental or predicted values for $\log K(s{:}a)$ and $\log K(o{:}a)$)

This part of the present study is concerned with building the QSPR models for rat blood:air partition coefficients using an extended pool of descriptors. The original pool of descriptors, calculated by CODESSA PRO, was extended by adding the logarithmic functions of the saline:air and olive oil:air partition coefficients. The experimental values ($D_{13}$, $D_{11}$) and the predicted values ($D_{14}$, $D_{12}$, obtained by using Eqs. 5 and 6) were each used in two separate treatments.

In both cases, QSPR models of up to ten parameters were built and the 'break point' procedure applied for determining the best equation. The break point appears at five descriptors in both cases (see Figs. 7 and 8).

The theoretical descriptors involved in each of the two equations are found to be identical. Thus, the only difference is that in one equation the experimental $\log K(s{:}a)$ ($D_{13}$) appears as a descriptor, whereas in the second equation the predicted $\log K(s{:}a)$ ($D_{14}$) is used.

These two models are given below, in Eqs. 15 and 16, again Eq. 15 using the experimental values for $K(s{:}a)$ is somewhat superior.

$$\begin{aligned}\log K(b:a) &= (-0.837 \pm 0.161) - (0.410 \pm 0.028)D_{15} \\ &\quad + (0.003 \pm 0.0002)D_4 + (0.348 \pm 0.037)D_{13} \\ &\quad + (338.09436.926)D_8 + (1.079 \pm 0.198)D_2 \\ N &= 100, \quad n = 5, \quad R^2 = 0.924, \quad R^2_{\text{cv}} = 0.913, \\ F &= 228.95, \quad s^2 = 0.075\end{aligned} \tag{15}$$

and

**Table 3.** QSPR models for $\log K(b{:}a)$ obtained using only $\log K(s{:}a)$ and $\log K(o{:}a)$ as descriptors

| Eq. no. | Model | $R^2$ | $R^2_{\text{cv}}$ | $F$ | $s^2$ |
|---|---|---|---|---|---|
| *Experimental values $D_{13}$ of $\log K(s{:}a)$ and $D_{11}$ of $\log K(o{:}a)$* | | | | | |
| 9 | $\log K(b{:}a) = 0.950 + 0.637\ D_{13}$ | 0.720 | 0.708 | 252.52 | 0.265 |
| 10 | $\log K(b{:}a) = -0.369 + 0.646\ D_{11}$ | 0.293 | 0.268 | 40.57 | 0.670 |
| 11 | $\log K(b{:}a) = -0.133 + 0.574\ D_{13} + 0.439\ D_{11}$ | 0.848 | 0.837 | 271.60 | 0.145 |
| *Predicted values $D_{14}$ of $\log K(s{:}a)$ and $D_{12}$ of $\log K(o{:}a)$* | | | | | |
| 12 | $\log K(b{:}a) = 0.965 + 0.606\ D_{14}$ | 0.605 | 0.590 | 150.06 | 0.374 |
| 13 | $\log K(b{:}a) = -0.344 + 0.636\ D_{12}$ | 0.261 | 0.235 | 34.73 | 0.700 |
| 14 | $\log K(b{:}a) = -0.369 + 0.568\ D_{14} + 0.533\ D_{12}$ | 0.786 | 0.772 | 178.17 | 0.205 |

