DISSERTATIONES CHIMICAE UNIVERSITATIS TARTUENSIS

52

QSPR MODELING OF SOME PROPERTIES
OF ORGANIC COMPOUNDS

KAIDO TÄMM

Department of Chemistry, University of Tartu, Estonia

Dissertation is accepted for the commencement of the degree of Doctor of Philosophy in Chemistry on 19 April, 2006, by the Doctoral Committee of the Department of Chemistry, University of Tartu.

Opponent: Dr. Cemil Öğretir, Professor of Organic Chemistry. Head of the Chemistry Department, Faculty of Arts & Sciences, Osmangazi University, 26020 Eskişehir, Turkey.

Commencement: June 12, 2006 at 18 Ülikooli St., room 204

CONTENTS

LIST OF ORIGINAL PUBLICATIONS

The present thesis consists of four articles listed below. All papers are denoted in text by Roman numerals I–IV.

I   Tämm, K., Fara, D. C., Katritzky, A. R., Burk, P., Karelson, M. A Quantitative Structure-Property Relationship Study of Lithium Cation Basicities. *J. Chem. Phys. A,* **2004**, *108*, 4812–4818.

II   Tämm, K., Burk, P. QSPR Analysis for Infinite Dilution Activity Coefficients of Organic Compounds. *J. Mol. Mod.* **2006**, *12*, 417–421

III Katritzky, A. R., Tämm, K., Kuanar, M., Fara, D. C. Oliferenko, A., Oliferenko, P., Huddleston, J. G., Rogers, R. D. Aqueous Biphasic Systems. Partitioning of Organic Molecules: A QSPR Treatment. *J. Chem. Inf. Comp. Sci*., **2004**, *44*, 136–142.

IV Katritzky, A. R., Dobchev, D. A., Fara, D. C., Hür, E., Tämm, K., Kurunczi, L., Karelson, M., Varnek, A., Solov'ev, V. P. Skin Permeation Rate as a Function of Chemical Structure. *J. Med. Chem*, **2006**, Accepted.

Publications I–III: The author is responsible for the preparation of the data set, calculation of the QSPR models, and interpretation of the results and the preparation of the manuscript.

Publication IV: The author is responsible for the preparation of the data set and writing the manuscript.

LIST OF ABBREVIATIONS

ABS	Aqueous Biphasic System
ANN	Artifical Neural Network
BMLR	Best Multilinear Regression
CODESSA	COmprehensive DEscriptors for Structural and Statistical Analysis
DFT	Density Functional Theory
IL	Ionic Liquid
ISIDA	In Silico Design and Data Analysis
LCA	Lithium Cation Affinity
LCB	Lithium Cation Basicity
LFER	Linear Free Energy Relationship
MOPAC	Molecular Orbital PACkage
PEG	Polyethylene glycol
QSAR	Quantitative Structure — Activity Relationship(s)
QSPR	Quantitative Structure — Property Relationship(s)
TLL	Tie-Line Length

INTRODUCTION

Chemical structure and properties of substances are of paramount importance, not only to chemists, but also to all scientists and to humanity in general. As soon as a chemical structure is written, all the physical, chemical, biological, and technological properties need to be defined. The fact that chemical structure is invariant is of great importance. Chemical structures offer an unique lasting and definitive means of representation of each compound.

The theoretical approach for the modeling of technologically important properties is very beneficial in chemistry and related disciplines, as experimental testing of properties of new compounds can be very time-consuming and expensive. Theoretical modeling tools allow very fast and inexpensive predictions for screening of candidate structures with desired properties from chemical databases and newly generated structures.

Several theoretical methods exist for the molecular modeling. The high-level quantum chemical *ab initio* and density functional theory (DFT) methods are mostly applicable to isolated and relatively small molecules. Also, these methods often require enormous amounts of computer time, memory, and disk space. Both semi-empirical and molecular mechanics methods can be applied to bigger systems but comparing with *ab initio* and DFT methods the results of the calculations can be erratic.

Most of the environmentally, technologically, and biologically relevant processes take place in the heterogeneous environment, thus, the complex intra- and intermolecular interactions should be considered. The proper theoretical description of such complex environments is not yet feasible with *ab initio* or semi-empirical methods.

The quantative structure property/activity relationship (QSPR/QSAR)[1–11] methods provide an alternative way for the description of complex molecular systems. A fundamental goal of QSPR studies is to relate simple or complex physical, chemical, biological, and technological properties of substances with molecular descriptors, calculated solely from molecular structure. The molecular descriptor associates a real number with a chemical, and then orders the set of chemicals according to the numerical value of the specific property. Each descriptor or property provides a scale for a particular set of compounds. Various statistical methods can be used to find the mathematical relationship between the molecular descriptors and investigated properties or activities. The QSPR/QSAR models developed can be either linear or nonlinear. In practice, linear models are commonly used, as they are easier to develop and interpret.

