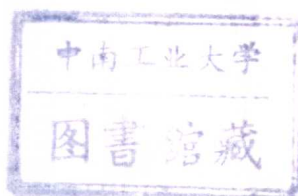


Quantitative Structure-Activity Relationship (QSAR)

■ Edited by

Wang Liansheng Han Shuokui



CHEMICAL INDUSTRY PRESS

This book is founded by the National Natural
Science Foundation of P. R. China

Quantitative Structure-Activity Relationship (QSAR)

Edited by
Wang Liansheng Han Shuokui



Chemical Industry Press
• Beijing, China •

(京)新登字 039 号

图书在版编目(CIP)数据

定量结构活性相关(QSAR)=Quantitative Structure-Activity Relationship(QSAR):英文/王连生主编.
北京:化学工业出版社,1998.10

ISBN 7-5025-2315-4

I. 定… II. 王… III. 有机化学-研究-文集-英文
IV. 062-53

中国版本图书馆 CIP 数据核字(98)第 26098 号

Quantitative Structure-Activity
Relationship(QSAR)

Edited by

Wang Liansheng Han Shuokui

责任编辑:陈丽 白洁

责任校对:陶燕华

封面设计:田彦文

*

化学工业出版社出版发行

(北京市朝阳区惠新里 3 号 邮政编码 100029)

新华书店北京发行所经销

快客印刷厂印刷

三河市东柳装订厂装订

*

开本 787×1092 毫米 1/16 印张 20¼ 字数 507 千字

1999 年 1 月第 1 版 1999 年 1 月北京第 1 次印刷

印数:1—500

ISBN 7-5025-2315-4/TQ·1092

定价:68.00 元

版权所有 违者必究

该书如有缺页、倒页、脱页者,本社发行部负责调换

Introduction

The chemical compounds registered on American "Chemical Abstract" have amounted to 10 million kinds and more, while most of them are new compounds that have never been found in the natural kingdom. It is essential for these chemicals to conduct risk assessment so as to control the production of some chemicals. But the assessing methods are fairly complicated, where it is essential to obtain not only the data on fate and persistence of the compounds, but the data on accumulation, distribution and toxicity in the organism as well. Because the properties of some compounds, such as partition coefficient and solubility, are difficult to be measured by chemical methods, it is essential to find out the relationships between the structure and activity as well as various physicochemical properties so that they are predicted and estimated. In order to complete this work, not only the reliable methods but also the theoretical studies are needed.

Up to now, the development on organic pollution chemistry and relevant subjects has tended towards deduction from description, towards quantitative analysis from qualitative analysis, towards microstructure from macrostate. The structure-activity relationship that studies and analyzes the relationships between the feature of the basic structure in the molecule or atom and some relevant properties obtained by experiments, has become the important content for the basic research on organic pollution chemistry. Therefore, the full classification for the relationships among the composition, structure and property of the compound or molecule, would accelerate greatly to carry out the transition of the structure-activity relationship from empirical science to theoretical science.

As we known, molecule is the basic unit to form compound. Both many physicochemical properties and bioactivity are mainly expressed and explained by the molecule. The structural and composite variation of molecule has a large effect on the physicochemical property and bioactivity. Therefore, it is useful for the structure-activity research to expound the molecular structure.

The goal of studying molecular structure, property and activity is to get hold of the casual relationship between the property of the compound and the structure, and the quantitative change law between the structure and the activity, to summarize the data and find the law so as to extend and perfect the structure-activity research step by step, and make it play a larger and larger role in the following scientific research.

There are a few books concerned on this aspects at home and abroad now. In order to propel environmental chemistry and other subjects, such as chemistry, life science, materia medica, etc. forward, we compile our research work in recent years into this book integrating the project funded by the National Natural Science Fund, introduce it to the readers, and hope it inspire the readers a little in their research. It would be the great pleasure for the authors that the readers can make use of the knowledge from this book in their research field and yield some achievements after they read the whole or partial chapter in the book.

Wang Liansheng
Nanjing University
1998. 4. 7

Contents

CHAPTER 1 QUANTITATIVE STRUCTURE-ACTIVITY RELATIONSHIP

(QSAR) 1

Using MTLSE Model and AM1 Hamiltonian in Quantitative Structure-Activity Relationship

Studies of Alkyl (1-phenylsulfonyl) cycloalkane-carboxylates 1

Using AM1 Hamiltonian in Quantitative Structure-Activity Relationship Studies of

Phenylthio-carboxylates 8

Toxicity of Organic Chemicals to Fathead Minnow: A United Quantitative Structure-Activity Relationship Model and Its Application 15

Using AM1 Hamiltonian in Prediction of Benzaldehydes Sorption on Soil 21

Quantitative Structure-Activity Relationship and Mixture Toxicity Studies of Heterocyclic Nitrogen Compounds 27

The Genotoxicity of Substituted Nitrobenzenes and the Quantitative Structure-Activity Relationship Studies 31

Quantitative Structure-Activity Relationship Studies of Selected Heterocyclic Nitrogen Compounds 36

Quantitative Structure-Activity Relationships of Hydrophobic Organic Chemicals 41

Analysis and Prediction of Structure-Reactive Toxicity Relationships of Substituted Aromatic Compounds 46

Toxicity QSAR of Substituted Benzenes to Yeast *Saccharomyces cerevisiae* 50

Quantitative Structure-Activity Relationships of Organic Acids and Bases 55

Quantitative Structure-Activity Relationships of Organic Pollutants 60

Quantitative Structure-Activity Relationship for Some Substituted Aromatic Compounds 63

Quantitative Structure-Activity Relationships for Substituted Aromatic Compounds 65

Quantitative Structure-Activity Relationships for Studying Alkyl(1-phenylsulfonyl)-cycloalkane-carboxylates 68

The Genotoxicity of Selected Heterocyclic Nitrogen Compounds: A Quantitative Structure-Activity Relationship Study 75

Quantitative Structure-Toxicity Relationships for Fluorine-Contained Aromatics to *Photobacterium phosphoreum* 80

Acute Toxicity of Substituted Biphenyls to *Daphnia magna* and Quantitative Structure-Activity Relationship Study 84