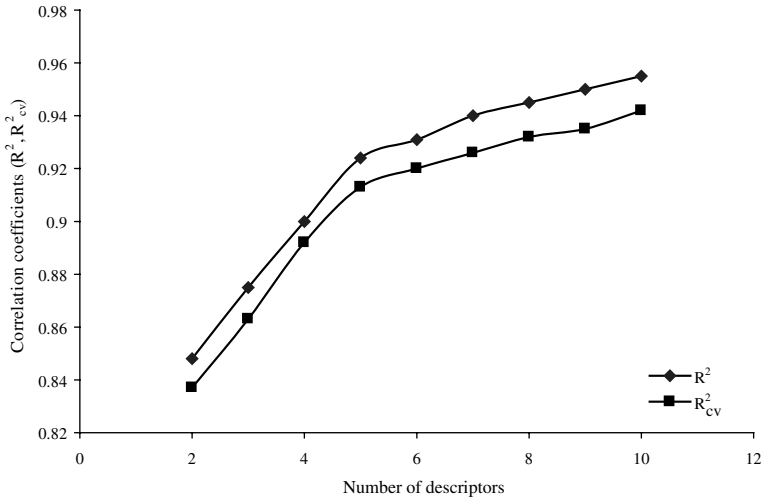


**Figure 7.** Correlation coefficients ($R^2$, $R^2_{cv}$) versus number of descriptors for $\log K(b{:}a)$—using $\log K(s{:}a)$ exp and $\log K(o{:}a)$ exp as external descriptors.

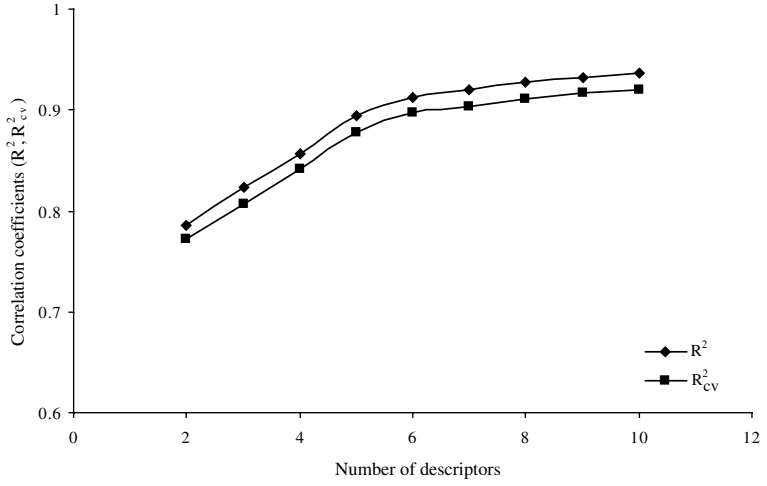


**Figure 8.** Correlation coefficients ($R^2$, $R^2_{cv}$) versus number of descriptors for $\log K(b{:}a)$—using $\log K(s{:}a)$ pred and $\log K(o{:}a)$ pred as external descriptors.

$$\begin{aligned}\log K(b:a) &= (-1.096 \pm 0.184) - (0.466 \pm 0.031)D_{15} \\ &\quad + (0.003 \pm 0.0002)D_4 + (395.57 \pm 44.700)D_8 \\ &\quad + (0.278 \pm 0.046)D_{14} + (1.371 \pm 0.227)D_2 \\ N &= 100, \quad n = 5, \quad R^2 = 0.894, \quad R^2_{cv} = 0.878, \\ F &= 158.66, \quad s^2 = 0.105 \end{aligned} \tag{16}$$

The corresponding plots for these models (Eqs. 15 and 16) are presented in Figures 9 and 10. The predictive power of these models is also demonstrated on these Figures 9 and 10 by the data for the test sets.

The optimal regression model, Eq. 15, for the rat blood:air partition coefficients using experimental data of saline:air partition coefficients together with theoretical descriptors correlates well ($R^2 = 0.924$). The cross-validated correlation coefficient ($R^2_{CV} = 0.913$) and Fisher indexes ($F$) also indicate the statistical significance of the model. Similarly, the regression model, Eq. 16, obtained using the predicted data for $\log K(s{:}a)$ together with theoretical molecular descriptors, also has good statistical characteristics: $R^2 = 0.894$ and $R^2_{CV} = 0.878$.

