The Use of Computer Based Structure−Activity Relationships in the Risk Assessment of Industrial Chemicals

Walter Karcher*,† and Stoyan Karabunarliev†,‡

European Commission, DG XII-JRC, Environment Institute, European Chemicals Bureau, TP 280, I-21020 Ispra, Italy, and Department of Physics, Bourgas University of Technology, 8010 Bourgas, Bulgaria

Received September 14, 1995⊗

The concept of a two-step approach toward the assessment of toxicity endpoints for a chemical is proposed. The first step involves the selection of chemical analogues for which toxicity data is available in a noncongeneric database. The next step is the derivation of a Quantitative Structure−Activity Relationship (QSAR) for the chemical domain, predetermined by the selection rules. The software tools needed for the computer implementation of such an approach are summarized. By making use of them, we have derived aquatic toxicity QSARs, of which two are given as example. The latter pertain to chemicals that have been automatically extracted from noncongeneric databases, after defining the substructure recognition rules implied by the putative mechanism of toxicity.

1. INTRODUCTION

The European Commission has charged the European Chemicals Bureau with scientific and technical tasks related with the practical implementation of EU legislation on chemicals. An important issue[1] in this context is the assessment of environmental and human health hazards, arising from industrial chemicals in use within the countries of the EU. Under the Dangerous Substances Directive,[2] new chemicals have to be notified before marketing, accompanied by a basic data set (identification, physicochemical, toxic, and ecotoxic properties) in order to permit the assessment of the environmental and human health risks that are associated with their use. Similarly, the EU regulation on the risk assessment of existing chemicals[3] obliges producers and importers to provide the European Commission through the European Chemicals Bureau with the available data that is pertinent for the risk assessment process. The data submission procedure has commenced with the so-called High Production Volume (HPV) chemicals (production level > 1000 t/year). The database created for the purpose[4] contains at present entries for around 2500 chemicals. In the survey of data collection, transpire the inevitable data gaps for important biological endpoints. Under these circumstances, Quantitative Structure−Activity Relationships (QSAR) have found regular application as a complementary tool for assessment. Moreover, the application of QSAR has been dictated by the necessity to prepare priority rankings and priority lists.[5] The latter specify the chemicals of major concern for which eventually more complete data is required, even at the expense of more thorough testing.

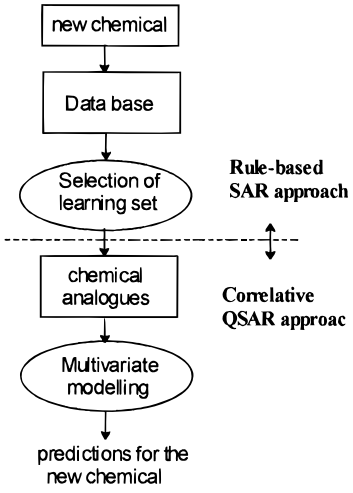
The practical application of QSAR[6] has become possible after research in the area has vastly expanded and has started to take the shape of a separate science.[7] Apart from the growing utilization of QSAR in pharmaceutical research and development, the increased concern of environmental hazards arising from chemical pollutants has opened another area of application. SAR research in environmental toxicology[8] bears specific features that are dictated by the theoretically unlimited number of chemicals which may represent potential environmental concern. Hence, much higher priority have acquired general chemical and toxicological problems about the discrimination and recognition of substances that invoke elevated or very elevated toxic effects.[9,10] It is evident that any large scale QSAR analysis of toxicity endpoints is only possible when categories of chemicals acting in a toxicologically similar manner are discerned.[11,12] The dichotomy of the problem has been long recognized in DNA-toxicity SAR research, where two approaches have been clearly outlined.[13] Rule-based approaches, comprising knowledge and expertise, aim at qualitative differentiation between noncongeneric chemicals. Classic correlative QSARs have limited range within congeneric series but provide more detailed and unbiased description of biological effects.

In the worst case, what is known of a compound for which predictions are requested is only its chemical structure. Indeed, all efforts for qualification of chemicals with respect to toxic action ultimately aim at rules[11,12] based exclusively on chemical structure. Recent progress in this direction has been made possible by the existence of representative collections[11] or large databases[14] of acute toxicity data that provide the background for such kind of analysis.[11,15] However, the prior qualitative relation of chemical structures to toxic action is still a challenging problem. Whereas chemical notion and nomenclature are based on frequently encountered functional groups, existing classification schemes suggest that trivial functional groups are not sufficient to provide discrimination of chemicals with respect to toxicity, even for standard endpoints like acute toxicity lethal doses ($LD_{\%}$) or concentrations ($LC_{\%}$). Apart from the specifically acting chemicals (e.g., pesticides, herbicides), the biological effect for a considerable number of the pronounced toxicants has been associated with the so-called[15a] reactive substructures that are not merely functional groups but their combinations or more complex structural patterns.[9−11] A number of reactive patterns that render a chemical intrinsically more toxic have been enumerated,[9−11] and in many cases the particular underlying molecular mechanism of toxicity has been suggested.[9,10] Nevertheless, the current

† European Chemicals Bureau.
‡ Bourgas University of Technology.
⊗ Abstract published in *Advance ACS Abstracts,* April 1, 1996.