Toxicity of Substituted Benzaldehydes to *Photobacterium phosphoreum* and Quantitative Structure-Activity Relationship 88

Quantitative Structure-Toxicity Relationships for Benzaldehydes to *Daphnia magna* 93

CHAPTER 1 APPLICATION OF QUANTUM CHEMISTRY PARAMETER ON

QSAR 98

Acute Toxicity of Alkyl (1-phenylsulfonyl) cycloalkane-carboxylates to *Photobacterium phosphoreum* and Quantitative Structure-Activity Relationship Study Based on the AM1 Hamiltonian 98

Correlation between Photolysis Rate Constants of Polycyclic Aromatic Hydrocarbons and Frontier Molecular Orbital Energy 104

Using AM1 Hamiltonian in Quantitative Structure-Properties Relationship Studies of Alkyl

(1-phenylsulfonyl) cycloalkane-carboxylates	111
Using AM1 Hamiltonian and Factor Analysis in Prediction of Partition Properties for Phenylthio, Phenylsulfinyl and Phenylsulfonyl Acetates	120
Relationship between Acute Toxicity to Aquatic Organism of Substituted Aromatic Compounds and Their Molecular Orbital Energy	130

CHAPTER III PREDICTING TOXICITY OF SUBSTITUTED AROMATIC HYDROCARBONS

134

Predicting Toxicities of Substituted Aromatic Hydrocarbons to Fish by Toxicities to <i>Daphnia</i> <i>magna</i> or <i>Photobacterium phosphoreum</i>	134
Application of <i>Photobacterium phosphoreum</i> to Toxicity Assessment of Phenylthio-carboxylates	138
Effects of Substituted Anilines and Phenols on Root Elongation of Cabbage Seed	141
Acute Toxicity of Alkyl(1-phenylsulfonyl) cycloalkane-carboxylates to <i>Daphnia magna</i> and Quantitative Structure-Activity Relationships	148
Mutagenicity of Organic Pollutants and Their Active Components in the Xi River Water at Shenyang	153
Analysis of Aromatic Sulfur-containing Compounds Based on Free-Wilson Method	157
Using Theoretical Solvatochromic Parameters in Prediction of Acute Toxicity of Substituted Aromatic Compounds to Aquatic Organisms	160

CHAPTER IV DETERMINATION AND ESTIMATION METHODS OF PHYSIC-CHEMICAL PROPERTY OF ORGANIC COMPOUNDS

164

Determination and Estimation of Physicochemical Properties for Phenylsulfonyl Acetates	164
Determination and Estimation of Partitioning Properties for Phenylthio carboxylates	170
Applicability of Molecular Connectivity Indices to Model Toxicity of Phenylthio carboxylates on <i>Photobacterium phosphoreum</i>	177
Quantification and Application of Lewis Acidity and Basicity of Benzene and Its Derivatives	178
Prediction of Partition Coefficient and Toxicity for Phenylthio, Phenylsulfinyl and Phenylsulfonyl Acetates	185
Prediction the Molar Volume of Aliphatic Compounds Using Edge Adjacency Index	192
Determination and Estimation of the Water Solubilities and the Octanol/Water Partition Coefficients for Derivates of Benzanilides	199
Use of Artificial Neural Network Models in Predicting Biodegradation Rate Constants of Aromatic Sulfur-containing Compounds	204
Aerobic Microbial Degradation of Aromatic Sulfur-containing Compounds and Effect of Chemical Structures	211

CHAPTER V TOXICOLOGY OF ORGANIC POLLUTANTS

217

Solubilities of Polycyclic Aromatic Hydrocarbons and Di-region Theory	217
Molecular Surface Areas of Polycyclic Aromatic Hydrocarbons and Di-region Theory	228
Bioconcentration and Toxicity Effect on Lipid Content of Aquatic Organisms	232
Discriminant Analysis of the Relationship Between Genotoxicity and Molecular Structure of Organochlorine Compounds	236
Relationships between Molecular Structure and Chromosomal Aberrations in <i>In Vitro</i> Human Lymphocytes Induced by Substituted Nitrobenzenes	239

CHAPTER VI HYDROLYSIS OF ORGANIC POLLUTANTS	242
Hydrolysis of Phenylthio, Phenylsulfinyl and Phenylsulfonyl Acetates and Neighboring Group Effect	242
Hydrolysis Kinetics of Phenylsulfonyl-cycloalkane Carboxylates	247
Hydrolysis Kinetics of Phenylthio-carboxylates	251
The Influential Factors of Hydrolysis of Organic Pollutants in Environment	256
 CHAPTER VII SOIL SORPTION OF ORGANIC COMPOUNDS	 261
Prediction of Soil Adsorption Coefficient K_{oc} for Phenylthio, Phenylsulfinyl and Phenylsulfonyl Acetates	261
Sorption of Phenylthioacetates on Natural Soil: Application of Partition-adsorption Mechanism and Model	266
The Prediction of Soil Sorption Coefficients of Heterocyclic Nitrogen Compounds by Octanol/Water Partition Coefficient, Water Solubility and by Molecular Connectivity Indices	273
Contribution of Various Molecular Forces to Sorption of Nonionic Organic Compounds in Natural Soils	277
Determination and Estimation of the Sorption of Benzaldehydes on Soil	283
 CHAPTER VIII MOLECULAR STRUCTURE AND CHROMATOGRAPHY	
RETENTION	288
Prediction Adsorption Coefficients (K_{oc}) for Aromatic Compounds by HPLC Retention Factors (K')	288
Reverse-phase High Performance Liquid Chromatography and Linear Solvation Energy Relationship for Studies of Alkyl(1-phenylsulfonyl) cycloalkane-carboxylates	294
The Estimation of Bioconcentration Factors of Aromatic Hydrocarbons by High-performance Liquid Chromatography	300
Retention of Aromatic Sulfur-containing Compounds on RP-HPLC: Correlation with Partition Coefficients and Molecular Connectivity Indices	305
Prediction of $\lg k'$ Using MCI's and LSER Methods for Heterocyclic Nitrogen Compounds	312

CHAPTER I

QUANTITATIVE STRUCTURE-ACTIVITY RELATIONSHIP (QSAR)