The basic strategy of QSPR is to find the optimum quantitative relationship, which can then be used for the prediction of the properties of molecular structures including those unmeasured or even unknown. QSPR has became more attractive for chemists with development of new software tools, which allowed them to discover and to understand how molecular structure influences properties, and very importantly, to predict and design the optimum structure.

QSPR equations are commonly used to predict spectroscopic, chromatographic and other analytical properties of compounds. In recent years, the QSPR approach has been rapidly expanding to diverse areas of industrial and environmental chemistry.

Current thesis provide an overviews for some QSPR/QSAR applications accomplished for the lithium cation basicities, infinite dilution activity coefficients in ionic liquids and partition coefficients in aqueous biphasic systems.

1. LITERATURE OVERVIEW

The QSPR approach has been applied in many different areas, including (i) properties of pure substances (boiling point, critical temperature, vapor pressure, flash point and auto ignition temperature, density, refractive index, melting point); (ii) interactions between different molecular species (octanol/water partition coefficients, aqueous solubilities of liquids and solids, aqueous solubilities of gases and vapors, solvent polarity scales, GC retention times and response factors); (iii) surfactant properties (critical micelle concentrations, cloud points); (iv) complex properties and properties of polymers (polymer glass transition temperatures, polymer refractive indexes, rubber vulcanization accelerations).[12]

1.1 QSPR — general algorithm

The molecular structure of an organic or inorganic compound determines its properties. However, the direct prediction of compound's properties using *ab initio* theory may be extremely difficult or even impossible. Therefore, the inductive establishment of QSPRs uses an indirect approach in order to tackle this problem. It depends on a set of compounds with known properties or activities that is used for the model building. The general QSPR scheme is presented in Fig 1.

Supplementary for the general QSPR scheme shown in Fig. 1 the QSPR model development and validation involves following major steps: (i) decision of the property to be modeled; (ii) construction of the data set; (iii) depiction of the molecular structures (calculation of the descriptors); (iv) selection of the most informative subset of descriptors; (v) mapping of the QSPR model; (vi) interpretation of the model; (vii) validation of the model; (viii) prediction of the property of interest.

This general QSPR methodology can be applied, in principle, to any property that can be related to the molecular structure, and it has been applied for the description and prediction of a wide variety of chemical, physical, and environmental properties and activities.[13]
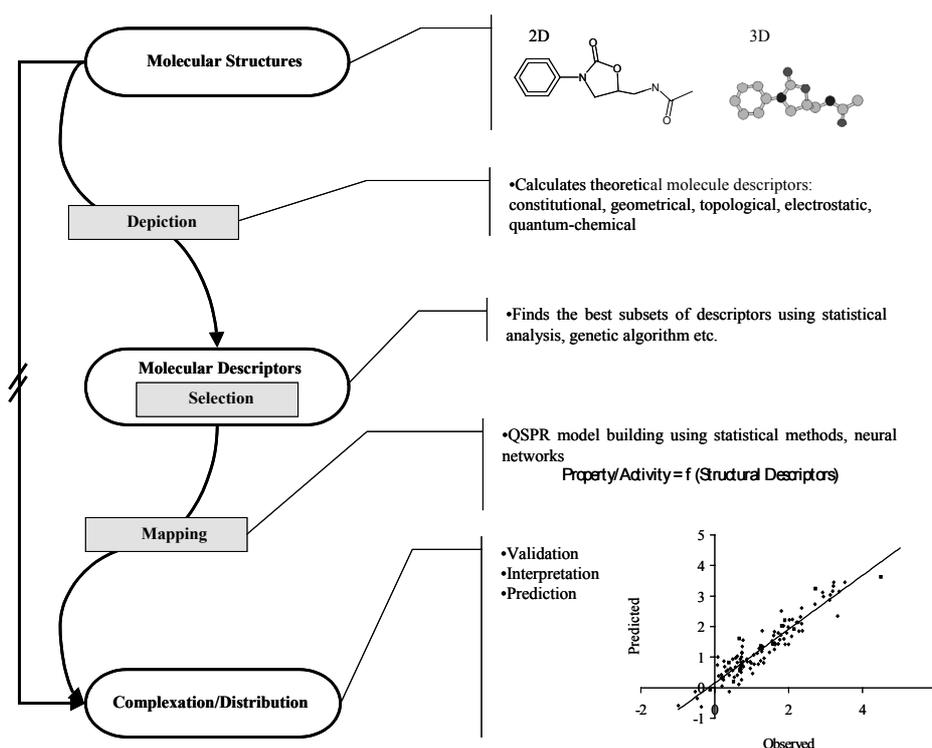
Figure 1. Flow chart of a QSPR model generation

1.2 Molecular descriptors

It is generally accepted that the chemical structure of compound is responsible for its distinct chemical and physical behavior. At the early stages of QSPR/ QSAR development the experimental data was mostly used to reflect the structure of the molecule. Later, especially with developments of quantum chemical methods, theoretical descriptors were calculated solely from molecular structure.

