

EXPERIMENTAL TOXICOLOGY *IN SILICO*

Miloň Tichý

National Institute of Public Health, Šrobárova 48, 10042 Praha 10
e-mail: mtichy@szu.cz

Received: June 10, 2005; Accepted: September 25, 2005

Key words: Toxicity Testing/*In Silico* Methods/QSAR/Validation of Computer Models

A review on applicability of *in silico* methods for toxicity testing by calculation for regulatory purposes, their survey and background of QSAR (Quantitative Structure-Activity Relationships) analysis are presented.

INTRODUCTION

The terms "experimental toxicology" and "methods *in silico*" jointly in one sentence may look like absurd conjunction. Experimental toxicology operates with experimental animals, cells, tissues, organs, with objects of living nature; methods *in silico* with computers. In spite of that the conjunction is not absurd. The *in silico* methods belong to the family of alternative methods in toxicology, which are especially suitable for testing toxicity of chemicals - to identify their hazard. These methods are not in our countries sufficiently exploited, the knowledge is of general nature and not always positive. They live to experience of a priority interest these days from several reasons.

The main reason for the increasing interest is a program of chemical safety, accepted by European Committee in 2003 called REACH, meaning "Registration, Evaluation and Authorization of Chemicals". The most important topic is that both "new" and "old" chemicals are involved. It means the demand to put forward materials on the hazard of both chemicals newly being distributed among people and in nature and those already used alone or in chemical preparations without any restriction for years. The aim is to improve environmental and human health and to strengthen competition ability of chemical industry in Europe. The program is a result of decades to find a legislative frame related to introduction of new and new chemicals and chemical preparations to the market, among people and to environment. The alternative methods of testing toxic and adverse effects of chemicals should make it possible to fulfil demands of REACH and regulatory norms of all states about categorization and registration of chemicals and chemical preparations and should be cheaper and faster than the classical ones. The alternative methods also should decrease a number of vertebrates, used for the testing according to accepted regulations and, thus, to fulfil also the demand of 3R', ie. Replace, Reduce and Refine the application of vertebrates.

METHODS

Citations of original papers and more details on history, development and procedures relating to QSAR models can be found elsewhere^{1,2}.

The experimental data on toxicity of chemicals obtained during centuries in experiments with laboratory animals have been possible to collect, to validate as much as possible and to generalize. This task in modern way has been succeeded by techniques of QSAR (Quantitative Structure - Activity Relationships) analysis to develop QSAR models. The modern time of these activities starts in fifties of the last century. Thus, the data with information adequate to that from direct experiments can be obtained; quicker, cheaper and saving experimental animals. The models are constructed for personal computers using methods of mathematical statistics. The models then generate data like the animals together with all statistical evaluation of their correctness, reproducibility or fit with the experiment General "category" of this and similar methods obtained a label *in silico* methods, meaning to obtain the prediction by computer calculation. Therefore we can easily use the idiom "experimental toxicology *in silico*".

Besides other types of *in silico* methods exist: (i) Physiologically-Based Kinetic Simulation Models (PBKM), (ii) biological similarity and allometric equations, (iii) expert systems based on basis of knowledge, rules and QSAR, (iv) models built by molecular graphics, (v) Artificial Neuron Network (ANN).

VALIDATION OF *IN SILICO* MODELS

The methods and models of QSAR and SAR were chosen among others to be able to fulfil demands and criteria of the programs of chemical safety namely REACH. OECD and JRC EC organize now activities to find criteria under which these models could be usable for regulatory purposes to provide materials on hazard identification

and useful for categorization of chemicals according their toxicities.

Joint meeting of International Council of Chemical Association and European Industry Council continued the discussion in a workshop on "Regulatory use of (Q)SAR for Human Health and Environmental Endpoints" in Setubal (Portugal) on spring 2002. The aim was to unify various regulatory usage of SAR among OECD member countries and even between different agencies in the same member country (not in Czech Republic where not much application activity of this type exists besides the research). Especially because no internationally harmonised principles for assessing the development and validation status of the (Q)SAR have existed.