These two models (Eqs. 15 and 16) involve the same four theoretical descriptors (number of fluorine atoms, $D_{15}$, gravitation index calculated over all bonds, $D_4$, HA dependent HDCA-2/TMSA (calculated from AM1 partial charges), $D_8$, and average bonding information content of zeroth order, $D_2$. The fifth parameter in Eq. 15 is the $\log K(s{:}a)$ experimental, $D_{13}$, and in Eq. 16, the $\log K(s{:}a)$ predicted, $D_{14}$,. With one exception, all of

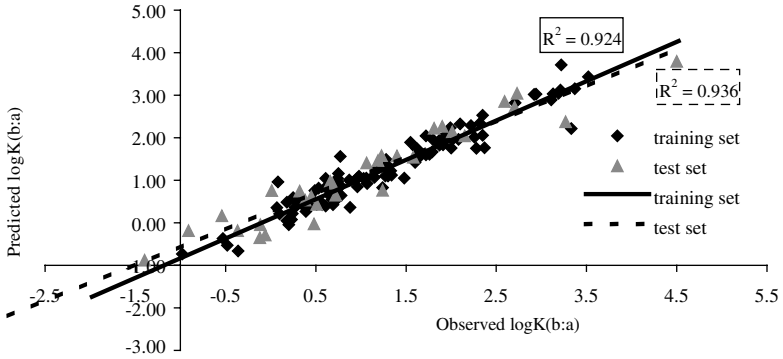


**Figure 9.** Predicted versus observed log $K(b{:}a)$ for training set and test set, using log $K(s{:}a)$ exp and log $K(o{:}a)$ exp as external descriptors.

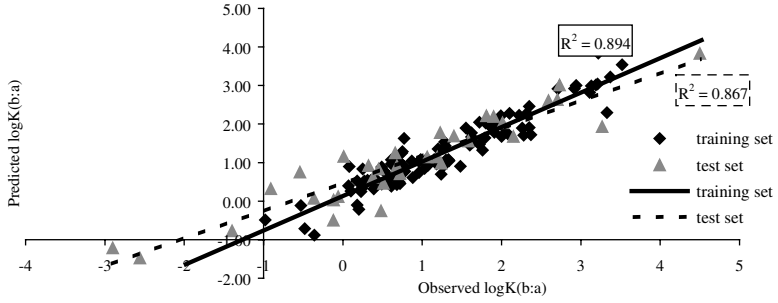


**Figure 10.** Predicted versus observed log $K(b{:}a)$ for training set and test set, using log $K(s{:}a)$ pred and log $(o{:}a)$ pred as external descriptors.

these theoretical descriptors were involved in the previously discussed models and described there. One additional descriptor is the average bonding information content of zeroth order, $^0$BIC, $D_2$, defined by Eqs. 17 and 18, where $n_i$ is the number of atoms in the $i$ th class, $n$ is the total number of atoms in the molecule, $k$ is the number of atomic layers in the coordination sphere around a given atom that are accounted for, and $q$ is the number of edges in the molecular graph. Descriptor $D_2$ encodes the degree of branching of a molecule.[82]

$$^{k}\mathrm{BIC} = {}^{k}\mathrm{IC}/\log_2 q \tag{17}$$

$$^{k}\mathrm{IC} = -\sum_{i=1}^{k} \frac{n_i}{n} \log_2 \frac{n_i}{n} \tag{18}$$

**Table 4.** Internal validation of the five-parameter models of Eqs. 2, 5, 6

| Set | $N$ | $R^2$ (fit) | $s^2$ (fit) | Subset | $N$ | $R^2$ (pred) | $s^2$ (pred) |
|---|---|---|---|---|---|---|---|
| *Eq.* 2 | | | | | | | |
| a + b | 67 | 0.898 | 0.110 | c | 33 | 0.823 | 0.151 |
| a + c | 67 | 0.881 | 0.118 | b | 33 | 0.879 | 0.123 |
| b + c | 66 | 0.881 | 0.118 | a | 34 | 0.857 | 0.144 |
| Average | | 0.887 | 0.115 | | | 0.853 | 0.139 |
| *Eq.* 5 | | | | | | | |
| a + b | 67 | 0.931 | 0.129 | c | 33 | 0.910 | 0.147 |
| a + c | 67 | 0.921 | 0.134 | b | 33 | 0.936 | 0.128 |
| b + c | 66 | 0.928 | 0.134 | a | 34 | 0.914 | 0.138 |
| Average | | 0.927 | 0.132 | | | 0.920 | 0.138 |
| *Eq.* 6 | | | | | | | |
| a + b | 67 | 0.936 | 0.053 | c | 33 | 0.867 | 0.063 |
| a + c | 67 | 0.928 | 0.049 | b | 33 | 0.901 | 0.071 |
| b + c | 66 | 0.906 | 0.059 | a | 34 | 0.947 | 0.045 |
| Average | | 0.923 | 0.054 | | | 0.905 | 0.060 |