Using MTLSER Model and AM1 Hamiltonian in Quantitative Structure-Activity Relationship Studies of Alkyl (1-phenylsulfonyl)cycloalkane-carboxylates^①

ABSTRACT Based on MTLSER model and quantum chemical descriptors computed by the AM1 hamiltonian, two Quantitative Structure-Activity Relationship (QSAR) equations for 28 alkyl (1-phenylsulfonyl) cycloalkane-carboxylates were obtained. Polarizability is the most significant term influencing toxicity. The toxicity increased with the increase of polarizability and the most negative net atomic charges on oxygen atoms in the sulfone group of the compounds and decreased with the increase of dipole moment and the most positive net atomic charges on a hydrogen atom of the compounds. This study proved the advantages of the MTLSER model. (a) The quantum chemical descriptors in the model can be easily and precisely obtained by computation instead of experiment, a large amount of expenses and time can be saved; (b) The quantum chemical descriptors have clear physicochemical interpretations, and interpretation of the correlation equations can suggest modes of interaction between toxicants and organisms; (c) The MTLSER mode can be used for compounds for which the solvatochromic parameters are not available.

INTRODUCTION Quantitative Structure-Activity Relationships (QSARs), which correlate and predict toxicity of chemicals from their physical and chemical descriptors, play an important role in effects assessment of organic compounds^[1,2]. One kind of QSAR equation is based on the linear free energy relationship (LFER). Burkhard^[3] and Hammett^[4] reviewed the existence of LFERs in 1935 and in 1937 Hammett^[5] proposed the equation that bears his name. Recently, a survey of LFER was given by Exner^[6]. Kamlet and coworkers^[7,8] extended the LFER to involve solute-solvent interactions. This linear solvation energy relationship (LSER) has been successfully used in developing predictive equations for toxicity of a wide range of environmental chemicals to *Photobacterium phosphoreum*, *Daphnia magna*, *Daphnia pulex*, and the fathead minnow (*Pimephales promelas*)^[9,10,11]. The coefficients of the descriptors in the correlation equation also can provide an insight into the nature of the interaction between toxicants and the so called "target molecules" of organisms. However, the use of the LSER model is limited because the descriptors of these models are either empirically determined with a time and money consuming course or only available for a finite number of chemicals. Even though there are tables of LSER parameters and predictive relationships for estimation, solvatochromic parameters for complex molecules are not as easily estimated.

The descriptors derived from quantum chemical computation have obvious advantages, they are not restricted to closely related compounds, can be easily obtained, and describe clearly defined molecular properties. For these reasons, there are many examples of the use of molecular orbital generated descriptors. Some examples were given by Lewis^[12], Stefan Baj *et*

① Reprinted from Jingwen Chen, Liansheng Wang, Using MTLSER Model and AM1 Hamiltonian in Quantitative Structure-Activity Relationship Studies of Alkyl (1-phenylsulfonyl) cycloalkane-Carboxylates, *Chemosphere*, vol. 35 (3) 623~631 (1997), copyright, with permission from "Elsevier Science".

al.^[13], Hansch *et al.*^[14], Nevalainen *et al.*^[15], Mekenyan *et al.*^[16] and Lu Xu *et al.*^[17]. Recently, Wilson, Famini and coworkers^[18,19,20] developed a theoretical set of LSER TLSEr descriptors that are determined solely from computation. Their TLSEr descriptors were developed to obtain optimum correlation with the LSER descriptors, to obtain TLSEr equations with correlation coefficients and standard deviations close in value to those for LSER, and to be as widely applicable to solute-solvent interactions as the LSER set.

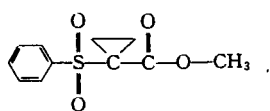
In our previous study^[21], we found it inappropriate to use the TLSEr model for developing prediction equations for partitioning properties of twenty-eight alkyl (1-phenylsulfonyl) cycloalkane-carboxylates, so we deduced a modified TLSEr (MTLSEr) model. The MTLSEr model was successful in developing predictive equations for octanol/water partition coefficients, aqueous solubility, adsorption coefficients for sediments of these compounds.

In the last several decades, aromatic sulfones are used more extensively as intermediates in the manufacture of pesticides, herbicides and anthelmintics^[22]. However, there has been only limited investigation on their QSARs^[23]. Recently, in our laboratory, twenty-eight alkyl (1-phenylsulfonyl) cycloalkane-carboxylates were synthesized, and their acute toxicity to *Daphnia magna straus* was reported^[23]. It is the purpose of our present study to obtain QSAR equations for these compounds based on quantum chemical descriptors and the MTLSEr model and to test and verify the MTLSEr model in fitting the toxicity.

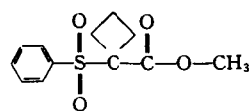
MATERIALS AND METHODS The structures for the alkyl (1-phenylsulfonyl) cycloalkane-carboxylates are given in figure 1-1. The negative logarithm of acute toxicity of the compounds in unit of mol/L, medium immobilization concentration (EC_{50}) and medium lethal concentration (LC_{50}) after 48 hours exposures, is reproduced in table 1-1. Those *Daphnia magna straus* that were unable to swim, but their hearts were still beating under microscope at the end of 48 hours were considered as immobilization, and those that were unable to swim and their hearts stopped beating under microscope at the end of 48 hours were considered as mortality^[23].

The molecular modeling package ALCHEMY I^[24] was used to construct and view all molecular structures. Internal coordinates were used to generate the molecular structure files. Molecular geometries were optimized and quantum descriptors were calculated using the AM1 Hamiltonian^[25] contained in the up-to-date version (Ver. 6.00) of the MOPAC program package^[26]. Compared to other hamiltonians like CNDO/2, MINDO/3, MNDO, etc., the AM1 hamiltonian is a much more recent and common semi-empirical method, gives good estimates of molecular energies^[25]. In addition, the computational time of AM1 is much shorter than that needed by *ab initio* methods. Therefore, AM1 algorithm was selected in this study. MOPAC was run with the following keywords: AM1, PRECISE, ESP, DIPOLE, POLAR, NOINTER. The Keyword ESP (electrostatic potential calculation) was used to obtain ESP derived net atomic charges. All calculations were run on a 80586/75 MHz computer equipped with 16 megabytes of internal memory and supported by Disk Operation System (DOS). The descriptors computed by MOPAC are listed in table 1-1. The computation showed that the most negative atomic charges on oxygen atoms in the sulfone group (SOO) are the largest negative net atomic charges in these molecules. Units of the energy and charge are electron volts (eV) and atomic charge units (a. c. u.) respectively. The unit for dipole moment and polarizability is atomic units (a. u.).