S0095-2338(95)00130-2 CCC: $12.00

Scheme 1. Flow Chart of a Model QSAR Expert System



state of knowledge is such that, loosely speaking specific or reactive, part of the inert chemicals can be recognized as lacking chemical functions with a few exceptions, and part of the intrinsically more toxic chemicals can be identified by the presence of chemically reactive patterns.

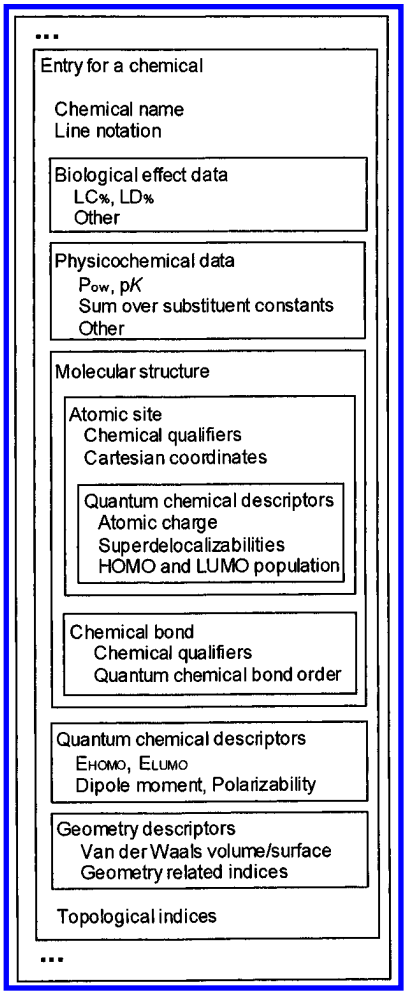
In existing practice, the environmental risk assessment for a "new" chemical starts usually with the identification of analogues, for which pertinent toxicologic and other data are available.[16] This approach clearly emphasizes the importance of large pools with chemical, physicochemical, and toxicologic data for noncongeneric compounds. However, as the number of compounds for which data are accumulated grows and provides more reference points in the "chemical space", the initial search for chemical analogues becomes less feasible for chemists and toxicologists, and the risk of missing important reference points increases. That is why this natural initial step in the risk assessment process is, as well, a necessary component of a model QSAR expert system. We have tentatively called this element "selection of a learning set" since the series of chemical analogues are most often the learning sets of correlative ecotoxicity QSARs.[17] We point out that this is the component of the flow Scheme 1 which essentially incorporates the features of chemical expertise and justifies the name "expert system". It bridges rule-based assessment schemes, derived from accumulated knowledge for large and heterogeneous collections of chemicals, to traditional correlative QSARs which usually operate with quite restricted homogenous series of congeners or other chemical analogues.

It is obvious in advance that the proposed general scheme for prediction may break at different points. Firstly, there might be no or not enough analogues of the "new" chemical in the database. Even if the appropriate learning set is large enough for further QSAR analysis, there is no guarantee that a mechanistic-consistent and statistically significant model can be derived. That is why any implementation is feasible only as the flexible and extendible computational environment that enhanced completion of the procedure.

2. SOFTWARE TOOLS FOR A DATA BASE INTEGRATED QSAR APPROACH

Next, we summarize the elements which, to our opinion, are indispensable for the implementation of such a computer system. Our experience is derived exclusively from QSAR modeling of aquatic toxicity in terms of LC_{50} to fish. The choice of this specific endpoint was dictated by (i) the existence of single source data collections[11,14] without methodological or interlaboratory variations and (ii) the considerable progress in the SAR research[8a] for this endpoint, in particular.

Chart 1. Structure of the OASIS Relational Database



As core for the implementation of a prototype, we use the OASIS system whose structure[18] and applications in ecotoxicity QSARs[19] have been described elsewhere. It incorporates the following more important elements that make it suitable for further upgrade.

1. Relational database with fields for molecular structure (chemical graph and atomic coordinates), observable properties, and theoretical descriptors. Theoretical descriptors pertain either to the molecule or to its distinct atomic sites (see Chart 1).

2. 3D molecular model builder[20] starting from the chemical graph in terms of line notation[21] or from 2D molecular input.

3. Calculation of theoretical descriptors related to molecular and electronic structure, including standard semiempirical quantum-chemicals methods.[22]

4. Variable selection technique within the multivariate regression approach. The module uses as input the database and enhances interpretation by immediate reference to the particular molecules.

5. Various visualization tools for numerical data plots, molecular structure, distribution of electron density, molecular orbitals, and site-specific electronic descriptors.

The OASIS system has been initially designed by Mekenyan and co-workers[18] as the software for standard cor-

relative QSAR studies. The latter usually start with input of the chemical graphs for the compounds under study and end up with the derivation of QSAR equation, if attainable. Due to the diversity of the computational methods typically required (e.g., those listed above), the system originally consisted of several modules whose input/output is practically a database that stores all pertinent data. Thus, OASIS comprises a number of software tools required for implementation of the model 1, but not those for the selection of a case-dependent learning set. That is why OASIS was augmented by several new elements and software tools which, to our opinion, will lie in the kernel of the learning set selection procedure.