A number of principles were suggested in Setubal. The 34th joint meeting of the Chemical Committee and the Working Party on Chemicals, Pesticides and Biotechnology decided to strengthen activities for regulatory acceptance of (Q)SARs. An Expert group for (Q)SAR validation was established which started to work in spring 2003 and continued in September 2003 and May 2004 at the site of OECD in Paris. Recommendations of those meetings were suggested to the 35th joint meeting and were agreed. Subsequently they were accepted by the Coordination group of the Expert group for (Q)SAR.

The original Setubal principles were modified and the following set of revised principles for validation of QSAR compatible with computer character of the models and usable for regulatory purposes were submitted:

The (Q)SAR models used for the regulatory purposes should be associated with the following information: (i) a defined endpoint, (ii) an unambiguous algorithm, (iii) a defined domain of applicability, (iv) appropriate measures of goodness-of-fit, robustness and predictivity, (v) a mechanistic interpretation, if possible.

QUANTITATIVE STRUCTURE - ACTIVITY RELATIONSHIPS

QSAR (Quantitative Structure-Activity Relationships analysis or modelling is the analysis of possible relations between biological and physicochemical, chemical or physical properties of a series of compounds, xenobiotics, ie. compounds foreign to organisms, but supplied by chemical industry or laboratories. The analysis is carried out by methods of mathematical statistics and a mathematical formulae usable for prediction of the biological properties is derived:

$$BA_i = f(X_i)$$

where BA_i is a biological activity (magnitude of an effect, for us toxic or adverse effect) of a compound i , X_i is a property connected with the chemical structure of the compound i , f is a mathematical function derived for a certain series of compounds. Consequently that formulae can be used to predict by calculation the index of toxicity (eg. LD50 with rats after oral application determined

after 24 hours after the application) knowing value of the physicochemical property X_i (i is the untested compound belonging to a series).

Regression analysis of the techniques of mathematical statistics is the most often used technique of the analysis. Not the only one, however, there are many other methods which are used for the QSAR analysis like: (i) factor analysis, (ii) principal component analysis, (iii) cluster analysis, (iv) pattern recognition analysis, (v) discriminant analysis, (vi) artificial neuron network (ANN), and (vii) cybernetic techniques like self-learning machine, genetic algorithm, linear surface decision, etc. This list is shown to draw attention to a variety of the methods used for the QSAR analysis.

Statistical evaluation is an important and necessary part of the model/equation. The only equation derived says nothing about its correctness (see the criteria of the validation). At least the minimum known statistical characteristics must be given like an amount of pairs (biological and physicochemical data) (n) in a series analyzed, correlation coefficient r , standard deviation of a value estimated by the equation (SD), statistical level of significance of the equation.

The description of chemical structure of a compound is joint with physicochemical, chemical or physical descriptors (properties). The properties which govern the magnitude of a biological effect (biological activity) are separated to three groups: (i) expressing hydrophobicity of a molecule, (ii) expressing its reactivity and electronic effects, (iii) expressing steric factors and spatial arrangement of the molecule. The fourth group can be arranged by those properties, which are complexed and cannot be simply defined in the three groups mentioned above (like molar refraction, parachor).

Decadic logarithm of partition coefficient of compounds between n-octanol and water (today already as $\log P$ which is slang, correctly $\log P_{ow}$) is the most often occurring physicochemical descriptor in the QSAR equations simulating hydrophobicity. The other hydrophobic constants are used, however, too like (i) partition coefficient of a compound between two immiscible phases (water and gas), (ii) solubility, (iii) a constant derived from chromatographic data, (iv) saturated vapour pressure above a liquid compound, (v) Hansch substituent hydrophobic constant π , (vi) fragmental constant by Rekker.

Electronic constants (constants of reactivity) are represented especially by Hammett constant σ and its modifications. Besides quantum chemical indices can be seen like (i) electron density at an atom, (ii) nucleophilic superdelocalizability, (iii) electrophilic superdelocalizability, (iv) energy of the highest occupied molecular orbital, (v) energy of the lowest unoccupied molecular orbital or (i) resonance Swain-Lupton constant, (ii) inductive Swain-Lupton constant, (iii) dipolemoment, (iv) Taft constant for polar compounds or for aliphatic compounds.