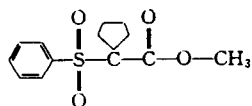
STATGRAPHICS (Ver. 4.0) software^[27] was used to perform regression analysis. Model adequacy was measured as the square of multiple correlation coefficient (R^2), the standard



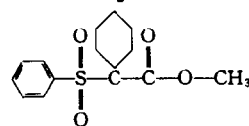
1 #



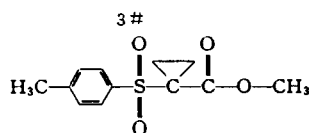
2 #



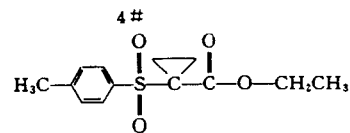
3 #



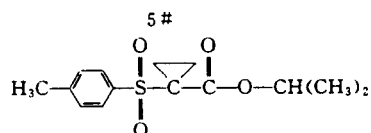
4 #



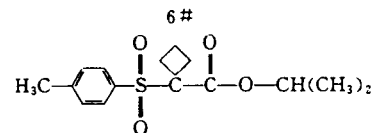
5 #



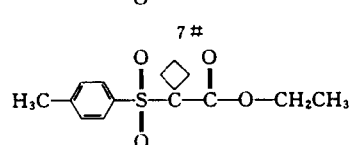
6 #



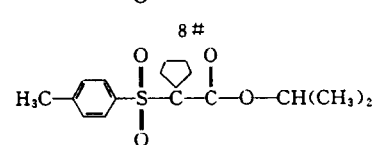
7 #



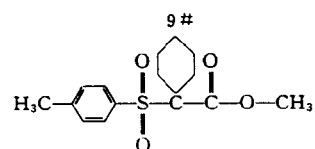
8 #



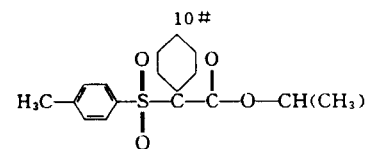
9 #



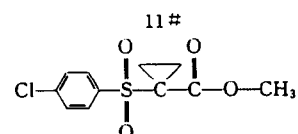
10 #



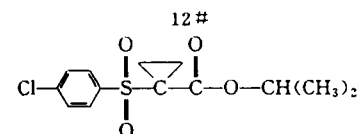
11 #



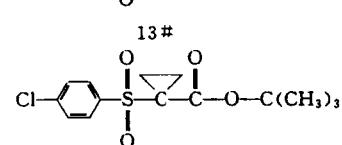
12 #



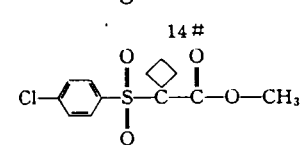
13 #



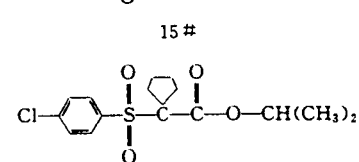
14 #



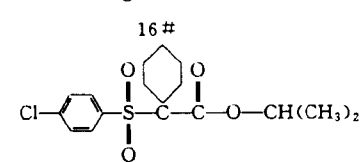
15 #



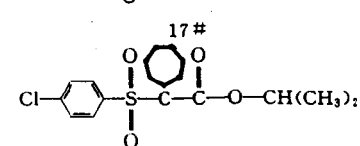
16 #



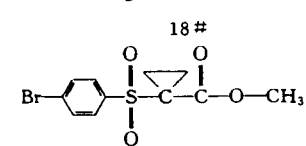
17 #



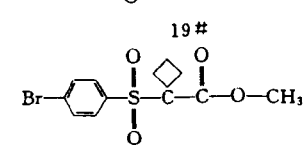
18 #



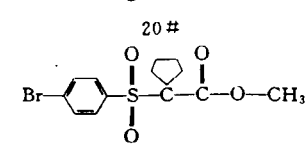
19 #



20 #



21 #



22 #

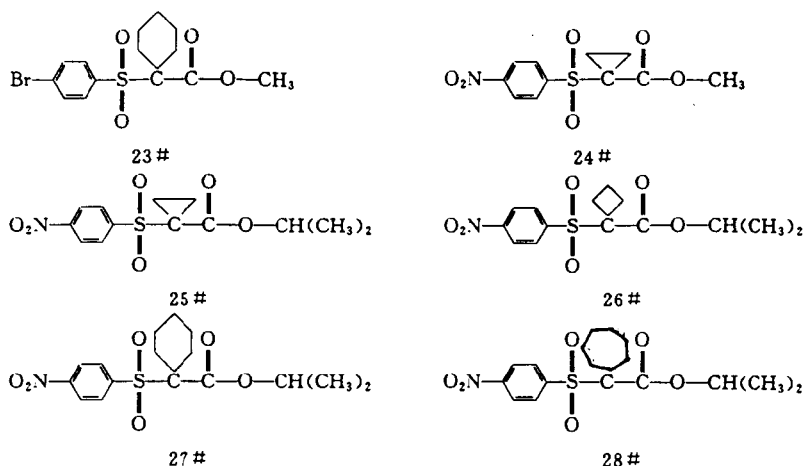


Figure 1-1 chemical structure of twenty-eight alkyl (1-phenylsulfonyl) cycloalkane-carboxylates

Table 1-1 toxicity and theoretical descriptors for 28 alkyl (1-phenylsulfonyl) cycloalkane-carboxylates *