6. Recognition of the largest common chemical subgraph for pairs of chemicals.

7. Chemical pattern recognition. A chemical pattern is a subgraph whose nodes are defined in terms of one or more allowed atomic sites or simple functional groups.

8. Flexible library of chemical patterns meant to store the known reactive molecular moieties.

9. Selection of chemicals from the database by chemical pattern recognition. The selection rule is given by a Boolean expression that consists of the occurrence values for an arbitrary set of chemical patterns.

10. Mapping of the atomic sites belonging to a chemical pattern that is common for a set of chemicals. That is needed to allow the use of site-specific descriptors in correlative QSAR.

Beyond doubt, software, as a bare computational background, is subordinate to the conceptual aspects of the problem. Hence, the software augmenting OASIS has been developed in parallel with QSAR studies, in which we tried to follow roughly the approach outlined, at least up to the automatic extraction of case-specific learning set from a heterogeneous database and the derivation of a QSAR with predefined range of validity in terms of rules that refer exclusively to chemical structure.

3. APPLICATIONS IN QSAR FOR ACUTE FISH TOXICITY

3.1. Sources of Toxicity Data, Methodology, and Background. We have used as databases two independent sources of fish toxicity data. The first one is the U.S.-EPA Environmental Research Laboratory-Duluth database[14] which is perhaps the most comprehensive and congruous aquatic toxicity data collection available. It contains entries for around 670 organic nonionic chemicals with LC_{50} values for fathead minnows (*Pimephales promelas*) from 96-h flow-through tests, performed under strict methodology. Another database was created using the collection of LC_{50} values to guppies (*Poecilia reticulata*) for 166 chemicals by Verhaar et al.[11] Both data set contained, in addition to the toxicity data, estimates of the octanol−water partition coefficient of the chemicals, P_{ow}. For the sake of compliance with prior QSAR modeling[11,15] of these toxicity data collections, the same $\log(P_{ow})$ estimates were used herein.

The conversion of the initial chemical data into OASIS file format and the further data processing were performed in three steps:

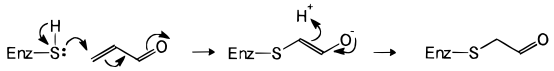
1. Conversion of the chemical line notations[21] for the compounds into internal chemical graph representations.

2. Generation of initial 3D molecular structures[23] in terms of Cartesian coordinates.

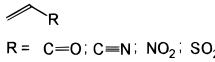
3. Complete quantum-chemical geometry optimization of molecular structures and calculation of electronic descriptors, employing the AM1 semiempirical all-valence electron Hamiltonian.[24]

The keystone of aquatic toxicity QSAR is the so-called baseline toxicity model[25,26] which pertains to the chemicals that act through the nonspecific mechanism of narcosis[27] and relates the toxic potency of the compounds exclusively to their lipophilicity. The baseline model is a univariate $\log(LC_{50}) - \log(P_{ow})$ relationship that predicts also the minimal toxicity a nonelectrolyte chemical may invoke. Hence, the challenging problem is to identify and assess the intrinsically more toxic chemicals which, apparently, constitute a considerable part of the existing ones.[11,15a] In view of what is known[9,10] of the underlying molecular mechanisms of toxicity, some manifest toxicants can directly bind to nucleophilic sites that are ubiquitous in biological macromolecules. Other chemicals exhibit toxicity more elevated than baseline, but their toxic action is not associated with a particular biological target, e.g., nonspecifically acting chemicals like polar narcotics[28] and uncouplers of oxidative phosphorylation.[29] In the QSAR studies, we concentrated on both categories of chemicals, namely three different series of direct electrophiles[30a] and a group of benzene derivatives[30b] which did not possess inherently reactive substituents. To outline series of toxicodynamically similar direct electrophiles, we proceeded from the mechanistic hypotheses about the specific molecular reactions[10] determining elevated toxic effect. Hereafter, the results for only one group of such electrophiles are summarized.

3.2. QSAR for the Toxicity of Michael-Type Acceptors. Michael-type acceptors are assumed[9,10] to bind to nucleophiles via addition to an activated carbon−carbon double bond (C═C). As activating moiety serves a polar group R (C═O in the scheme below) across a single bond:



Most prominent examples of such toxicants are acrolein, acrylamide, and α-naphthoquinone,[10] but other polarizable substituents such as nitrile, nitro, and sulfone may also enable the reaction.[9,11] In the screening, we extracted from the ERL-Duluth database the compounds possessing the chemical substructure[9,11]



with the additional requirements that the active carbon site at the β position to the polar group R has at least one attached hydrogen and the α,β carbon−carbon double bond does not participate in a stable six π-electron aromatic ring that precludes addition. The correlation sample generated involved 18 chemicals for which toxicity tests have been performed. No chemicals with polar group other than R = C═O were encountered. The toxicity of these probable Michael-type acceptors, mostly acrylates, is definitely not determined only by partitioning since $\log(P_{ow})$ explains 12% of the variance of $\log(LC_{50})$ (see Figure 1a). The pool of electronic descriptors considered consisted of atomic charges, superdelocalizability indices,[31] and bond orders[32] pertaining to the common reactive chemical pattern. Among them several correlated already better with $\log(LC_{50})$ than log-