Evidently, the combination of the gravitation index and bonding information content descriptors comprises the size and shape information about molecules, and represents adequately the effective dispersion and cavity formation effects for the nonpolar solutes in water.

The observed and predicted values of $\log K(b{:}a)$, $\log K(s{:}a)$, and $\log K(o{:}a)$ for all 133 organic compounds using the regression Eqs. 2, 5, 6, 11, 14, 15 and 16 are given in Table 1.

### 4.4. Additional validation of the models

Following our previous practices,[18,19] we validated Eqs. 2, 5, 6 by dividing the training data set into three subsets: a, b, and c as follows: the set a includes the 1st, 4th, 7th, etc. data points, set b includes 2nd, 5th, 8th, etc. points and set c has the 3rd, 6th, 9th, etc. points. In each of the three validation runs, two subsets were combined into one and the best correlation equation was derived for this set using the identical descriptors as given for the relevant Eqs. 2, 5, or 6 but allows the coefficients to vary. In each case, the equation obtained was used to predict the data for the remaining subset. The results show that the predicted $R^2$ values are in good agreement with our original QSPR models (see Table 4), where $N$ is the number of data points, $R^2$ is the squared correlation coefficient and $s$ is the standard deviation. These characteristics are given for both the fitted equation (fit) and for the predicted data (pred).

## 5. Conclusions

Three QSPR models (Eqs. 2, 5, 6) form a useful tool for predicting the logarithmic function of rat blood:air, saline:air, and olive oil:air partition coefficients. The models were developed using 100 diverse organic compounds as the training set. The three regression equations (Eqs. 2, 5, 6) correlate the above mentioned partition coefficients with theoretical molecular descriptors with squared correlation coefficients ($R^2$) of 0.881, 0.926, and 0.922 for Eqs. 2, 5, 6, respectively.

The predictive power of these models was verified on a test set of 33 organic chemicals. Good statistical characteristics were obtained for all three models: the respective squared correlation coefficients were 0.791 for rat blood:air, 0.794 for saline:air and 0.846 for olive oil:air.

The relationship between the rat blood:air, and saline:air and olive oil:air partition coefficients was also investigated, using the last two partition coefficients in order to predict the first (Eqs. 9–13 and 14).

Based on these results, two alternative 5-parameters QSPR models were developed to calculate the rat blood:air partition coefficients, using four theoretical molecular descriptors and one external parameter (experimental or predicted value of $\log K(s{:}a)$)—Eq. 15 ($R^2 = 0.924$) and Eq. 16 ($R^2 = 0.894$).

Taking into account these promising results, we expect that CODESSA PRO will also derive satisfactory correlation equations for predicting the partition coefficients of different organic compounds in various other biological systems (i.e. tissue:air, tissue–blood etc. for both human and rat) using theoretical molecular descriptors calculated from chemical structures.

## Acknowledgements

We thank Professor Alexandre Varnek (University Louis Pasteur, Strasbourg, France) for helpful discussions and the University of Florida for support.

The Estonian Science Foundation (Grant No. 4548) is gratefully acknowledged for the partial support of this work.