Theoretical molecular descriptors can be divided into several subclasses according to the molecular features they reflect. The classification as shown in Table 1 is somewhat arbitrary since many molecular descriptors are related to several molecular features.

11

Table 1. Classification of molecular descriptors.[14]

MOLECULAR DESCRIPTORS	
CONSTITUTIONAL *(Derived from atomic composition of compound)* * Molecular weight * Counts of atoms and bonds * Counts of rings	TOPOLOGICAL *(Describe the atoms connectivity in the molecule)* * Balaban index * Kier and Hall indices * Wiener index * Randic indices * Information indices, etc.
GEOMETRIC *(Derived from 3D atomic coordinates)* * Principal moments of inertia * Molecular volume * Total solvent-accessible surface * Molecular shadow	ELECTROSTATIC *(Derived from partial charge substitution)* * Partial charges * Polarity indices * Charged partial surface areas
QUANTUM-CHEMICAL *(Derived from the molecular electron wavefunctions)*	
* Net atomic charges * Polarizability * MO energies * Normal modes * Thermodynamical properties	* Dipole moment components * σ- and π-bond orders * FMO reactivity indices * Energy partitioning terms * Electrostatic surface descriptors

The constitutional descriptors (see Table 1) are derived on basis of the molecular formula. Topological descriptors are constructed from the two-dimensional connectivity matrix of the molecule by using algorithms from graph theory. Topological descriptors describe connectivity (chemical bonds) and branching of the molecule. Geometrical descriptors describe size and shape of the molecule and require three-dimensional coordinates of atoms, which are usually extracted from the output of quantum-chemical or molecular mechanics calculations. Electrostatic descriptors reflect charge distribution of the molecule. Quantum chemical descriptors are directly extracted or calculated from the output of semi-empirical or *ab initio* quantum-chemical calculations.

1.3 Methodology

The CODESSA PRO RC3[15] software was used to calculate (i) constitutional, (ii) geometrical, (iii) topological, (iv) electrostatic, (v) quantum chemical, and (vi) thermodynamical molecular descriptors, based on the geometric, electronic, energetic, and thermodynamic characteristics from the MOPAC[16] calculations.

All these descriptors are derived solely from molecular structure and do not require any experimental data.

There are several statistical methods implemented in CODESSA PRO software[17] such as heuristic method, partial least squares method (PLS), multi-linear regression (MLR) method, etc.

In current theses, the best multilinear regression method (BMLR) was commonly used to calculate up the models. The best multilinear regression method (BMLR) is based on the (i) selection of the orthogonal descriptor pairs, (ii) extension of the correlation (saved on the previous step) with the addition of new descriptors until the *Fisher's*-criteria becomes less than that of the best two-parameter correlation. The best N correlations (by R^2) are saved.

Both methods successfully solve the initial selection problem by reducing the number of pairs of descriptors in the "starting set". The major limitations of these methods are the pairwise selection on the first step and the low consistence of the presentation of the upper (according to the selected criteria) segment of the search (N in both cases is 400) due to the small size of the correlation selection.

2. DEVELOPMENT OF QSPR MODELS USING CODESSA-PRO METHODOLOGY

In the present work, the CODESSA-PRO approach and software was used to perform QSPR correlations of diverse physico-chemical properties with whole theoretical molecule descriptors. The main target of these studies was (i) to explore the applicability of QSPR methodology, (ii) to find the best regression equation for the prediction of the property of interest, and (iii) to explain how molecular interactions characterize the studied property.

2.1 A Quantitative Structure-Property Relationship Study of Lithium Cation Basicities

Numerous theoretical techniques at different levels (*ab initio* and DFT) have been used to study the structure and thermodynamic properties (including lithium cation affinity, LCA) of the interaction between the Li^+ cation and different neutral and anionic bases. Most of these studies comprise a fairly small number of similar bases and provide limited relationships between experimental and theoretically calculated LCA values. High level calculations are needed for predictions of LCBs (for e.g. G2 level) and even then some empirical correction is needed to refine absolute LCB values.[18]

Lithium cation gas-phase basicity (LCB) is defined in Eq. 1 as the Gibbs free energy of Eq. 2 associated with the thermodynamic equilibrium of Eq. 3.

$$LCB = -\Delta G_{Li+} \tag{1}$$

$$\Delta G_{Li+} = -RT \ln K_1 \tag{2}$$

$$B + Li^+ \overset{K_1}{\rightleftharpoons} [B - Li^+] \tag{3}$$

In article **I** we attempt to correlate lithium cation gas phase basicities with theoretical descriptors, calculated solely from molecular structure of neutral base using a QSPR technique.