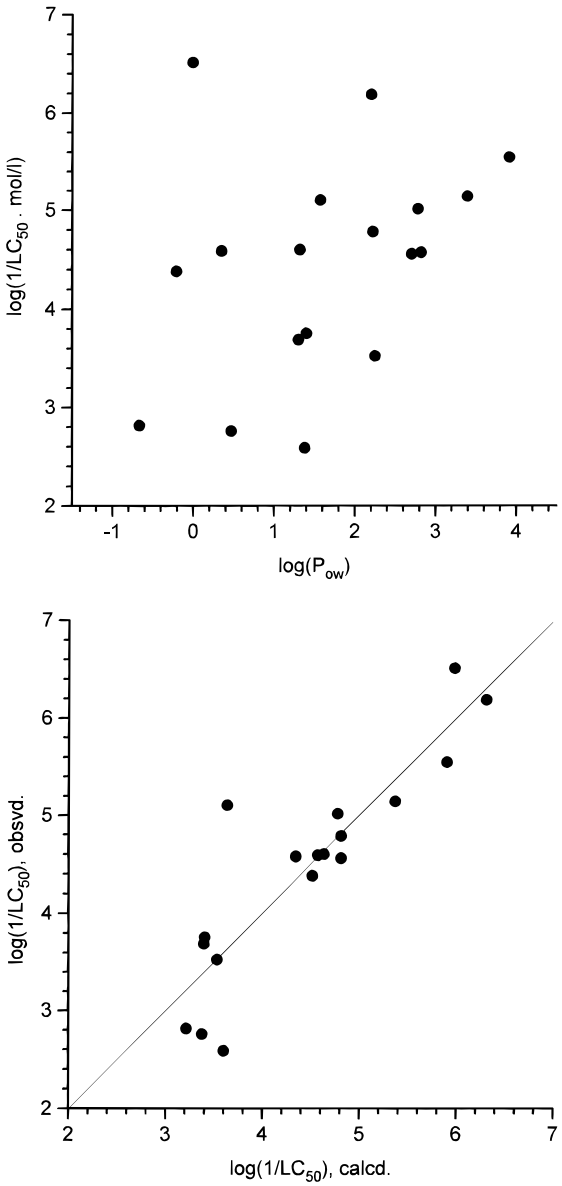


Figure 1. Michael-type acceptors. (a) Plot of $\log(1/LC_{50})$ vs $\log(P_{ow})$. (b) Plot of $\log(1/LC_{50})$ observed vs $\log(1/LC_{50})$ calculated according eq 1. LC_{50} in mol/L.

(P_{ow}). Starting from the $\log(1/LC_{50}) - \log(P_{ow})$ linear equation and allowing for addition of any next independent descriptor at 70% confidence level, the following three-factor QSAR (see also Figure 1b) was obtained:

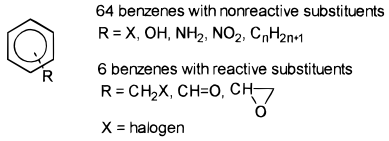
$$\log(1/LC_{50}) = 28.638(\pm 0.525)A_R + 81.330(\pm 0.306)B_{\alpha\text{-}R} + 0.359(\pm 0.142)\log(P_{ow}) - 89.090(\pm 0.290)$$

$$n = 18, \quad R^2 = 0.78, \quad s^2 = 0.33, \quad F = 16.49, \quad Q^2 = 0.43 \quad (1)$$

Apart from its acceptable statistical significance, the QSAR (1) is consistent with the proposed reaction mechanism. Namely, A_R is the acceptor superdelocalizability for the carbon of the polar group to which the electron transfer occurs, and $B_{\alpha\text{-}R}$ is the bond order for the bridging single bond that acquires double bond character at the intermediate step of the reaction (see scheme).

3.3. QSAR for the Toxicity of Substituted Benzenes. Substituted benzenes, including anilines and phenols, are a significant portion of the chemicals that are of environmental concern. Their toxic potency may vary, however, in a wide range. According to Verhaar et al.,[11] representatives of this category belong to all toxicity classes, namely inert chemicals, less inert chemicals, reactive chemicals, and specifically acting chemicals. Similarly, according to Fish Acute Toxicity Syndromes[33] for fathead minnows, substituted benzenes can act as nonpolar narcotics, polar narcotics, uncouplers of oxidative phosphorylation, or even invoke reactive mode of toxic action. On the other hand, earlier QSAR studies[19a−c] showed that, for a large set of nonspecifically acting aromatics (so-called soft electrophiles), the variation of toxic potency is related to chemical electrophilicity in terms of the average acceptor superdelocalizability for the π-conjugated moiety. Hence, we proceeded from the mechanistic assumption that there is an approximate distinction between toxic action associated with the perturbed π-electronic structure of the benzene ring and excess toxicity due to inherently reactive substituents that are known to give rise to elevated toxic effect also for nonaromatic compounds. To check this hypothesis, we used both the fathead minnow and guppy toxicity data sets and obtained very similar results.[30b] Herein, the QSAR derived for the acute toxicity of substituted benzenes to guppies by using the original $\log(P_{ow})$ estimates[11] for the chemicals of the data collection is summarized.