## References and notes

1. Poulin, P.; Krishnan, K. A. *Hum. Exp. Toxicol.* **1995**, *14*, 273.
2. Abraham, M. H.; Kamlet, M. J.; Taft, R. W.; Doherty, R. M.; Weathersby, P. K. *J. Med. Chem.* **1985**, *28*, 865.
3. Kamlet, M. J.; Abraham, D. J.; Doherty, R. M.; Taft, R. W.; Abraham, M. H. *J. Pharm. Sci.* **1986**, *75*, 350.
4. Paterson, S.; Mackay, D. *Br. J. Ind. Med.* **1989**, *46*, 321.
5. Abraham, M. H.; Weathersby, P. K. *J. Pharm. Sci.* **1994**, *83*, 1450.
6. Poulin, P.; Krishnan, K. A. *Toxicol. Appl. Pharmacol.* **1996**, *136*, 131.
7. Gargas, M. L.; Burgess, R. J.; Voisard, D. E.; Cason, G. H.; Andersen, M. E. *Toxicol. Appl. Pharmacol.* **1989**, *98*, 87.
8. Sato, A.; Nakajima, T. *Toxicol. Appl. Pharmacol.* **1979**, *47*, 41.
9. Meulenberg, C. J. W.; Vijverberg, H. P. M. *Toxicol. Appl. Pharmacol.* **2000**, *165*, 206.
10. Meulenberg, C. J. W.; Wijnker, A. G.; Vijverberg, H. P. M. *J. Toxicol. Env. Health A* **2003**, *66*, 1985.
11. Poulin, P.; Krishnan, K. *Toxicol. Methods* **1996**, *6*, 117.
12. Lam, C. W.; Galen, T. J.; Boyd, J. F.; Pierson, D. L. *Toxicol. Appl. Pharmacol.* **1990**, *104*, 117.
13. Beliveau, M.; Krishnan, K. *Toxicol. Lett.* **2000**, *116*, 183.
14. Beliveau, M.; Krishnan, K. *J. Toxicol. Env. Health A* **2000**, *60*, 377.
15. Kaneko, T.; Wang, P.-Y.; Sato, A. *Toxicol.* **2000**, *143*, 203.
16. Basak, S. C.; Mills, D.; Hawkins, D. M.; El-Masri, H. A. *SAR QSAR Env. Res.* **2002**, *13*, 649.
17. Katritzky, A. R.; Mu, L.; Karelson, M. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 1162.
18. Katritzky, A. R.; Wang, Y.; Sild, S.; Tamm, T.; Karelson, M. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 720.
19. Maran, U.; Karelson, M.; Katritzky, A. R. *Quant. Struct.-Act. Relat.* **1999**, *18*, 3.
20. Katritzky, A. R.; Tatham, D. B.; Maran, U. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 1162.