The lithium cation gas phase basicity (LCB) for a diverse set of 205 compounds was satisfactorily described with the six-parameter QSPR model ($R^2 = 0.801$; $R^2_{cv} = 0.785$; $F = 133.11$; $s^2 = 8.78$; $N = 205$).

Our model confirms the electrostatic nature of the lithium cation – base interaction, as the minimum net atomic charge, the HOMO-1 energy, the FPSA-1 (fractional partial positive surface area), the total point-charge component of

the molecular dipole, and the HASA-2 (area-weighted surface charge of hydrogen bonding acceptor atoms) are directly charge related descriptors.

The model also contains the relative number of S atoms. Its importance in the model can be rationalized as a measure of the local polarizability — the sulfur atom has relatively large radius and is easily polarizable.

All the descriptors involved were calculated solely from the chemical structures and have a definite physical meaning corresponding to different intermolecular interactions. Statistical validation of the model demonstrated that the proposed model has a normal statistical stability and validity.

2.2 QSPR Analysis for Infinite Dilution Activity Coefficients of Organic Compounds in Ionic Liquids

In article **II** a quantitative structure property relationship (QSPR) study was carried out for the infinite dilution activity coefficients of organic compounds in ionic liquids, such as 1-methyl-3-ethylimidazolium bis((trifluoromethyl)sulfonyl)imide ([emim][N(Tf)$_2$]), 1,2-dimethyl-3-ethylimidazolium bis((trifluoromethyl)sulfonyl)-imide ([em$_2$im][N(Tf)$_2$]), and 4-methyl-*N*-butylpyridinium tetrafluoroborate ([bmpy][BF$_4$]).

Ionic liquids (ILs) are liquids composed entirely of ions and are defined as the salts with the melting points below 150 ^{o}C. They have garnered increasing interest in the last few years as novel solvents for electrochemistry, biochemistry, and for synthesis and catalysis.

For ILs to be used effectively as solvents, it is essential to know their interaction with different solutes. A quantitative measure of this property is given by the activity coefficient, γ_i, which describes the degree of nonideality for species *I* in a mixture. The infinite dilution activity coefficient, γ_i^{inf}, is especially important because it describes the extreme case, in which only solute-solvent interactions contribute to nonideality. In addition to its theoretical importance, γ_i^{inf} has several practical implications.[19]

A three-parameter model was developed for describing the infinite dilution activity coefficients for a set of 38 organic compounds in three different ILs at temperatures 313 and 343 K. The R^2 values for the models vary from 0.943 up to 0.966. The models include following descriptors: (i) the complementary information content of zeroth order, which describes the atomic connectivity in the molecule and encodes the size and the atomic constitution of the compound; (ii) the fractional partial negative surface area (FNSA-2) describing the polar interactions between molecules, and (iii) the count of hydrogen donor sites directly relates to the hydrogen donor ability of the molecule — the compounds with higher counts of hydrogen donors are more soluble in the ionic liquids.

All descriptors were calculated solely from the chemical structures of solutes and have a definite physical meaning to describe the dilution mechanism of

15

organic compounds in ionic liquids. The example of the internal validation demonstrates the stability and the reliability of the models.

2.3 Aqueous Biphasic Systems. Partitioning of Organic Molecules: A QSPR Treatment

Aqueous biphasic systems (ABS) are formed by the addition of two (or more) water-soluble polymers or a polymer and salt to an aqueous solution above certain critical concentrations or temperature. ABS are unique because each of the two nonmiscible phases is over 80% water on a molal basis and possesses different solvent properties.[20,21]

Article **III** reports the partitioning of 29 small organic probes in a PEG-2000/$(NH_4)_2SO_4$ biphasic system using a QSPR approach. A three-parameter model ($R^2 = 0.967$; $R^2_{cv} = 0.956$; $F = 245.070$ $s^2 = 0.017$) was obtained to describe the partitioning PEG-2000/$(NH_4)_2SO_4$ biphasic system (LogD). The descriptors in the model, such as the gravitational index for all bonds, the image of the Onsager-Kirkwood solvation energy and the minimal net atomic charge were calculated solely from the chemical structures and have definite physical meaning corresponding to different intermolecular interactions.