Among the benzene derivatives of the guppy collection, 64 chemicals do not possess any known inherently reactive substituent and are consequently included into the correlation sample. The rest of the chemicals of the data set with a single benzene ring is comprized of 12 specifically acting phosphorothionates[17b] and six benzenes possessing inherently reactive substituents,[9−11] such as α-unsaturated halide, aldehyde, and epoxy functions:



The latter six benzenes which probably act as direct electrophiles due to reactive substituents are further considered solely to illustrate that they lie out of the domain of the model which takes into account only electronic descriptors that pertain to the common benzene ring. For the actual correlation sample of 64 benzene derivatives, $\log(P_{ow})$ explains 54% of the variance of $\log(LC_{50})$ (see Figure 2a) which is in line with the presence of chemicals from all four toxicity classes.[11] The descriptor pool of quantum-chemical quantities included average and maximal acceptor and donor superdelocalizabilities for the common benzene ring. Stepwise variable selection starting from the linear $\log(LC_{50}) - \log(P_{ow})$ equation unambiguously discloses the importance of complementary electronic factors. In a highly significant QSAR, the maximal acceptor superdelocalizabily, A_{max}, and

$$\log(1/LC_{50}) = 0.511(\pm 0.022)\log(P_{ow}) + 24.006(\pm 0.230)A_{max} + 38.420(\pm 0.351)D_{max} - 11.754(\pm 0.068)$$

$$n = 64, \quad R^2 = 0.874, \quad s^2 = 0.075, \quad F = 138.21, \quad Q^2 = 0.722 \quad (2)$$

the maximal donor superdelocalizability, D_{max}, are sequentially accepted at 99% confidence level with respect to the

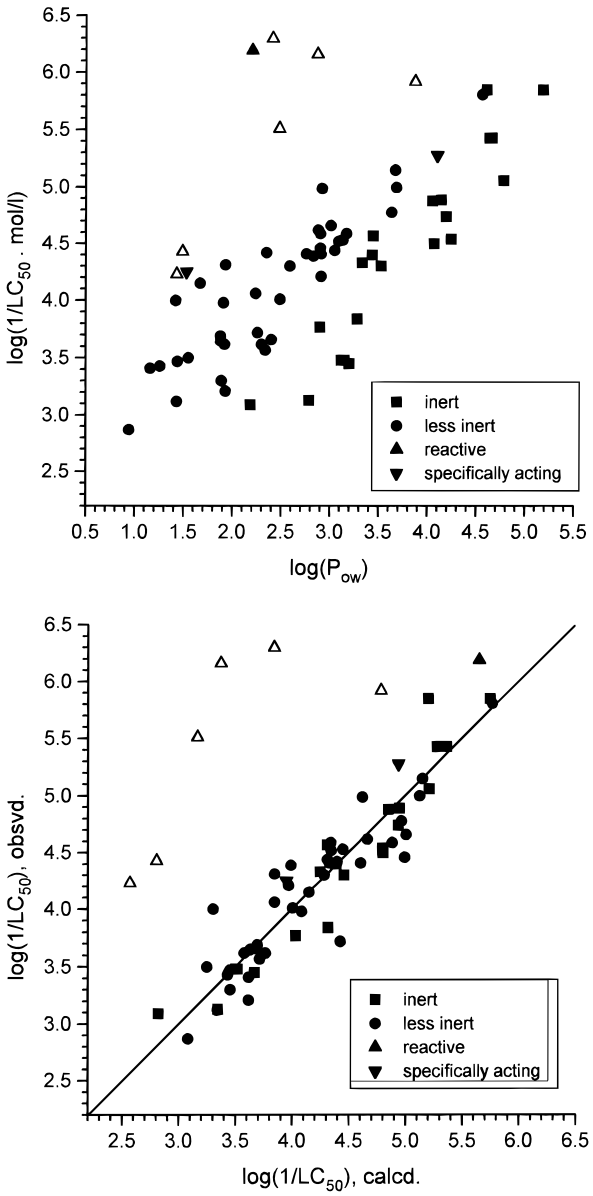


Figure 2. Benzenes derivatives included in the correlation sample—solid tags, and benzene derivatives with reactive substituents—empty tags. The classification is according Verhaar et al.[11] (a) Plot of $\log(1/LC_{50})$ vs $\log(P_{ow})$. (b) Plot of $\log(1/LC_{50})$ observed vs $\log(1/LC_{50})$ calculated according eq 2. LC_{50} in mol/L.

preceding models with one and two independent descriptors, respectively. Actually, in the stepwise variable selection, each additional variable explains ca. 50% of the unexplained variance of the preceding multivariate regression. The QSAR (2) employing electronic descriptors in addition to $\log(P_{ow})$ provides several advantages with respect to the generic $\log(LC_{50}) - \log(P_{ow})$ linear relationship (compare Figure 2 (parts b and a)). Substituted benzenes from the correlation sample (solid tags) which otherwise considerably deviate from the baseline[11] (less inert, reactive and specifically acting) now fall almost equally well into the assessment range of the model. In line with the assumption that the significant electrophilic function is outside of the aromatic ring, benzenes with reactive substituents (empty tags) exhibit considerably more elevated toxic potencies than predicted from QSAR (2). In general, the latter equation provides better segregation between benzenes with and without reactive substituents than the $\log(LC_{50}) - \log(P_{ow})$ linear relationship.