No.	-lgEC ₅₀			-lgLC ₅₀			theoretical descriptors						
	exptl. equation (4) diff.			exptl. equation (5) diff.			α	μ	E_{homo}	E_{lumo}	qH^+	q^-	$V_i/100^*$
1	3.45	3.43	0.02	3.03	3.12	-0.09	114.0	1.780	10.179	0.762	0.172	0.930	1.186
2	3.62	3.60	0.02	3.18	3.22	-0.04	131.1	1.541	8.717	0.073	0.159	0.876	1.281
3	3.81	3.90	-0.09	3.39	3.56	-0.17	128.7	1.318	10.300	0.916	0.161	0.920	1.376
4	3.96	4.16	-0.20	3.73	3.80	-0.07	139.0	1.352	10.231	0.955	0.157	0.921	1.471
5	3.54	3.67	-0.13	3.21	3.36	-0.15	124.9	1.991	10.114	0.713	0.173	0.940	1.284
6	3.86	3.86	0.00	3.49	3.54	-0.05	133.2	1.944	10.085	0.682	0.174	0.942	1.382
7	4.09	4.08	0.01	3.72	3.74	-0.02	141.9	1.775	10.134	0.711	0.171	0.937	1.470
8	4.31	4.29	0.02	3.96	3.92	0.04	150.1	1.655	10.277	0.882	0.162	0.924	1.568
9	4.08	4.03	0.05	3.75	3.66	0.09	141.4	1.986	10.333	0.901	0.160	0.921	1.480
10	4.51	4.40	0.11	4.11	4.03	0.08	157.9	1.551	10.174	0.875	0.168	0.924	1.666
11	4.37	4.38	-0.01	4.01	4.00	0.01	149.9	1.500	10.219	0.920	0.156	0.922	1.578
12	4.72	4.73	-0.01	4.25	4.34	-0.09	166.9	1.448	10.176	0.876	0.158	0.924	1.764
13	3.77	3.66	0.11	3.43	3.34	0.09	122.8	1.806	10.425	1.046	0.170	0.934	1.276
14	3.95	3.99	-0.04	3.67	3.66	0.01	139.9	1.553	10.327	0.962	0.178	0.935	1.462
15	4.06	4.12	-0.06	3.78	3.79	-0.01	147.3	1.624	10.296	0.929	0.179	0.936	1.560
16	3.83	3.84	-0.01	3.59	3.49	0.10	132.1	1.543	10.499	1.167	0.166	0.917	1.374
17	4.28	4.44	-0.16	3.97	4.08	-0.11	156.2	1.215	10.361	1.089	0.166	0.921	1.658
18	4.52	4.59	-0.07	4.16	4.22	-0.06	165.6	1.134	10.377	1.117	0.172	0.921	1.756
19	4.93	4.71	0.22	4.40	4.33	0.07	174.0	1.344	10.312	1.075	0.173	0.922	1.854
20	3.89	3.75	0.14	3.54	3.42	0.12	126.0	1.737	10.404	1.110	0.170	0.934	1.317
21	4.02	3.90	0.12	3.70	3.55	0.15	133.9	1.486	10.473	1.220	0.166	0.918	1.415
22	4.21	4.15	0.06	3.83	3.81	0.02	141.0	0.919	10.415	1.191	0.167	0.917	1.513
23	4.47	4.40	0.07	4.17	4.04	0.13	151.7	1.042	10.371	1.201	0.163	0.918	1.611
24	3.60	3.63	-0.03	3.31	3.31	-0.00	132.1	1.987	10.875	1.837	0.182	0.928	1.326
25	3.92	3.86	0.06	3.60	3.53	0.07	148.7	2.182	10.808	1.776	0.190	0.931	1.526
26	3.97	4.08	-0.11	3.70	3.73	-0.03	157.1	1.877	10.894	1.860	0.184	0.918	1.610
27	4.34	4.45	-0.11	3.97	4.07	-0.10	174.5	1.780	10.876	1.871	0.182	0.913	1.806
28	4.75	4.72	0.03	4.36	4.35	0.01	180.9	1.232	10.940	1.914	0.183	0.917	1.904

* exptl. = experimental toxicity; equation (4) and equation (5) = toxicity calculated by equation (4) and equation (5) respectively; diff. = experimental toxicity - calculated toxicity; $V_i/100$ values are from reference^[23].

error of estimates (SE), the F value for analysis of variance (F), the significance level (p) and the variance inflation factor (VIF). The VIF is defined as $1/(1-R^2)$, where R is the multiple correlation coefficient of one independent variable against the others. "Large" VIF values imply strong correlation^[19]. The reported R^2 values were adjusted for degree of freedom (d. f.).

RESULTS AND DISCUSSION Many forms of toxicity are dependent on the pharmacokinetics and distribution of chemicals within the organism. Increasing water solubility (hydrophilicity) slows down the passage of chemicals through membranes and favors excretion of the solute from the organism, thus lessening the tendency to build up to toxic concentrations. In contrast, increasing lipophilicity (hydrophobicity) leads to easier passage through membranes and greater distribution into the organisms, thus resulting in greater toxicity^[1,8,9,10].

The LSER concept is a general approach to describe solvation and partitioning or related properties in diverse media. Within this approach there are three contributions to the total solute-solvent interactions^[8].

$$XYZ = XYZ_0 + \text{Cavity term} + \text{Dipolar term} + \text{Hydrogen bonding terms} \quad (1)$$

Where XYZ represents solubility or solvent-dependent properties (often expressed as the logarithm of measured properties). The cavity term measures the free energy or enthalpy input necessary to separate the solvent molecules, i. e., to overcome solvent-solvent cohesive interactions to provide a suitably sized cavity for the solute. The dipolar term measures the (typically) exoergic effects of solute-solvent dipole-dipole and dipole-induced dipole interactions. The hydrogen bonding terms measures the exoergic effects of complexation between hydrogen bond donor (HBD) solvents and hydrogen bond acceptor (HBA) solutes or vice versa.