Using the same descriptors, a three-parameter model was also obtained for LogP (octanol/water, $R^2 = 0.878$; $R^2_{cv} = 0.837$; $F = 64.68$; $s^2 = 0.236$) and the predicted LogP values were used as an external descriptor for modeling LogD ($R^2 = 0.940$; $R^2_{cv} = 0.930$; $F = 424.09$; $s^2 = 0.029$).

All the QSPR models with one, two or three parameters contain the gravitational index or the Kier and Hall indices as major descriptors. This indicates the importance of the "shape descriptors" for describing the partition behavior of small organic compounds in aqueous biphasic systems.

The models obtained in article **III** have some limitations as they have been calculated using LogD values in a specific ABS (PEG-2000/$(NH_4)_2SO_4$). The salt and PEG may vary and the current models may not work properly for predicting LogD in different ABS. Recently, Rogers *et al.*[22] showed the importance of the tie line length [TLL = $(\Delta PEG^2 + \Delta salt^2)^{1/2}$] for describing the ABS, that can be used for normalizing the LogD scale (e.g. LogD/TLL). Anyhow, the models obtained in current work help to improve and understand the prediction of partition coefficients in PEG/salt aqueous biphasic systems (ABS) for structurally diverse compounds.

2.4 Skin Permeation Rate as a Function of Chemical Structure

Article **IV** describes a multilinear and nonlinear QSAR models built for the skin permeation rate ($LogK_p$) of a set of 143 diverse compounds. Permeation rates can be measured by (i) monitoring *in vivo* drug release in live animals or human volunteers, (ii) employing excised skin from human or animal sources, or (iii) using synthetic model membranes as diffusion barriers in *in vitro* experiments.[23]

The main goal of the present study was to relate $LogK_p$ with theoretical descriptors (constitutional, topological, geometrical, charge related, and semi-empirical) calculated solely from the molecular structure. The target property was related to:

(i) Molecular descriptors, which are certain physico-chemical parameters calculated either by quantum mechanical methods or by an empirical technique (CODESSA PRO software). The mathematical form of this QSAR model has the following statistical characteristics: $N = 143$; $n = 5$; $R^2 = 0.800$; $R^2_{cv} = 0.781$; $F = 109.6$; $s = 0.54$.

(ii) Fragment descriptors, which use counts of different types of topological fragments of molecules as variables in a multiple regression analysis (ISIDA package).[24] The statistical parameters of this model ($N = 143$, $n = 41$, $R^2 = 0.907$, $s = 0.43$, $R_{cv}^2 = 0.812$, $F = 24.8$) are slightly better than those obtained for the CODESSA PRO models.

(iii) In addition, an in-house built artificial neural network (ANN) was used for studying the nonlinear dependency between the $LogK_p$ and theoretical molecular descriptors. In neural network modeling, it was implied that the model has to be as simple as possible and to have as good prediction as possible. With only four descriptors in the input, the network was able to predict the data for 40 additional compounds with a reasonable correlation of $R^2 = 0.72$.

Our attempt to correlate the skin permeation rate ($LogK_p$) with theoretically calculated molecular descriptors, fragment descriptors and external descriptor as LogP has led to successful QSAR models that relate this complex molecular property to structural characteristics of the molecules. Notably, all descriptors appearing in the multilinear equations (CODESSA PRO and ISIDA approach) and the ANN model have been derived from theoretical molecular calculations. The exception is the external descriptor LogP (octanol/water), for which the experimental data were taken from the literature. The possibility of using calculated LogP values together with the theoretical molecular descriptors for the prediction of $LogK_p$ was also examined and discussed.

3. CONCLUSIONS

The general algorithm of the QSPR/QSAR methodology was applied for various experimantally measured physico-chemical properties of the compounds. The results are presented in four articles. In article I a QSPR model for the lithium cation gas phase basicities was developed. In article II infinite dilution activity coefficients in ionic liquids were related with the theoretical molecular descriptors. In article III a QSPR models for the partitioning in aqueous biphasic systems were developed. The skin permeation rate was modeled and analyzed in article IV.

I. The lithium cation gas phase basicity (LCB) for a diverse set of 205 compounds was satisfactorily described with the six-parameter QSPR model ($R^2 = 0.801$). Statistical validation of the model demonstrated that the proposed model has a normal stability and validity. The model confirms the electrostatic nature of the lithium cation – base interaction, as the descriptors involved, such as the minimum net atomic charge, the HOMO-1 energy, the FPSA-1, the total point-charge component of the molecular dipole, and the HASA-2 are directly charge related.