4. CONCLUSIONS AND OUTLOOK

Existing database linked QSARs for aquatic toxicity use comprehensive, but fixed classification schemes, in combination with a univariate toxicity/octanol-water partition coefficient relation for each particular class. We here envisage another, more flexible computer aided approach for QSAR assessment because of the following: Strict division of chemicals is hardly possible or rational since classes inevitably overlap[34] and the qualitative variation of toxic action does not necessarily parallel the chemical notion of different classes. For groups of chemical analogues that are more toxic, presumably because of the presence of a reactive chemical pattern, $\log(P_{ow})$ is often insufficient to account for the toxicity variance. In many cases it is possible to select or derive complementary quantities related to electronic structure that, in combination with $\log(P_{ow})$, provide (i) less approximate quantitative description of toxicity endpoints, (ii) improved level of statistical significance and (iii) better insight into the particular toxic mechanism.

The notion of reactive chemical patterns has already granted rationale[9−11] for the explanation of qualitative and quantitative inter-chemical variations of acute aquatic toxicities. Hence, it is the appropriate conceptual background for the selection of case-specific learning sets of known chemicals. A more elaborate approach[35] that relates a single chemical to a multitude of chemicals on the basis of a set chemical subgraphs has been developed. However, there are several knowledge-based prerequisites for its use for the purpose discussed. Namely, a more complete and detailed enumeration of chemical patterns that are relevant for toxicity is required. Some reactive functions give rise to much more pronounced excess toxicity than others. Moreover, the toxic effect of chemicals is most probably a cumulative result of different biological interactions. Hence, in many cases the straightforward assignment of a single toxicity relevant moiety is excluded.

ACKNOWLEDGMENT

One of us (S.K.) acknowledges a visiting scientist grant from the European Commission. We thank Dr. C. L. Russom, Dr. S. P. Bradbury and Prof. G. D. Veith (U.S.-EPA Environmental Research Laboratory − Duluth) for providing us with the database of acute toxicities of organic chemicals to fathead minnows, Dr. J. Sadowski and Prof. J. Gasteiger (Universität Nürnberg-Erlangen) for a temporary license for the 3D molecular model builder CORINA, and Prof. O. G. Mekenyan (Bourgas University of Technology) for helpful discussions.

Supporting Information Available: Tables of the chemicals used as learning sets, the toxicity data, and the descriptor variables of QSARs (1) and (2) (3 pages). This material is contained in many libraries on microfiche, immediately follows this article in the microfilm version of the journal, can be ordered from the ACS, and can be downloaded from the Internet; see any current masthead page for ordering information and Internet access instructions.

REFERENCES AND NOTES

(1) Karcher, W. The Role of the European Chemicals Bureau in the Risk Assessment of Chemicals. *4th Annual SETAC-Europe Conference*; Brussels, April 1994.
(2) Directive 67/548/EEC Dangerous Substances, O.J.L. 196, 16 August 1967, p 1; O.J.L. 259, Vol. 22, 15 October 1979, p 10.
(3) EU-Council Regulation on the Risk Assessment of Existing Chemicals No.793/93 (1993), O.J.L. 84/36, 5 April 1993.