Kamlet and coworkers^[7,8,9,10] used molecular volume and the so-called "solvatochromic parameters" to estimate the different interaction terms. These solvatochromic parameters are measured by different spectroscopic methods and are afflicted with bad availability and experimental errors. To be independent of experiments, Wilson, Famini and coworkers^[18,19,20] developed the TLSEER. In the TLSEER model, cavity term is described by the intrinsic molecular volume (V_i); the dipole term uses the polarizability index, π_i , obtained by dividing the polarizability (α) by the intrinsic molecular volume. The HBA basicity is composed of covalent, ϵ_b , and electrostatic, q^- , basicity terms, analogously, the HBD acidity is made up of ϵ_a , and electrostatic, qH^+ , acidity terms. Where ϵ_b is the magnitude of the difference between the energy of the highest occupied molecular orbital (E_{homo}) of the solute and the lowest unoccupied molecular orbital (E_{lumo}) of water, the result is divided by 100 for convenience and comparison of coefficients; ϵ_a is the magnitude of the difference between the E_{lumo} of the solute and the E_{homo} of water, again the result is divided by 100. q^- and qH^+ are the magnitude of the largest negative formal charge on an atom and the most positive formal charge on a hydrogen atom respectively.

In our previous study^[21], we found the correlation between intrinsic molecular volume ($V_i/100$) and polarizability (α) for these sulfur-containing compounds was significant. The obtained correlation equation was:

$$\alpha = 9.81 + 89.1V_i/100 \quad (2)$$

$$n=28, r=0.990, SE=2.5, F=1273.2, p=0.0000$$

Where n is the number of chemicals, r is the simple correlation coefficient.

Average molecular polarizability (α) measures the ability of the electrons in a molecule

to move easily as a result of a stimulus. According to the opinion of Kaliszan^[28], it is essentially a structurally nonspecific bulk property. For this reason, it is not appropriate to use $\pi_1 = \alpha/V_i$ as a measure of dipolar term in equation (1). For convenience of computation, α was used as a measure of cavity term in our MTLSE model. In addition, we proved it was reasonable to use dipole moment (μ) as a measure of dipolar term. The generalized MTLSE equation was:

$$XYZ = XYZ_0 + m\alpha + s\mu + a_1E_{\text{lumo}} + a_2qH^+ + b_1E_{\text{homo}} + b_2q^- \quad (3)$$

Where q^- is the most negative net atomic charge on an atom, qH^+ is the most positive net atomic charge on a hydrogen atom, m, s, a_1, a_2, b_1 and b_2 are regression coefficients. The coefficients and constant in equation (3) are to be determined by multiple regression analysis.

Based on the MTLSE, forward stepwise variable selection regression analysis of the $\lg EC_{50}$ and $\lg LC_{50}$ of the compounds under study versus the quantum chemical descriptors resulted in table 1-2. From table 1-2, we obtained the following equations:

$$-\lg EC_{50} = -3.350 + 0.0208\alpha - 0.226\mu - 12.0qH^+ + 7.39q^- \quad (4)$$

$$n=28, R^2=0.924, SE=0.105, F=83.0, p=0.0000$$

$$-\lg LC_{50} = -4.332 + 0.0196\alpha - 0.240\mu - 10.3qH^+ + 7.99q^- \quad (5)$$

$$n=28, R^2=0.933, SE=0.093, F=95.5, p=0.0000$$

Table 1-2 model fitting results for toxicity of alkyl (1-phenylsulfonyl) cycloalkane-carboxylates

independent variables	coefficient		SE		t-statistics		p	
	$-\lg EC_{50}$	$-\lg LC_{50}$	$-\lg EC_{50}$	$-\lg LC_{50}$	$-\lg EC_{50}$	$-\lg LC_{50}$	$-\lg EC_{50}$	$-\lg LC_{50}$
constant	-3.350	-4.332	1.643	1.454	-2.038	-2.9792	0.0532	0.0067
α	0.0208	0.0196	0.0014	0.0013	14.657	15.6005	0.0000	0.0000
μ	-0.226	-0.240	0.082	0.073	-2.746	-3.2913	0.0115	0.0032
qH^+	-12.0	-10.3	2.8	2.5	-4.278	-4.1813	0.0003	0.0004
q^-	7.39	7.99	1.85	1.63	4.001	4.8902	0.0006	0.0001

According to the principles of statistics, a regression equation is of no relevance when the explanatory variables applied were mutually interrelated by simple or multiple correlations. However, as it was shown by the correlation coefficients of the independent variables and the VIF values given by table 1-3, the largest simple correlation coefficient between the four independent variables is 0.4198, and the largest VIF value is 1.473, which corresponds with the multiple correlation coefficient of 0.3211, correlation coefficient tests showed all the correlation coefficients are not significant at significance levels $p \leq 0.05$, so equation (4) and equation (5) are interpretable.

The F and p values of equation (4) and equation (5) show the correlation of equation (4) and equation (5) are significant. In addition, equation (4) and equation (5) have great correlation coefficients and small SE, and the differences between the experimental toxicity and the predicted toxicity are very small (table 1-1 and figure 1-2), so equation (4) and equation (5) may be used to predict toxicity of similar series of compounds.

All the variables in equation (4) and equation (5) are significant ($p < 0.0115$). The importance of the descriptors is indicated by the sign and magnitude of the t statistics in table 1-2. Examination of table 1-2 leads to the following observations: (a) The average molecular polarizability (α) is the most statistically significant term influencing toxicity ($p = 0.0000$). The greater the polarizability (α), the greater is the toxicity. The polarizability is expected to be involved since it is in direct proportion to intrinsic molecular volume, and molecular vol-

ume is a measure of the energy needed to form a cavity in the solvent. The positive sign indicates that larger molecules tend to be in the less polar phase and be partitioned into the organism and result in greater toxicity. The fact is consistent with greater dispersive forces on larger molecules. (b) Increasing dipole moment (μ) leads to lower toxicity. This is reasonable since a greater μ would imply greater dipole-dipole and dipole-induced dipole interactions between solutes and more polar water, resulting a greater tendency to be partitioned into water and a lesser toxicity. (c) Increasing qH^+ leads to lower toxicity. This implies the hydrogen bonding interactions between the oxygen atoms in water molecules and the hydrogen atoms in the phenylthio-carboxylates. Thus the phenylthio-carboxylates with greater qH^+ tend to be partitioned into water and result in lower toxicity. (d) Increasing q^- , leads to greater toxicity. This implies the hydrogen bonding interactions between the compounds and the "target molecules" of *Daphnia magna straus*, with the "target molecules" accept electrons and the oxygen atoms in the sulfone group of these compounds provide electrons.