II. Three descriptor QSPR models with good statistical parameters ($R^2 = 0.943$–0.966) were obtained to correlate with infinite dilution activity coefficients of organic compounds in three different ionic liquids ([emim][N(Tf)$_2$], [em$_2$im][N(Tf)$_2$], and [bmpy][BF$_4$]) in two different temperatures (313 K and 343 K). All the descriptors involved were calculated solely from the chemical structures and have a definite physical meaning to describe the dilution mechanism of organic compounds in ionic liquids. The example of the internal validation demonstrates the stability and the reliability of the models.

III. Three-parameter QSPR model was developed for the distribution of 29 organic compounds in PEG-2000/(NH$_4$)$_2$SO$_4$ biphasic system with $R^2 = 0.967$. Using the same descriptors, a three-parameter model was also obtained for LogP (octanol/water, $R^2 = 0.878$, $R^2_{cv} = 0.837$) and predicted LogP values were used as an external descriptor for modeling LogD.

IV. Multilinear and nonlinear QSAR models were built for the skin permeation rate (LogK_p) of a set of 143 diverse compounds. Satisfactory models were obtained by three approaches applied: (i) CODESSA PRO, (ii) Neural Network modeling using large pools of theoretical molecular descriptors, and (iii) ISIDA modeling based on fragment descriptors. The predictive abilities of the models were assessed by internal and external validations. The descriptors involved in the models are discussed from the physico-chemical point of view to illuminate the factors that influence skin permeation.

REFERENCES

1. Hansch, C.; Leo, A. *Exploring QSAR, Fundamentals and Applications in Chemistry and Biology*; ACS: Washington, 1995.
2. Abraham, M. H. New Solute Descriptors for Linear Free Energy Relationships and Quantitative Structure-Activity Relationships, In *Quantitative Treatments of Solute/Solvent Interactions*; Politzer, P.; Murray, J. S.; Eds; Elsevier: Amsterdam, 1994; pp 83–133.
3. Abraham, M. H.; Chadha, H. S.; Dixon, J. P.; Rafols, C.; Treiner, C. *J. Chem. Soc. Perkin Trans. 2* **1995**, 887–894.
4. Balaban, A. T. *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 645–650.
5. Hilal, S. H.; Carreira, L. A.; Karickhoff, S. W. Estimation of Chemical Reactivity Parameter and Physical Properties of Organic Molecules Using SPARC. In *Quantitative Treatments of Solute/Solvent Interactions*; Politzer, P.; Murray, J. S.; Eds; Elsevier: Amsterdam, 1994; pp 291–353.
6. Stuper, A. J.; Brugger, W. E.; Jurs, P. C. *Computer-assisted Studies of Chemical Structure and Biological Function*; New York: While, 1979.
7. Katritzky, A. R.; Lobanov, V. S.; Karelson, M. *Chem. Soc. Rev.* **1995**, 279–287.
8. Kier, L. B.; Hall, L. H. *Molecular Connectivity in Structure-Activity Analysis*; John Wiley & Sons: New York, 1986.
9. Murray, J. S.; Politzer, P. A General Interaction Properties Function (GIPF): An Approach to Understanding and Predicting Molecular Interactions. In *Quantitative Treatments of Solute/Solvent Interactions*; Politzer, P.; Murray, J. S.; Eds.; Elsevier: Amsterdam, 1994; pp 243–289.
10. Randic, M.; Razinger M. On the Characterization of Three-Dimensional Molecular Structure. In *From Chemical Topology to Three-Dimensional Geometry*; Balaban, A. T.; Ed.; Plenum Press: New York, 1996; pp 159–236.
11. Lucic, B.; Trinajstic, N. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 121–132.
12. Katritzky, A. R.; Maran, U.; Lobanov, V. S.; Karelson, M. *J. Chem. Inf. Comput. Sci.* **2000,** *40*, 1–18. (and references therein)
13. Katritzky, A. R.; Fara, D.; Tatham, D.; Petrukhin, R.; Maran, U.; Lomaka, A.; Karelson, M. *Curr. Top. Med. Chem.* **2002**, *2*, 1333–1356
14. Katritzky, A. R.; Fara, D. C. *Annals of West University of Timisoara*, Ser. Chem. **2003**, *12*, 1205–1214.
15. http://www.codessa-pro.com
16. Stewart, J. J. P. MOPAC 7.0, QCPE # 455, http://qcpe.chem.indiana.edu
17. Katritzky, A. R.; Lobanov V.; Karelson, M. CODESSA Reference Manual. University of Florida, Gainesville, 1996.
18. Burk, P.; Koppel, I. A.; Koppel, I.; Kurg, R.; Gal, J-F.; Maria, P-C.; Herreros, M.; Notario, R.; Abboud, J.-L. M.; Anvia, F; Taft, R. W. *J. Phys. Chem.* **2000**, *104*, 2824–2833
19. Sandler, S. I. *Fluid Phase Equilib.* **1996**, *116*, 333–342.
20. Albertsson, P.-A. *Nature* **1958**, *182*, 709–711.
21. Albertsson, P.-A. *Partition of Cell Particles and Macromolecules*; Wiley & Sons: New York, 1986.

