QSAR

QSAR method

Risk assessment

In environmental risk assessment, knowledge of the acute toxicity, chronic toxicity and environmental fate and behaviour of a chemical substance is a basic need. Factors affecting the environmental fate and behaviour of a chemical comprise its water solubility, adsorption to soil and sediments, volatilisation, biotic and abiotic degradation, and bioaccumulation. Quantitative knowledge of these processes enables one to model the concentrations of a certain chemical substance in the different environmental compartments (soil, air, water, and sediment). The knowledge of the toxicity of a chemical to aquatic organisms is normally limited to simple effects as lethality, growth or reproduction inhibition. The effect concentration for acute toxicity is expressed as the LC50 (EC50), the aqueous concentration that produces 50% lethality (and/or other effects). The effect concentrations for long-term or chronic effects is expressed as the NOEC, the highest test concentration with no observed effect on e.g. reproduction, population growth or other kinds of sublethal toxicity. An approach towards the toxicity of a compound with regard to environmental risk assessment would be to determine a “safe level”, a concentration at or below which, no organism or only a certain percentage of organisms in an ecosystem would be affected by the compound. Methods to predict the level of no-effect use the lowest acute or chronic values (TGD 1996) or interpolation of several values (e.g. Straalen & Denneman 1989, Wagner and Løkke 1991) and multiply with a relevant assessment factor.

Pesticides

For pesticides, the comprehensive data requirements demanded for authorisation normally mean that sufficient data for a risk assessment are present. This is not the case for the additives, impurities and substances used in the formulation of the pesticide product and usually not the case for the degradation or transformation products from biocidal active substances. The research devoted to develop reliable estimation procedures for the toxicity of environmental pollutants may therefore have a potential in estimating the needed data for the groups of substances. Today, the most promising technique for estimating the toxicity of pollutants is QSAR. However, it should be noted that QSARs should be applied within its recognised limits of applicability, e.g.

validity within a certain range of parameters (Kow-values, pH, etc.), certain groups of chemicals (carbamates, phenylureas, triazines, etc.), or mode of action.

Structure-Activity Relationships

SAR is based on the knowledge that substances with a similar (analogous) chemical structure may have the same biological activity. SAR is a qualitative comparison of the structures of chemical compounds and their effects in the biological system. From this evaluation of the influence of the chemical structure on the biological system, combined with experience in how changes in the chemical structure affect the magnitude and type of biological effect, unknown toxic effects to the biological system of unknown compounds with related chemically structure are predicted.

Quantitative Structure-Activity Relationships

QSAR is a statistical data analytical procedure in which quantitative endpoints of compounds (e.g. toxicity) are correlated with one or more structural parameters of these compounds, normally through uni- or multivariate linear regression (Chapman & Shorter 1978), non-linear regression (Könemann 1981), bilinear (Veith et al. 1983) or exponential regression. Commonly used structural parameters for inclusion in QSAR correlations are for instance:

- octanol-water partition coefficients (log Kow)
- aqueous solubility (log S)
- Molar Refraction or Parachor (dispersion forces)
- dipole moment • ionisation potentials
- molar volumes • molar surface areas (Hermens 1989).

Several variables have been used in attempts to obtain the best-fitted parameter(s).

N-Octanol/water partition coefficient

The parameter n-octanol/water partition coefficient, Kow, is an experimental data describing the lipophilicity of the substance. It has been shown that a non-linear relation between biological activity and lipophilicity exists. Substances of very low lipophilicity may be less able to pass

lipidous membranes and substances with a high lipophilicity will accumulate in fat tissue and other lipophile phases and may therefore not release a biological response.

Polarity

The polarity is an expression of the electronic distribution in the substance. The polarity is essential to the binding or release of the substance to an organism's membranes and/or macromolecules and thus determines whether a biological response may take place or not.

Stereochemical structure

The stereochemical structure may influence the possibility of interaction with the macromolecules of an organism. Size and shape should be suitable to fit into the receptor or enzyme before biological action may take place.

Scope

QSARs were originally mathematical models relating biological activity of chemicals to their structures and were developed and used mainly on the drug design area. Today, the scope has been broadened to predict any kind of data related to both toxicity and exposure of chemicals i.e. the two categories of data that integrated together should permit the risk assessment of chemicals. In ecotoxicology, QSAR models are used in the estimation of physicochemical and effect related properties of chemicals in non-tested endpoints to assess if testing is needed or not. QSARs are empirical models indicating that the results of evaluated studies are used in the further development of more precise models. The result of this iterative process is that QSARs change over time. As an example of the scope of the structure-activity based modelling, the following parameters are considered:

- physico-chemical properties
- the partitioning of pesticides among environmental compartments
- bioaccumulation potential
- aquatic toxicity

QSAR modelling

A QSAR model is a mathematical expression that relates the variation of the biological activity in a series of structurally similar compounds to the variation in their chemical structure. Thus, a QSAR model is a mathematical equation describing the activity for a specific class of substances and derived from the quantified measured data belonging to these substances. The strategy

mainly rests on the concept that biological data measured for a few compounds selected may form the basis for a QSAR of a class. The developed QSAR models may permit the estimation of the corresponding missing data for all the non-tested compounds belonging to the class, regardless of their number. In order to validate the QSAR models, measured and predicted values are compared. The experimental values used in this report are mostly obtained from letters of approval or denial for sale or import given according to the current statutory order from the Danish Ministry of the Environment where information from the applicants are evaluated by the Danish EPA's Pesticide Division. Other major sources of information are the Pesticide Manual (Tomlin 1994, 1997) and Linders et al. (1994). The experimental values are compared by linear regression analyses with QSAR estimates derived from QSAR models. The QSAR models used are the currently most preferred models.