(4) Grein, K.; Heidorn, Ch.; Karcher, W.; McCutcheon, P.; Murphy, P. The EUCLID-Database Industry Data for Risk Assessment, *207th ACS National Meeting*, San Diego, March 1994.
(5) van der Zandt, P. T. J.; van Leuwen, C. J. A Proposal for Priority Setting of Existing Chemical Substances (prepared for the European Commission Directorate General for Environment, Nuclear Safety and Civil Protection), *VROM*, Directorate for Chemicals, External Safety and Radiation Protection: The Hague, 1992.
(6) The area of ecotoxicity applications has been covered in: *Practical Applications of Quantitative Structure−Activity Relationships in Environmental Chemistry and Toxicology*; Karcher, W.; Devillers, J., Eds.; Kluwer: Dordrecht, 1990.
(7) Hansch, C. Quantitative Structure−Activity Relationships and the Unnamed Science. *Acc. Chem. Res*. **1993**, *26*, 147−153.
(8) An excellent review on acute toxicity QSAR has appeared recently: (a) Cronin, M. T. D.; Dearden, J. C. QSAR in Toxicology. 1. Prediction of Aquatic Toxicity, *Quant. Struc.-Act. Relat*. **1995**, *14*, 1−7. (b) Cronin, M. T. D.; Dearden, J. C. QSAR in Toxicology. 2. Prediction of Acute Mammalian Toxicity and Interspecies Correlation. *Quant. Struc.-Act. Relat*. **1995**, *14*, 117−120.
(9) Hermens, J. L. M. Electrophiles and Acute Toxicity to Fish. *Environ. Health Perspect*. **1990**, *87*, 219−225.
(10) Lipnick, R. L. Outliers: Their Origin and Use in the Classification of Molecular Mechanisms of Toxicity. *Sci. Total Environ*. **1991**, *109/110*, 131−153.
(11) Verhaar, H. J. M.; van Leeuwen, C. J.; Hermens, J. L. M. Classifying Environmental Pollutants. 1. Structure−Activity Relationships for Prediction of Aquatic Toxicity, *Chemosphere* **1992**, *25*, 471−491.
(12) Russom, C. L.; Bradbury, S. P.; Broderius, S. J.; Drummond, R. A.; Hammermeister, D. E. Predicting Modes of Toxic Action from Chemical Structure: Acute Toxicity in Fathead Minnow (*Pimephales Promelas*). Part I − Development of a Rule-Based Expert System. *Environ. Toxicol. Chem*. (submitted).
(13) (a) Richard, A. M. Application of SAR Methods to Noncongeneric Data Bases Associated with Carcinogenicity and Mutagenicity: Issues and Approaches, *Mutat. Res*. **1994**, *305*, 73−97. (b) Richard, A. M.; Rabinowitz, J. R.; Waters, M. D. Strategies for the Use of Computational SAR Methods in Assessing Genotoxicity. *Mutat. Res*. **1989**, *221*, 181−196.
(14) Center for Lake Superior Environmental Studies, Acute Toxicities of Organic Chemicals to Fathead Minnows (*Pimephales promelas*); Geiger, D. L., Brooke, L. T., Call, D. J., Eds.; University of Wisconsin-Superior: Superior, WI, 1990; Vol. 5, also Vols. 1−4.
(15) (a) Jäckel, H.; Nendza, M. Reactive Substructures in the Prediction of Aquatic Toxicity Data. *Aquatic Toxicol*. **1994**, *29*, 305−314. (b) Nendza, M.; Russom, C. L. QSAR Modelling of the ERL-D Fathead Minnow Acute Toxicity Database. *Xenobiotica* **1991**, *2*, 147−170.
(16) (a) Auer, C. M.; Nabholz, J. V.; Baetcke, K. P. Mode of Action and the Assessment of Chemical Hazards in the Presence of Limited Data: Use of Structure−Activity Relationships (SAR) under TSCA, Section 5. *Environ. Health. Perspect*. **1990**, *87*, 183−197. (b) Auer, C. M.; Zeeman, M.; Nabholz, J. V.; Clements, R. G. SAR − The U.S. Regulatory Perspective (Review). *SAR QSAR Environ. Res*. **1994**, *2*, 29−38.
(17) (a) Mekenyan, O. G.; Veith, G. D.; Bradbury, S. P.; Russom, C. L. Structure-Toxicity Relationship for α,β-unsaturated Alcohols in Fish. *Quant. Struc.-Act. Relat*. **1993**, *12*, 132−136. (b) Verhaar, H. J. M., Erikson, L.; Sjöström, M., Schüürmann, G.; Seinen, W.; Hermens, J. L. M. Modelling the Toxicity of Organophosphates. *Quant. Struc.-Act. Relat*. **1994**, *13*, 133−143. (c) Erikson, L.; Verhaar, H. J. M.; Sjöström, M.; Hermens, J. L. M. Multivariate Characterization and Modelling of the Chemical Reactivity of Epoxides. *Quant.Struc.-Act. Relat*. **1993**, *12*, 357−366. (d) Deneer, J. W.; Seinen, W.; Hermens, J. L. M. The Acute Toxicity of Aldehydes to the Guppy. *Aquatic Toxicol*. **1988**, *2*, 185−192.
(18) (a) Mekenyan, O. G.; Karabunarliev, S.; Bonchev, D. The Microcomputer OASIS System for Predicting of the Biological Activity of Chemical Compounds. *Comput. Chem*. **1990**, *14*, 193−200. (b) Mekenyan, O. G.; Karabunarliev, S.; Ivanov. J.; Dimitrov D. A New Development of the OASIS Computer System for Modelling of Molecular Properties, *Comput. Chem*. **1994**, *18*, 173−187.