22. Willauer, H. D.; Huddleston, J. G.; Rogers, R. D. *Ind. Eng. Chem. Res.* **2002**, *41*, 1892–1904.
23. Cronin, M. T. D.; Dearden, J. C.; Gupta, R.; Moss, G. P. *J. Pharm. Pharmacol.* **1999**, *50*, 143–152.
24. http://infochim.u-strasbg.fr/recherche/isida/index.php

SUMMARY IN ESTONIAN

Mõningate orgaaniliste ühendite omaduste modelleerimine QSPR meetodiga

Kvantitatiivsete struktuur-omadus sõltuvuste (QSPR) meetodit kasutati erinevate ainete eksperimentaalsete füsiko-keemiliste omaduste modelleerimiseks. Uurimistöö tulemused on avaldatud neljas artiklis. Artiklis I korreleeriti liitiumi katiooni gaasifaasilisi aluselisusi (LCB) teoreetiliste molekulaarsete deskriptoritega. Artiklis II tuletati QSPR mudelid orgaaniliste ühendite aktiivsuskoefitsientidele lõpmatul lahjendusel ioonsetes vedelikes. Artiklis III modelleeriti jaotuskoefitsiente kahefaasilistes süsteemides (PEG2000/$(NH_4)_2SO_4$, oktanool/ vesi). Artiklis IV töötati välja erinevad mudelid naha läbitavuse hindamiseks.

I. Liitiumi katiooni gaasifaasilisi aluselisusi kirjeldati kuueparameetrilise QSPR mudeliga (R^2 = 0.801) 205 ühendi suhtes. Mudel kinnitab liitiumi katiooni ja aluse vahelise interaktsiooni elektrostaatilist loomust, kuna viis deskriptorit kuuest on otseselt seotud laengutega. Statistiline valideerimine kinnitab mudeli stabiilsust ja ennustusvõimet.

II. QSPR meetodiga uuriti orgaaniliste ühendite aktiivsuskoefitsente lõpmatul lahjendusel kolmes ioonses vedelikus ([emim][N(Tf)$_2$], [em$_2$im][N(Tf)$_2$] [bmpy][BF$_4$]) kahel erineval temperatuuril (313 K and 343 K). Arvutatud mudelitel olid korrelatsioonikoefititsiendid vahemikus 0.943–0.966. Mudelites sisalduvad deskriptorid on arvutatud molekuli struktuuri põhjal ning on selgitatud nende tähendust orgaanilise ühendi lahustumise kirjeldamiseks ioonsetes vedelikes. Kasutatud valideerimismeetodid annavad kinnitust arvutatud mudelite stabiilsusest ja usaldusväärsusest.

III. QSPR meetodit rakendati 29 madalmolekulaarse aine jaotumise uurimiseks PEG2000/$(NH_4)_2SO_4$ kahefaasilises süsteemis. Jaotuskoefitsiendi (LogD) jaoks saadi kolmeparameetriline võrrand korrelatsioonikoefitsiendiga R^2 = 0.967 Samade deskriptoritega tuletati ka kolmeparameetriline mudel LogP (oktanool/ vesi jaotuskoefitsient) jaoks, R^2 = 0.878. Viimase järgi arvutatud LogP väärtusi kasutati lisadeskriptorina LogD modellerimiseks.

IV. Kasutades 143 ühendi eksperimentaalseid andmeid, töötati välja lineaarsed ja mittelineaarsed QSAR mudelid naha läbitavuse koefitsendi määramiseks. Mudelite arvutamiseks kasutati kolme erinevat lähenemist: (i) CODESSA PRO tarkvara, (ii) Neural Network (NN) modelleerimine, kasutades laialdast teoreetiliste deskriptorite andmebaasi; (iii) ISIDA modelleerimine, mis baseerub fragmendi deskriptoritel. Saadud mudelid valideeriti ning selgitati mudelites sisalduvate deskriptorite tähendust ja olulisust naha läbitavuse määramiseks.