The QSAR models or mathematical equations have been developed on the basis of experimental data on model substances. During the development of QSAR models, the calculations and testing were performed by using a great number of substances, e.g. high production volume substances or other industrial substances. These industrial substances had mostly simple structures. QSAR was previously used in the chemical industry in the development of new substances and only within the last decade the models have been refined to the use in assessment of chemical substances effects, fate and behaviour in the environment. The American Environmental Protection Agency (US-EPA) has developed a system of QSARs which are connected to a database (AQUIRE) and can therefore use the latest evaluated endpoint-values, whether physico-chemical, effect or fate data. This should improve the models as the reliability of model estimations relies on the precision of the input data. The model system iscalled ASTER: Assessment Tools for the Evaluation of Risk (Russom 1991, Pedersen et al. 1995). The US-EPA has also developed a computer programme for estimating the ecotoxicity of industrial chemicals based on structure activity relationships: ECOSAR. The programme uses specific QSARs for different chemical classes (US-EPA 1994). Because the programme was not complete at the time of this work, it was not used. The value in using QSARs in the environmental assessment is that in the absence of experimental data employing QSAR may derive the missing variables. Besides, when several experimental studies on the same chemical substance are giving information on single endpoints or parameters which are not complementary or in the same range, the decision on which results to use may be supported by QSAR (TGD 1996). When applying QSAR, it

should be taken into account that a QSAR is an estimation method and therefore, there is a certain probability that the estimate is poor even for well-evaluated models. QSAR model estimates cannot be the only basis for preparing risk assessment. QSAR estimates should be seen as a complementary tool which, evaluated together with test results, can provide a more complete understanding of the physico-chemical and ecotoxicological characteristics of the substance. This means that QSARs are no better than the data on which they are based. Furthermore, it should be noted that QSAR models, generally, only exists for discrete organic substances and not for more complex substances or reaction mixtures.

Thus, QSARs can be used to assist data evaluation

- to contribute to the decision on whether further testing is necessary to clarify an endpoint of concern
- to establish input parameters which are necessary to conduct the exposure or effect assessment.

QSAR models should only be used in risk assessment if the models have been thoroughly evaluated and no experimental data or conflicting validated experimental data exist. As the work on QSAR model development and evaluation is being performed in national and international programmes, the various models change currently.

Environmental risk assessment is based on a comparison of two variables:

- the concentration of the chemical in the environment (exposure)
- the concentration of the chemical at which no adverse effects on the environment are expected or estimated to occur.

Concentration in the environment

Measurements of the actual concentration in the environment are to be preferred. However, in many cases the concentration that can be expected after the release of the chemical in the environment (exposure) is the most interesting issue. Fate modelling techniques may be applied to estimate these expected concentrations.

Fate models require an input of data for the various fate processes, e.g.:

- abiotic degradation (hydrolysis, photolysis, oxidation)
- biodegradation
- adsorption to soil, sediments, suspended particles
- volatilisation, evaporation
- leaching
- bioaccumulation

The rate or equilibrium constants of these fate processes can be measured in the laboratory and technical guidelines are developed to ensure comparable results (e.g. EU 1992, OECD 1993). To ensure a comparable result that can be used in the risk assessment, the laboratory results are used instead of field data where factors affecting the results may be less controllable. The fate processes determine the extent to which the organisms are exposed to the substance, i.e. the extent to which the chemical is bioavailable. Several models are developed to predict/estimate the environmental fate processes and based on the physico-chemical parameters. Physico-chemical properties are important data for the exposure analysis. Especially, a few physico-chemical parameters and variables are observed to be important, i.e.:

- the size and structure of the chemical
- the water solubility
- vapour pressure
- octanol/water partition coefficient (lipophilicity)
- adsorption coefficient.

For a long time, the lipophilicity character has been shown to play a basic role in determining distribution phenomena as well as influencing the mechanisms of ecotoxicity of organic chemicals. The classic procedure for measuring lipophilicity of organic chemicals is based on the partition between octanol and water.

Conclusion

The development of QSAR is based on the assumption that chemical substances, which reach and interact with a target site by the same mechanism, perform likewise due to their similar chemical properties. Effect concentration Conclusion 20 The analysis for QSAR is through regression method affording transparent relations and simple mathematical equations and leading to a quantitative correlation. However, for a meaningful regression analysis, precise and

accurate input data are required which tend to limit the number of samples in the testing set. It also means that the data used should be carefully evaluated and not taken from any available handbook unless the data quality is known to be validated (Hart 1991).

Reference Pesticides Research No. 94 2004 Bekæmpelsesmiddelforskning fra Miljøstyrelsen