Replacing Experimentation

Use of a holistic approach in validation: QSAR models versus experimental data

Pascal BICHEREL, Franklin BAUER, Faizan SAHIGARA, Paul THOMAS

KREATiS, Knowledge & Research in Environment And Toxicology in Silico

KREATiS SAS, 23, rue du Creuzat, ZAC St Hubert, 38080 L'ISLE D'ABEAU, FRANCE

www.kreatis.eu

Introduction

Experimental results are often considered as reliable provided methodological and experimental conditions are scientifically appropriate (e,g. following an OECD guideline). However there is often limited information on methods in studies which are sourced from databases. This is especially the case for handbooks which may not contain further information than the study result. This hinders precise validation, forcing the reader to locate the original literature source. In the worst case, no information can be obtained and the study must be classified as invalid (eg. K4 according to the Klimisch scoring method[1]).

One option is to use High Accuracy Quantitative Structure-Activity Relationship models (known as 'HA-QSARs') as a supporting study without resorting to further laboratory testing. KREATiS has developed a series of HA-QSARs which are included in the iSafeRat® Toolbox[2]. In comparison to standard QSAR models these offer predicted values reliable enough to replace experimental studies for substances within the applicability domain and are validated following criteria specified in regulatory guidelines. The models follow a "holistic approach" where predicted properties are related to each other by the laws of phase-equilibrium thermodynamics. Octanol-Water Partition Coefficient (K_{OW}) is obviously linked to water solubility but further physicochemical or even biological properties can also be interlinked. Mayer and Reichenberg (2006)[3], Mackay *et al.* (2009)[4] and ECETOC (2013)[5] introduced the chemical activity of a substance as a central property which can accurately explain its toxicological properties. The results presented here are an example of the « holistic approach ». The aim of this approach is to increase confidence in the final predicted result in order to evaluate experimental data found in the literature devoid of details on study quality.

Principle of "holistic approach"

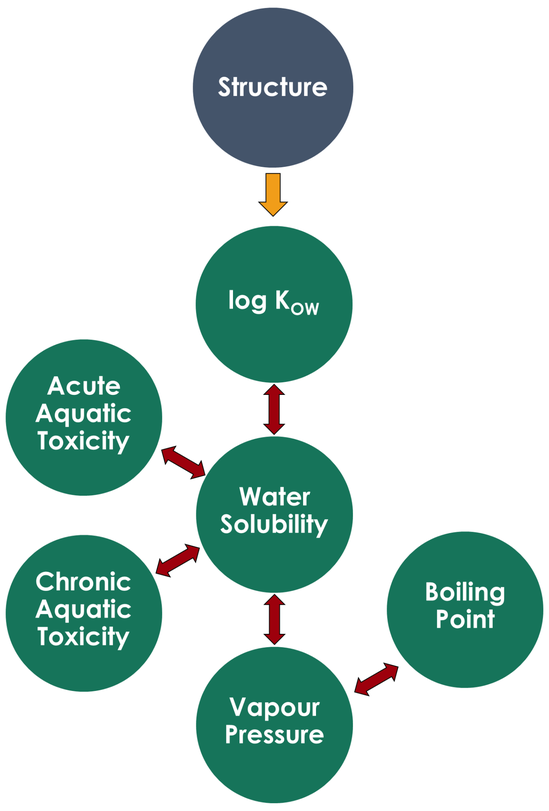

Figure 1: Holistic approach

The "holistic approach" as defined by KREATiS and used in iSafeRat® Toolbox interlinks six properties : Octanol-Water Partition Coefficient (K_{OW}), water solubility, vapour pressure, boiling point, acute and chronic aquatic toxicity as shown on **Figure 1**.

Initially the log K_{OW} can be determined from the structure using a fragment approach developed by KREATiS where the log K_{OW} of a substance is the combination of the contributions of fragments that compose it. This step is the most critical since it is the input of all the following predictions. The other properties are then determined through simple linear regression in a cascade calculation. Based on phase-equilibrium thermodynamic principles, this method provides simple relationships with strong correlations between few parameters thus guarantying high prediction accuracy along the cascade.

1,3,5-trimethylbenzene (mesitylene) (CAS# 108-67-8, SMILES c(cc(cc1C)C)(c1)C, **Figure 2**) is used as an example. In this case, molecular structure is the input used to generate a log K_{OW} value. Other parameters are then calculated from it. In **Table 1** calculated data with iSafeRat® v1.3 are compared to experimental values. The results are consistent with experimental data found on ECHA disseminated dossier.

Table 1: Predicted properties of mesitylene with iSafeRat® v.1.3 from structure compared to experimental values.