21. Toropov, A. A.; Schultz, T. W. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 560.
22. Reitz, R. H.; McCroskey, P. S.; Park, C. N.; Andersen, M. E.; Gargas, M. L. *Toxicol. Appl. Pharmacol.* **1990**, *105*, 37.
23. Gearhart, J. M.; Seckel, C.; Vinegar, A. *Toxicol. Appl. Pharmacol.* **1993**, *119*, 258.
24. Loizou, G. D.; Anders, M. W. *Drug Metab. Dispos.* **1993**, *21*, 634.
25. Kedderis, G. L.; Carfagna, M. A.; Held, S. D.; Batra, R.; Murphy, J. E.; Gargas, M. L. *Toxicol. Appl. Pharmacol.* **1993**, *123*, 274.
26. Liu, J.; Laster, M. J.; Taheri, S.; Eger, E. I., II; Chortkoff, B.; Halsey, M. J. *Anesth. Analg.* **1994**, *79*, 1049.
27. Andersen, M. E.; Clewell, H. J., III; Mahle, D. A.; Gearhart, J. M. *Toxicol. Appl. Pharmacol.* **1994**, *128*, 158.
28. Evans, M. V.; Crank, W. D.; Yang, H.-M.; Simmons, J. E. *Toxicol. Appl. Pharmacol.* **1994**, *128*, 36.
29. Loizou, G. D.; Urban, G.; Dekant, W.; Anders, M. W. *Drug Metab. Dispos.* **1994**, *22*, 511.
30. Vinegar, A.; Williams, R. J.; Fisher, J. W.; McDougal, J. N. *Toxicol. Appl. Pharmacol.* **1994**, *129*, 103.
31. Kaneko, T.; Wang, P.-Y.; Sato, A. *Occup. Environ. Med.* **1994**, *51*, 68.
32. Teo, S. K. O.; Kedderis, G. L.; Gargas, M. L. *Toxicol. Appl. Pharmacol.* **1994**, *128*, 92.
33. Barton, H. A.; Creech, J. R.; Godin, C. S.; Randall, G. M.; Seckel, C. S. *Toxicol. Appl. Pharmacol.* **1995**, *130*, 237.
34. Loizou, G. D.; Anders, M. W. *Drug Metab. Dispos.* **1995**, *23*, 875.
35. Knaak, J. B.; Al-Bayati, M. A.; Raabe, O. G. *Toxicol. Lett.* **1995**, *79*, 87.
36. Fang, Z.; Sonner, J.; Laster, M. J.; Ionescu, P.; Kandel, L.; Koblin, D. D.; Eger, E. I., II; Halsey, M. J. *Anesth. Analg.* **1996**, *83*, 1097.
37. Fassoulaki, A.; Eger, E. I., II. *Br. J. Anaesth.* **1986**, *58*, 327.
38. Fang, Z.; Ionescu, P.; Chortkoff, B. S.; Kandel, L.; Sonner, J.; Laster, M. J.; Eger, E. I., II. *Anesth. Analg.* **1997**, *84*, 1042.
39. Tardif, R.; Charest-Tardif, G.; Brodeur, J.; Krishnan, K. *Toxicol. Appl. Pharmacol.* **1997**, *144*, 120.
40. Cantoreggi, S.; Keller, D. A. *Toxicol. Appl. Pharmacol.* **1997**, *143*, 130.
41. Plowchalk, D. R.; Andersen, M. E.; Bogdanffy, M. S. *Toxicol. Appl. Pharmacol.* **1997**, *142*, 386.
42. Knaak, J. B.; Smith, L. W.; Fitzpatrick, R. D.; Olson, J. R.; Newton, P. E. *Inhal. Toxicol.* **1998**, *10*, 65.
43. Zhang, Y.; Trudell, J. R.; Mascia, M. P.; Laster, M. J.; Gong, D. H.; Harris, R. A.; Eger, E. I., II. *Anesth. Analg.* **2000**, *91*, 1294.
44. Hays, S. M.; Elswick, B. A.; Blumenthal, G. M.; Welsch, F.; Conolly, R. B.; Gargas, M. L. *Toxicol. Appl. Pharmacol.* **2000**, *163*, 67.
45. Eger, E. I., II; Laster, M. J. *Anesth. Analg.* **2001**, *92*, 1477.
46. Beliveau, M.; Charest-Tardif, G.; Krishnan, K. *Chemosphere* **2001**, *44*, 377.
47. Eger, E. I., II; Liu, J; Koblin, D. D.; Laster, M. J.; Taheri, S.; Halsey, M. J.; Ionescu, P.; Chortkoff, B. S.; Hudlicky, T. *Anesth. Analg.* **1994**, *79*, 245.
48. Meyer, M.; Tebbe, U.; Piiper, J. *Pflug. Arch.* **1980**, *384*, 131.