Table 1-3 correlation coefficient matrix for significant independent variables and the variance inflation factors (VIFs)

	correlation coefficients matrix				VIFs for equation (4) and equation (5)
	α	μ	qH^+	q^-	
α	1.0000				1.281
μ	-0.3541	1.0000			1.473
qH^+	0.2100	0.4198	1.0000		1.413
q^-	-0.2009	0.3781	0.3326	1.0000	1.223

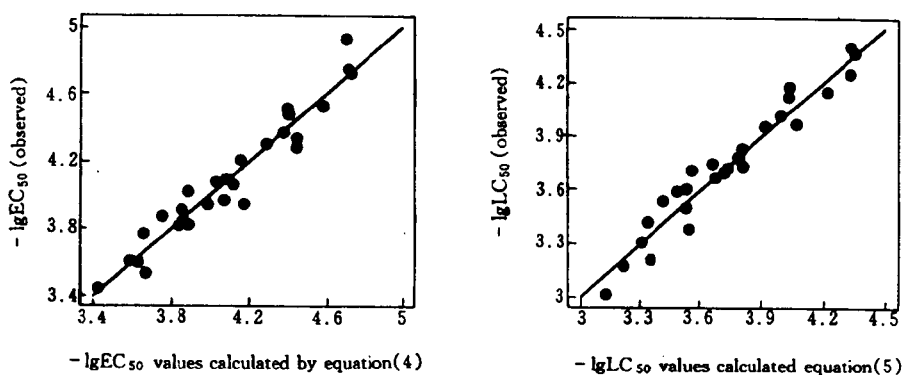


Figure 1-2 plot of observed toxicity vs. calculated toxicity

CONCLUSION The MTLSE model is successful in developing QSAR equations for the 28 alkyl (1-phenylsulfonyl) cycloalkane-carboxylates. Polarizability is the most significant term influencing toxicity. The toxicity increased with the increase of polarizability (α) and the most negative net atomic charges on oxygen atoms in the sulfone groups (q^-) of the compounds, decreased with the increase of dipole moment (μ) and the most positive net atomic charges on some a hydrogen atom (qH^+) of the compounds.

This study proved the advantages of the MTLSE model. (a) The quantum chemical descriptors in the model can be easily and precisely obtained by computation instead of experiment, a large amount of expenses and time can be saved; (b) The quantum chemical descriptors have clear physicochemical interpretations, and interpretation of the correlation equa-

tions can suggest modes of interaction between toxicants and organisms; (c) The MTLSE mode can be used for compounds for which the solvatochromic parameters are not available.

REFERENCES

1. D. J. W. Blum, and R. E. Speece, *Environ. Sci. Technol.*, **24**, 284 (1990).
2. H. J. M. Verhaar, C. J. Van Leeuwen, J. Bol, and J. L. M. Hermens, *SAR and QSAR in Environmental Research*, **2**, 39 (1993).
3. G. N. Burkhardt, *Nature (London)*, **17**, 584 (1935).
4. L. P. Hammett, *Chem. Rev.*, **17**, 125 (1935).
5. L. P. Hammett, *J. Am. Chem. Soc.*, **59**, 96 (1937).
6. O. Exner, in *Correlation Analysis of Chemical Data*, edited by J. Shorter, Plenum Press, New York (1988).
7. M. J. Kamlet, R. W. Taft, and J.-L. M. Abboud, *J. Am. Chem. Soc.*, **91**, 8325 (1977).
8. M. J. Kamlet, R. M. Doherty, J.-L. M. Abboud, M. H. Abraham, and R. W. Taft, *CHEMTECH*, **16**, 566 (1986).
9. M. J. Kamlet, R. M. Doherty, G. D. Veith, R. W. Taft, and M. H. Abraham, *Environ. Sci. Technol.*, **20**, 690 (1986).
10. M. J. Kamlet, R. M. Doherty, R. W. Taft, M. H. Abraham, G. D. Veith, and D. J. Abraham, *Environ. Sci. Technol.*, **21**, 149 (1987).
11. J. P. Hickey, D. R. Passino-Reader, *Environ. Sci. Technol.*, **25**, 1753 (1991).
12. D. F. V. Lewis, *J. Comput. Chem.*, **10**, 145 (1989).
13. Stefan Baj, and Malgorzata Dawid, *Journal of Liquid Chromatography*, **17**, 3933 (1994).
14. C. Hansch, and E. Coats, *J. Pharm. Sci.*, **59**, 731 (1970).
15. T. Nevalainen, and E. Kolehmainen, *Environ. Toxicol. Chem.*, **13**, 1699 (1994).
16. O. G. Mekenyan, G. T. Ankley, G. D. Veith, D. J. Call, *Chemosphere*, **28**, 567 (1994).
17. Lu Xu, J. W. Ball, S. L. Dixon, and P. C. Jurs, *Environ. Toxicol. Chem.*, **13**, 841 (1994).
18. L. Y. Wilson, G. R. Famini, *J. Med. Chem.*, **34**, 1688 (1991).
19. G. R. Famini, C. A. Penski, and L. Y. Wilson, *J. Phy. Org. Chem.*, **5**, 395 (1992).
20. A. D. Headley, S. D. Starnes, L. Y. Wilson, and G. R. Famini, *J. Org. Chem.*, **59**, 8084 (1994).
21. J. Chen, L. Feng, Y. Liao, S. Han, L. Wang, *Chemosphere*, **33**, 537 (1996).
22. S. N. Pandeya, T. N. Ojha, V. Srivastava, *J. Scient. Ind. Res.*, **44**, 150 (1985).
23. Y. He, L. Wang, Z. Liu, Z. Zhang, *Chemosphere*, **32**, 2739 (1995).
24. Tripos Associates, Inc. A molecular modeling system for the IBM PC, 1988.
25. M. J. S. Dewar, E. G. Zoebisch, E. F. Healy, and J. J. Stewart, *J. Am. Chem. Soc.*, **107**, 3902 (1985).
26. J. J. P. Stewart, *MOPAC MANUAL* (Sixth Edition), Frank J. Seiler Research Laboratory, U. S. Air Force Academy, CO 80840 (1990).
27. STSC, Inc. and Statistical Graphics Corporation, *STATGRAPHIC*, Version 4. 0, (Serial Number: 3017435), 1985.
28. R. Kaliszan, *Analytical Chemistry*, **64**, 619A (1992).