Properties	log K_{OW}	Water Solubility (mg/L)	Aquatic Toxicity to daphnid (mg/L)		Vapour Pressure (Pa)	Boiling Point (°C)
			acute[a]	chronic[b]		
iSafeRat® v1.3	3.53	68	3.8	0.32	505	157
Experimental	3.42	48.2	6	0.4	320	164.7

[a] 48h-EC50
[b] 21d-NOEC

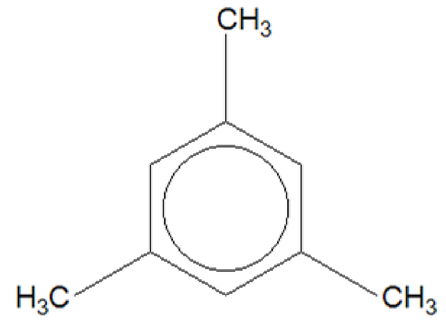

Figure 2: Molecular structure of mesitylene

HA-QSAR models for study evaluation

As HA-QSAR models are strictly validated, they can precisely determine the properties of a substance which falls within their applicability domain, thus allowing confirmation or invalidation of a study result. For example, for mesitylene a 48h-EC50 on algae based on biomass has been reported as 25 mg/L in an ECHA disseminated dossier (assessed as a K2 study). However deviations from OECD guideline (eg. no analytical monitoring) can introduce bias and the reliability of the result cannot be ensured. And therefore iSafeRat® was used to test the validity of this result.

Model predictions for the iSafeRat® HA-QSAR model can be based on information at any point in the cascade. Thus using the structure as an input, the HA-QSAR predicts a 72h-ErC50 of 3.5 mg/L based on algae growth rate. Moreover, when valid experimental data on other endpoints (eg. log K_{OW}, water solubility or other ecotoxicity endpoint values) are used as an input for iSafeRat® HA-QSAR models, the 72h-ErC50 for algae may also be derived which provide a whole set of predictions for the same endpoint. This approach is called '**holistic reinforcement approach**'. Predictions are reported in **Table 2**. Using this holistic approach it is possible to conclude that the ErC50 is below 10 mg/L. The mean (3.6 mg/L) is very close to the initial value predicted from structure. Therefore the value of 25 mg/L should not be used for regulatory purposes as a value reflecting the toxicity to algae.

Table 2: HA-QSAR predictions of algae growth rate inhibition (72h-ErC50) from properties using 'holistic reinforcement approach'.

Input properties	72h-ErC50 (mg/L) to algae (derived from the experimental values of several properties)
log K_{OW}	4.3
Water Solubility	2.6
Acute Toxicity to daphnid	5.5
Chronic Toxicity to daphnid	4.3
Vapour Pressure	2.2
Boiling Point	2.6
	mean derived value 3.6

Summary

In this exercise, an initial prediction was made of ErC50 based on structure. The results were corroborated by a series of predictions from experimental data on related endpoints using the '**holistic reinforcement approach**' (see **Table 2**). The mean derived value can then be compared to the original prediction derived from the structure

This method enhances consistency between predictions. Therefore it is a powerful tool to validate experimental results and QSAR predictions alike.

In this case a predicted ErC50 of 3.5 mg/L based on structure and 3.6 mg/L based on the mean derived value from six different experimental values can be considered as strong evidence that the experimental value of 25 mg/L is not valid.

References

[1] Klimisch H.J., Andreae M., Tillmann U. (1997) A Systematic Approach for Evaluating the Quality of Experimental Toxicological and Ecotoxicological Data. Regul Toxicol Pharmacol. 25(1), 1-5.
[2] iSafeRat® – in Silico Algorithms For Environmental Risk And Toxicity version 1.3
[3] Mayer P., Reichenberg F. (2006) Can highly hydrophobic organic substances cause aquatic baseline toxicity and can they contribute to mixture toxicity? Environmental toxicology and chemistry 25 (10), 2639-2644
[4] Mackay D., Arnot J.A., Petkova E.P., Wallace K.B., Call D.J., Brooke L.T., Veith G.D. (2009) The physicochemical basis of QSARs for baseline toxicity. SAR QSAR Env. Res. 20(3-4), 393-414
[5] Thomas P., Dawick J., Lampi M., Lemaire P., Presow S., van Egmond R., Burgos M.G. (2013) Activity based relationships for aquatic ecotoxicology data: use of the activity approach to strengthen MOA predictions. ECETOC Technical Report no. 120.

Poster No. WE016

This poster is presented at SETAC Europe 25th annual meeting 2015, Barcelona, Spain.