ACKNOWLEDGEMENTS

PUBLICATIONS

CURRICULUM VITAE

Kaido Tämm

Born: 28 April 1975
Citizenship: Estonian
Marital Status: Married, two daughters (1 and 4 years old)
Address: University of Tartu, Institute of Chemical Physics,
              Jakobi 2, Tartu, 51014, Estonia
              Tel: +372 737 5258
              E-mail: karu@ut.ee

Education

1994–1998	*B.Sc.* student, Department of Chemistry, University of Tartu, Estonia
1999–2001	*M.Sc.* student, Institute of Chemical Physics, University of Tartu, Estonia
2001–2004	Ph.D. student, Department of Chemistry, University of Tartu, doctoral advisor Prof. Peeter Burk

Professional experience

2002–2003	Visiting Scholar, Center for Heterocyclic Compounds, Department of Chemistry, University of Florida, USA
2004–present	Reseacher, Institute of Chemical Physics, University of Tartu, Estonia

Publications

1) Tämm, K., Burk, P. QSPR Analysis for Infinite Dilution Activity Coefficients of Organic Compounds. *J. Mol. Mod.* **2006**, *12*, 417–421.
2) Katritzky, A. R., Dobchev, D. A., Fara, D. C., Hür, E., Tämm, K., Kurunczi, L., Karelson, M., Varnek, A., Solov'ev, V. P. Skin Permeation Rate as a Function of Chemical Structure. *J. Med. Chem.*, **2006**, Accepted.
3) Tämm, K., Fara, D.C., Katritzky, A. R., Burk, P., Karelson, M. A Quantitative Structure-Property Relationship Study of Lithium Cation Basicities. *J. Chem. Phys. A,* **2004**, *108*, 4812–4818.
4) Katritzky, A. R., Tämm, K., Kuanar, M., Fara, D. C. Oliferenko, A., Oliferenko, P., Huddleston, J. G., Rogers, R. D. Aqueous Biphasic Systems.

Partitioning of Organic Molecules: A QSPR Treatment. *J. Chem. Inf. Comp. Sci*., **2004**, *44*, 136–42.

5) Katritzky, A. R., Fara, D. C., Yang, H., Tämm, K., Tamm, T., Karelson, M. Quantitative Measures of Solvent Polarity. *Chem. Rev*., **2004**, *104*, 175–98.

6) Burk, P., Mölder, U., Tämm, K. Phosphorus-Carbon Diad Tautomerism in Phosphonium Compounds. Theoretical Study. Proc. *Estonian Acad. Sci. Chem*., **2001**, *50*, 241–253.

ELULOOKIRJELDUS

Kaido Tämm

Sündinud:	28. aprill 1975
Kodakondsus:	Eesti
Perekonnaseis:	abielus, kaks tütart (1. ja 4. a.)
Address:	Tartu Ülikool, Keemilise füüsika instituut, Jakobi 2, Tartu, 51014, Eesti
Tel:	+372 737 5257
E-mail:	karu@ut.ee

Haridus

1994–1998	Tartu Ülikool, Füüsika-keemia teaduskond, keemiaosakond
1999–2001	Magistriõpe, Tartu Ülikool, Keemilise füüsika instituut
2001–2004	Doktoriõpe, Tartu Ülikool, Keemilise füüsika instituut, juhendaja prof. Peeter Burk

Teenistuskäik

2002–2003	Külalisteadur, Florida Ülikool, Gainesville, USA
2004–	Erak. teadur, Keemilise füüsika instituut

Publikatsioonid

1) Tämm, K., Burk, P. QSPR Analysis for Infinite Dilution Activity Coefficients of Organic Compounds. *J. Mol. Mod.*, **2006**, *12*, 417–421.
2) Katritzky, A. R., Dobchev, D. A., Fara, D. C., Hür, E., Tämm, K., Kurunczi, L., Karelson, M., Varnek, A., Solov'ev, V. P. Skin Permeation Rate as a Function of Chemical Structure. *J. Med. Chem.*, **2006**, accepted.
3) Tämm, K., Fara, D. C., Katritzky, A. R., Burk, P., Karelson, M. A Quantitative Structure-Property Relationship Study of Lithium Cation Basicities. *J. Chem. Phys. A,* **2004**, *108*, 4812–4818.
4) Katritzky, A. R., Tämm, K., Kuanar, M., Fara, D. C. Oliferenko, A., Oliferenko, P., Huddleston, J. G., Rogers, R. D. Aqueous Biphasic Systems. Partitioning of Organic Molecules: A QSPR Treatment. *J. Chem. Inf. Comp. Sci*., **2004**, *44*, 136–42.
5) Katritzky, A. R., Fara, D. C., Yang, H., Tämm, K., Tamm, T., Karelson, M. Quantitative Measures of Solvent Polarity. *Chem. Rev*., **2004**, *104*, 175–98.
6) Burk, P., Mölder, U., Tämm, K. Phosphorus-Carbon Diad Tautomerism in Phosphonium Compounds. Theoretical Study. Proc. *Estonian Acad. Sci. Chem*., **2001**, *50*, 241–253.