Using AM1 Hamiltonian in Quantitative Structure-Activity Relationship Studies of Phenylthio-carboxylates^①

ABSTRACT AM1 Hamiltonian contained in the MOPAC (6. 0) program package was used to compute descriptors for eighteen phenylthio-carboxylates. The Modified TLSE (MTLSE) model was used to develop a QSAR equation for toxicity of these compounds. The obtained equation can be used to predict toxicity to *Photobacterium Phosphoreum* of this series of compounds. The polarizability (α) increases the toxicity. The dipole moment (μ) and the most positive net atomic charge on a hydrogen atom (qH^+) decrease the toxicity. The study showed the advantages of the MTLSE descriptors: they permit near a *priori* prediction of toxicity; they can be easily and precisely obtained by computation instead of experiment, thus a large amount of expenses and time can be saved; and they have clear physicochemical interpretations, and interpretation of the correlation equations can suggest modes of interaction between toxicants and organisms.

① Toxicological and Environmental Chemistry, Vol. 60, pp. 211~221, 1997. Authors: JINGWEN CHEN, YANJUN ZHAO, LIU FENG, SHUOKUI HAN, LIANSHENG WANG and ZHENG ZHANG.

INTRODUCTION In the last several decades, phenylthio-carboxylates are used more extensively as intermediates in the manufacture of pesticides, herbicides and anthelmintic drugs^[1]. Accompanying a rapid advance in production, the toxicity effects caused by these compounds should be studied. However, there has been only limited investigation of their toxicity. Recently, eighteen phenylthio-carboxylates were synthesized in our laboratory. The acute toxicity of these compounds to *Photobacterium Phosphoreum* was reported^[1].

However, it is impossible to comprehensively test all existing chemicals to species of concern because of large expenditures of money and time. In addition, for the assessment of risks of exposure of these compounds, it is essential to make a prior prediction of toxicity of these compounds. One technique for the efficient development of toxicity data is Quantitative Structure-Activity Relationships (QSARs). QSARs correlate and predict toxicity of chemicals from their physical and chemical descriptors^[2].

Many forms of toxicity are dependent on the pharmacokinetics and distribution of chemicals within the organism. Increasing water solubility (hydrophilicity) slows the passage of chemicals through membranes and favors excretion of the solute from the organism, thus lessening the tendency to build up to toxic concentrations. Increasing lipophilicity (hydrophobicity) leads to easier passage through membranes and greater distribution into the organisms, thus resulting in greater toxicity^[3].

The linear solvation energy relationship (LSER) developed by Kamlet and coworkers^[4-6] has been successfully applied in developing predictive equations for toxicity of a wide range of environmental chemicals to *Photobacterium Phosphoreum*, *Daphnia magna*, *Daphnia pules*, and the fathead minnow (*Pimephales Promelas*)^[2,5,7,8]. However, the LSER methods are limited in their ability to make a prior prediction because the descriptors of these models, solvatochromic parameters, are either empirically determined with a time and money consuming course or only available for a finite number of chemicals. Even though there are tables of LSER parameters and predictive relationships to help in their estimation, solvatochromic parameters for complex molecules are not as easily estimated.

The descriptors derived from quantum computation have obvious advantages; they are not restricted to closely related compounds, can be easily obtained, and describe clearly defined molecular properties. For these reasons, there are quite a few more examples of the use of molecular orbital generated descriptors. Some examples were given by Lewis^[9], Stefan Baj et al.^[10], Hansch et al.^[11], Nevalainen et al.^[12], Mekenyan et al.^[13] and Lu Xu et al.^[14] Recently, Wilson, Famini and coworkers^[15-17] developed a theoretical set of LSER (TLSER) descriptors that are determined solely from computation. Their TLSER descriptors were developed so as to give optimum correlation with the LSER descriptors, to give TLSER equations with correlation coefficients and standard deviations close in value to those for LSER, and to be as widely applicable to solute-solvent interactions as the LSER set. In our previous study^[18], we found it inappropriate to use the TLSER model for developing prediction equations of twenty-eight alkyl (1-phenylsulfonyl) cycloalkane-carboxylates, so we deduced a modified TLSER (MTLSER) model. The MTLSER model was successful in developing predictive equations for octanol/water partition coefficients, aqueous solubility, adsorption coefficients for sediments of these compounds.

In short, it is the purpose of our present study to obtain a QSAR equation for the phenylthio-carboxylates on the basis of quantum chemical descriptors and the MTLSER model.

MATERIALS AND METHODS The structures and toxicity for the phenylthio-carboxylates are reproduced here in figure 1-3 and table 1-4, respectively, to aid discussion. The toxicity listed in table 1-4 was expressed as negative logarithm of EC_{50} , where EC_{50} is concentration values causing a 50% inhibition of bioluminescence after 15 minutes exposure.

The molecular modeling package ALCHEMY I^[19] was used to construct and view all molecular structures. Internal coordinate was used to write all molecular structures. Molecular geometry was optimized with the Broyden-Fletcher-Goldfarb-Shanno (BFGS) method^[20] and quantum descriptors were calculated using the AM1 Hamiltonian^[21] contained in the up-to-date version (Ver. 6.00) of the MOPAC program package^[22]. The AM1 algorithm was selected because it is a much more recent semi-empirical methods and gives good estimates of molecular energies and the computational time is much shorter than needed by *ab initio* methods. The MOPAC was run with the following keywords: AM1, PRECISE, ESP, DIPOLE, POLAR, NOINTER. The Keyword ESP (electrostatic potential calculation) was used to obtain ESP derived net atomic charges. All calculations were run on a 80586/75 MHz computer equipped with 16 megabytes of internal memory and supported by Disk Operation System (DOS). The descriptors computed by the MOPAC were listed in table 1-4. Units of the energy, charge, dipole and polarizability were electron volts (eV), atomic charge units (a. c. u.) and atomic units (a. u.) respectively.