49. Levitt, M. D.; Levitt, D. G. *J. Clin. Invest.* **1973**, *52*, 1852.
50. Loizou, G. D.; Anders, M. W. *Drug Metab. Dispos.* **1993**, *21*, 634.
51. Niazi, S.; Chiou, W. L. *J. Pharm. Sci.* **1975**, *64*, 1538.
52. Niazi, S.; Chiou, W. L. *J. Pharm. Sci.* **1974**, *63*, 532.
53. Koblin, D. D.; Chortkoff, B. S.; Laster, M. J.; Eger, E. I., II; Halsey, M. J.; Ionescu, P. *Anesth. Analg.* **1994**, *79*, 1043.
54. Filser, J. G.; Schmidbauer, R.; Rampf, F.; Baur, C. M.; Pütz, C.; Csanady, G. A. *Toxicol. Appl. Pharmacol.* **2000**, *169*, 40.
55. Aarstad, K.; Becker, R.; Dahl, J.; Dybing, E.; Nilsen, O. G. *Pharmacol. Toxicol.* **1990**, *67*, 284.
56. Morris, J. B.; Hassett, D. N.; Blanchard, K. T. *Toxicol. Appl. Pharmacol.* **1993**, *123*, 120.
57. Chortkoff, B. S.; Laster, M. J.; Koblin, D. D.; Taheri, S.; Eger, E. I., II; Halsey, M. J. *Anesth. Analg.* **1994**, *79*, 234.
58. Allerheiligen, S. R. B.; Ludden, T. M.; Burk, R. F. *Drug Metab. Dispos.* **1987**, *15*, 794.
59. Guitart, R. *Rev. Esp. Fisiol.* **1993**, *49*, 195.
60. Csanady, Gy. A.; Denk, B.; Pütz, C.; Kreuzer, P. E.; Kessler, W.; Baur, C.; Gargas, M. L.; Filser, J. C. *Toxicol. Appl. Pharmacol.* **2000**, *165*, 1.
61. Taheri, S.; Laster, M. J.; Liu, J.; Eger, E. I., II; Halsey, M. J.; Koblin, D. D. *Anesth. Analg.* **1993**, *77*, 7.
62. Abraham, M. H.; Andonian-Haftvan, J.; Whiting, G. S.; Leo, A.; Taft, R. S. *J. Chem. Soc., Perkin Trans. 2* **1994**, 1777.
63. Wilhelm, E.; Battino, R.; Wilcock, R. J. *Chem. Rev.* **1977**, *77*, 219.
64. Abraham, M. H.; Fuchs, R. *J. Chem. Soc., Perkin Trans. 2* **1988**, 523.
65. Abraham, M. H.; Grellier, P. L.; McGill, R. A. *J. Chem. Soc., Perkin Trans. 2* **1987**, 797.
66. Cabala, R.; Svobodova, J.; Feltl, L.; Tichy, M. *Chromatographica* **1992**, *34*, 601.
67. Falk, A.; Gullstrand, E.; Löf, A.; Wigaeus-Hjelm, E. *Br. J. Ind. Med* **1990**, *47*, 62.
68. www.mdl.com.
69. www.hyper.com.
70. www.codessa-pro.com.
71. Karelson, M. *Molecular Descriptors in QSAR/QSPR*; J. Wiley & Sons: New York, 2000.
72. Stanton, D. T.; Jurs, P. C. *Anal. Chem.* **1990**, *62*, 2323.
73. Stanton, D. T.; Egolf, L. M.; Jurs, P. C.; Hicks, M. G. *J. Chem. Inf. Comput. Sci.* **1992**, *32*, 306.
74. Katritzky, A. R.; Mu, L.; Lobanov, V. S.; Karelson, M. *J. Phys. Chem.* **1996**, *100*, 10400.
75. Huibers, P. D. T.; Lobanov, V. S.; Katritzky, A. R.; Shah, D. O.; Karelson, M. *Langmuir* **1996**, *12*, 1462.
76. Roesch, N.; Zerner, M. C. *J. Phys. Chem.* **1994**, *98*, 5817.
77. Amos, A. T.; Burrows, B. L. *Adv. Quant. Chem.* **1973**, *7*, 289.
78. Randic, M. *J. Am. Chem. Soc.* **1975**, *97*, 6609.
79. Kier, L. B.; Hall, L. H. *Molecular Connectivity in Structure–Activity Analysis*; J. Wiley & Sons: New York, 1986.
80. McQuarrie, D. A. *Statistical Thermodynamics*; Harper & Row: New York, 1973.
81. Zefirov, N. S.; Kirpichenok, M. A.; Ismailov, F. F.; Trofimov, M. I. *Dokl. Akad. Nauk SSSR* **1987**, *296*, 883.
82. Basak, S. C.; Harriss, D. K.; Magnuson, V. R. *J. Pharm. Sci.* **1984**, *73*, 429.