(19) (a) Veith, G. D.; Mekenyan, O. G. A QSAR Approach for Estimating the Aquatic Toxicity of Soft Electrophiles. *Quant. Struct.-Act. Relat*. **1993**, *12*, 349−356. (b) Mekenyan, O. G.; Veith, G. D. Relationships Between Descriptors for Hydrophobicity and Soft Electrophilicity in Predicting Toxicity. *SAR QSAR Environ. Res*. **1994**, *1*, 335−344. (c) Mekenyan, O. G.; Veith, G. D. The Electronic Factor in QSAR: MO-Parameters, Competing Interactions, Reactivity and Toxicity. *SAR QSAR Environ. Res*. **1994**, *2*, 129−143. (d) Mekenyan, O. G.; Ankley, G. T.; Veith, G. D.; Call, D. J. QSAR For Photo-Induced Toxicity: I. Acute Lethality of PAHs to Daphnia Magna. *Chemosphere* **1994**, *28*, 567−582. (e) Mekenyan, O. G.; Ankley, G. T.; Veith, G. D.; Call, D. J. The Relation Between Ground and Excited Molecular States of PAHs and Their Photo-induced Toxicity to Daphnia Magna. *SAR QSAR Environ. Res*. **1994**, *2*, 237−247. (f) Mekenyan, O. G.; Ankley, G. T.; Veith, G. D.; Call, D. J. A QSAR Analysis of the Photoinduced Toxicity of Thiophenes, Trinitrotoluenes and Related Compounds. *Environ. Sci. Technol*. **1995**, *29*, 1267−1272.
(20) Ivanov, J.; Karabunarliev, S.; Mekenyan, O. G. 3DGEN: A System for Exhaustive 3D Molecular Design Proceeding from Molecular Topology. *J. Chem. Inf. Comput. Sci*. **1994**, *34*, 234−243.
(21) Two chemical line notations are supported: a) Karabunarliev, S.; Ivanov, J.; Mekenyan, O. G. Coding of Chemical Structures Based on Line Notation, *Comput. Chem*. **1994**, *18*, 189−193. (b) Weininger, D. SMILES, a Chemical Language and Information System. 1. Introduction to Methodology and Encoding Rules. *J. Chem. Inf. Comput. Sci*. **1988**, *28*, 31−36.
(22) Stewart, J. J. P. MOPAC: A General Molecular Orbital Package (Version 7), Q.C.P.E. no. 455, 1995. The program has been augmented with calculation of site-specific electronic descriptors like superdelelocalizability indices and polarizabilities.
(23) Sadowski, J.; Gasteiger, J. From Atoms and Bonds to Three-dimensional Atomic Coordinates: Automatic Model Builders. *Chem. Rev*. **1993**, *93*, 2567−2581. The 3D molecular model builder CORINA has been used for this particular study. By its beginning, the OASIS 3D molecular builder[20] was only practical for smaller series since it generated exhaustive sets of conformers for each chemical.
(24) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. AM1: A New General Purpose Quantum Mechanical Molecular Model. *J. Am. Chem. Soc*. **1985**, *107*, 3902−3909.
(25) (a) Könemann, W. H. QSAR in Fish Toxicity Studies. Part 1. Relationship for 50 Industrial Pollutants, *Toxicology* **1981**, *19*, 209−221. (b) Veith, G. D.; Call, D. J.; Brooke, L. T. Structure-Toxicity Relationships for the Fathead Minnow, Pimephales promelas: Narcotic Industrial Chemicals. *Can. J. Fish. Aquat. Sci*. **1983**, *40*, 743−748.
(26) Lipnick, R. L. Narcosis: Fundamental and Baseline Toxicity Mechanism for nonelectrolyte organic chemicals. In *Practical Applications of QSARs in Environmental Chemistry and Toxicology*; Karcher, W., Devillers, J., Eds.; Kluwer: Dordrecht, 1990; pp 281−294.
(27) Franks, N. R.; Lieb, W. R. Mechanisms of General Anesthesia. *Environ. Health. Perspect*. **1990**, *87*, 199−205.
(28) Veith, G. D.; Broderius, S. J. Rules for Distinguishing Toxicants that Cause Type I and Type II Narcosis Syndromes. *Environ. Health Perspect*. **1990**, *87*, 207−211.
(29) Terada, H. Uncouplers of Oxidative Phosphorylation, *Environ. Health Perspect*. **1990**, *87*, 213−218.
(30) (a) Karabunarliev, S.; Mekenyan, O. G.; Karcher, W.; Russom, C. L.; Bradbury, S. P. Quantum-chemical Descriptors for Estimating the Acute Toxicity of Electrophiles to the Fathead Minnow: An Analysis Based on Molecular Mechanisms. *Quant. Struc.-Act. Relat*. (accepted). (b) Karabunarliev, S.; Mekenyan, O. G.; Karcher, W.; Russom, C. L.; Bradbury, S. P. Quantum-chemical Descriptors for Estimating the Acute Toxicity of Substituted Benzenes to the Guppy and Fathead Minnow. *Quant. Struc.-Act. Relat*. (accepted).
(31) (a) Fukui, K.; Yonezawa, T.; Nagata, C. Theory of Substitution in Conjugated Molecules. *Bul. Chem. Soc. Jpn*. **1954**, *27*, 423−427. A generalization for superdelocalizability indices beyond π-electronic approximation is given: (b) Schüürmann, G. QSARs for the Polarizability, Solvatochromic Parameters and Lipophilicity. *Quant. Struc.-Act. Relat*. **1990**, *9*, 326−333.
(32) Armstrong, D. R.; Perkins, P. G.; Stewart, J. J. P. Bond Indices and Valency. *J. Chem. Soc., Dalton Trans*. **1973**, 838.
(33) Bradbury, S. P.; Henry, T. R.; Carlson, R. W. Fish Acute Toxicity Syndromes in the Development of Mechanism-Specific QSARs. In *Practical Applications of QSARs in Environmental Chemistry and Toxicology*; Karcher, W., Devillers, J., Eds.; Kluwer: Dordrecht, 1990; pp 295−316.
(34) Brüggemann, R.; Altschuh, J.; Matthies, M. QSAR for Estimating Physicochemical Data. In *Practical Applications of QSARs in Environmental Chemistry and Toxicology*; Karcher, W., Devillers, J., Eds.; Kluwer: Dordrecht, 1990; pp 197−212.
(35) Downs, G. M.; Willett, P. The Use of Similarity and Clustering Techniques for the Prediction of Molecular Properties. *In Applied Multivariate Analysis in SAR and Environmental Studies*; Karcher, W., Devillers, J., Eds.; Kluwer: Dordrecht, 1991; pp 247−380, and references therein.

CI9501305