The spacial arrangement can be characterized by (i) Taft steric substituent constant E_s , (ii) van der Waals radius of a substituent, (ii) volume of a molecule, (iii) molar volume (relative molecular weight/density), (iv) surface of

a molecule, (v) relative molecular weight, (vi) topological index of molecular connectivity.

QSAR MODELS

The main and largely used is the analysis by Hansch. This type of QSAR models I discussed as *in silico* alternative method for toxicity prediction (testing). The physicochemicals descriptors used in this model can be easily found in tables, data basis or can be calculated by a computer software. The parabolic form of the Hansch equation makes it possible to estimate the optimal hydrophobicity of a molecule for an effect desired.

The basic Hansch equation is:

$$\begin{aligned} -\log C &= k_1 (\log P_{ow})^2 + k_2 \log P_{ow} + k_3 \sigma + k_4 E_s + k_5 \\ \text{or} \\ -\log C &= k_1 \pi^2 + k_2 \pi + k_3 E_s + k_4 \sigma + k_5 \end{aligned}$$

where $-\log C$ is a logarithm of an effective concentration, constants k 's are a result of the regression analysis. The Hansch equation covers all three properties of a compound responsible for its biological effect. A weight of the descriptors is calculated and its significance for the prediction evaluated. Often $\log P_{ow}$ is sufficient, the other descriptors contribute to the correlation insignificantly. Instead of the descriptors $\log P_{ow}$, π , σ or E_s another descriptor mentioned in their review above, or also unmentioned, are used by authors. It is possible even to see a list of let us say 50 descriptors, the author applies commercial software of factor analysis or of principle component analysis. today such studies bring rarely anything new.

RESULTS

A total 11 case studies were already made by individual experts or small groups of experts by applying the "QECD principles for (Q)SAR validation": (i) QSARs for acute fish toxicity, (ii) QSARs for mutagenicity and carcinogen-

ity, (iii) QSARs for atmospheric degradation, (iv) METI biodegradation model, (v) ECOSAR expert system, (vi) BIOWIN model, (vii) DEREK expert system, (viii) human NOEL model by Multi-CASE, (ix) human NOEL model by MDL, (x) skin sensitisation by DEREK, (xi) rat oral chronic toxicity model by TOPKAT.

The case studies are analyzed and the application evaluated. The tuning and acceptation of the validation principles for the (Q)SAR models for regulatory purposes is under the progress. Real usage of the models for "testing" is, thus, not a fantasy (details of statistics and criteria used can be found¹⁻⁶).

ACKNOWLEDGEMENT

The presentation and work were supported by grants of GA ČR 305/03/1169, of IGA MZ ČR NJ7435-3 and by National Institute of Public Health.

REFERENCES

1. Tichý M. Biological activities of xenobiotics and their chemical structure (in Czech). Praha: Avicenum, 1983.
2. Tichý M, Halousková O. (eds.) Alternative methods of toxicity testing of chemicals (in Czech). Proceedings of a workshop on the same title, Chrudim-Pišťov, Czech Republic, April 27–28, 2005. Vodní zdroje EKOMONITOR, 2005.
3. Cronin MTD, Walker JD, Jaworska JS, Comber MHI, Watts CD, Worth AP. (2003) Use of QSARs in international decision-making frameworks to predict ecologic effects and environmental fate of chemical substances. Environ Health Persp *III*, 1376–90.
4. Cronin MTD, Jaworska JS, Walker JD, Comber MHI, Watts CD, Worth AP. (2003) Use of QSARs in international decision-making frameworks to predict health effects of chemical substances. Environ Health Persp *III*, 1391–401.
5. Erriksson L, Jaworska J, Worth AP, Cronin MTD, McDowell RM, Gramatica P. (2003) Methods of reliability and uncertainty assessment and for applicability evaluations of classification – a regression-based QSARs. Environ Health Persp *III*, 1361–75.
6. Jaworska J, Comber M, Auer C, Van Leeuwen CJ. (2003) Summary of a workshop on regulatory acceptance of (Q)SAR for human health and environmental endpoints. Environ Health Persp *III*, 1358–60.