National Institute
for Public Health
and the Environment

Report 607220011/2008
W.J.G.M. Peijnenburg | H.A. den Hollander

Supporting REACH – Development of building blocks of a module for intelligent testing of data-poor organic substances

RIVM Report 607220011/2008

Supporting REACH – Development of building blocks of a module for intelligent testing of data-poor organic substances

W.J.G.M. Peijnenburg
H.A. den Hollander

Contact:
W.J.G.M. Peijnenburg
Laboratory for Ecological Risk Assessment
wjgm.peijnenburg@rivm.nl

This investigation has been performed by order and for the account of Directorate-General for Environmental Protection, Directorate for Chemicals, Waste and Radiation Protection VROM/DGM/SAS, within the framework of Project M/607220 (Ondersteuning REACH)

RIVM, PO Box 1, 3720 BA Bitlhoven, the Netherlands. Tel +31 30 274 91 11 www.rivm.nl

Abstract

Supporting REACH – Development of building blocks of a module for intelligent testing of data-poor organic substances

The new EU legislation for industrial chemicals, REACH, obliges registrants to collect all available relevant information on the intrinsic properties of a substance. Many of these properties are unknown and/or even impossible to measure. For this reason, one of the so-called REACH implementation projects provides a general guidance on Intelligent (or Integrated) Testing Strategies (ITSs), with the aim of optimizing the use of available data and reducing animal testing. In this context, data-poor chemicals are a particularly difficult challenge. We report here the first steps towards the development of a module for dealing with data-poor chemical classes. The focus of our research was on the development of methods that use chemical structure as the sole input parameter for predicting the toxicity of specific organic substances (in this case, carbamates and organophosphate esters, and their metabolites) to aquatic organisms. Methods such as these are eagerly awaited as they are essential for the successful implementation of REACH. Carbamates and organophosphate esters were selected as the chemical classes to be studied because, despite their large application volumes, a relatively limited amount of information is available on their fate and effects in the environment. The report describes how the formation of metabolites of the specific chemicals can be predicted with the QSAR-based computer programme CATABOL. Quantum-chemical descriptors of the substances and their metabolites were computed by CHEM3D. Based on these descriptors, structure–activity relationships were developed to predict the toxicity of the starting compounds and their metabolites to aquatic organisms.

Key words:
REACH, risk assessment, integrated testing strategy, RIP, carbamates, organophosphates, QSAR

Rapport in het kort

Ondersteuning van REACH – Ontwikkeling van bouwstenen voor een module voor het intelligent testen van data-arme organische stoffen

De nieuwe EU-regelgeving met betrekking tot de productie en het gebruik van chemische stoffen (REACH) streeft naar een verbetering van de kwaliteit van een gezonde leef- en werkomgeving. Stoffen komen direct en indirect in het arbeids- en leefmilieu terecht. Op dit moment is voor veel stoffen onbekend wat de gevaren zijn voor de volksgezondheid en de effecten op de leefomgeving. Binnen REACH wordt onder andere gestreefd naar een minimalisatie van het gebruik van proefdieren. Aan de andere kant moet in de komende jaren een inhaalslag gemaakt worden om essentiële kennislacunes weg te nemen. Hierbij dient zoveel mogelijk gebruik gemaakt te worden van bestaande stofgegevens, waarbij het essentieel is om de beschikbare informatie zo efficiënt mogelijk te gebruiken. In dit rapport worden enkele bouwstenen uitgewerkt van een module voor een geïntegreerde teststrategie (ITS in REACH-termen) voor data-arme stoffen. Dit is gedaan voor twee stofgroepen: carbamaten en organofosfaat-esters. Dit zijn twee stofgroepen met diverse toepassingen terwijl het ontbreekt aan kennis over hun lot en effecten in het milieu. De aandacht ligt bij dit uitgewerkt voorbeeld op het voorspellen van de aquatische toxiciteit van zowel de uitgangsstoffen als van hun metabolieten.

Trefwoorden:
REACH, risicobeoordeling, teststrategie, carbamaten, organofosfaten, QSAR

Contents

Summary

The new EU regulation regarding the production and use of chemical substances (REACH) is aimed at improving the quality of the environment for humans (include workers) and ecosystems. Chemicals are emitted directly and indirectly in these environments. At this moment, the extent of adverse effects on humans and ecosystems following emissions is virtually unknown for many chemicals. REACH requires making up for the lack of essential information on chemical substances within a limited period of time. Thereupon, REACH is aimed at minimizing animal testing. Within so-called Integrated Testing Strategies (ITSs), existing chemical data are to be used as efficiently as possible in connection with newly to be developed assessment tools like read-across methodologies, structure-activity relationships, weight-of-evidence reasoning based on several independent sources of information, and *in vitro* testing. Some of the building blocks of a module for an ITS for data-poor chemicals are designed in this report. This design was performed for two chemical classes that despite their widespread use are to be considered as being data-poor: carbamates and organophosphate esters. The focus is on prediction of the aquatic toxicity of the metabolites of the chemical substance classes investigated.

1. Introduction

The implementation of the new EU legislation concerning the Registration, Evaluation, Authorization and restriction of CHemicals (REACH) requires amongst others the implementation of a multitude of tools that will assist in meeting the main objectives of REACH of efficient safety management of chemicals whilst minimizing the use of test animals. The array of tools to be optimized in Intelligent (or Integrated) Testing Strategies (ITSs) include non-testing information on top of a minimum amount of newly generated data. Read-across methodologies, computational methods like Quantitative Structure-Activity Relationships (QSARs), as well as Weight-Of-Evidence reasoning (WOE) based on several independent sources of information, and *in vitro* testing are to supplement existing experimental and historical data and substance-tailored exposure driven testing.

Despite efforts to supplement available data by newly generated properties, most chemicals to be assessed are to be considered as being data-poor with regard to physico-chemical properties and effect data. This necessitates to have operational a multitude of alternatives like the ones mentioned above. RIP 4 (RIP = REACH Implementation Project, RIP 4 deals with the technical guideline documents for authorization. This includes issues like dossier evaluation (4.1), substance evaluation (4.2), priorities for authorisation (4.3), restriction / derogation of chemicals (4.4), and priorities for evaluation (4.5)) provides general guidance on dealing with (data-poor) chemicals. RIP 4 deals (amongst others) in a generic sense with ITS, although guidance on this issue is not detailed. ITSs have been defined for a number of assessment endpoints but the building blocks are in a state of design at best and further efforts are needed to generate robust ITSs.

Metabolites are a special class of data-poor chemicals. Metabolites are chemicals that are formed following release of a chemical in the environment as a result of the interaction with abiotic or biotic phenomena; metabolites are the result of incomplete degradation of the parent compound. As for most chemicals to be evaluated, the chemical structure is the sole piece of information that is always available. Predictive tools based solely on the chemical structure are therefore one of the methods of choice for inclusion in ITSs. Quantumchemical descriptors (i.e. descriptors based upon the basic properties of a chemical, like the charge of atoms in the molecule, the energy of molecular orbitals, dipole moment, polarity, the total energy of the chemical, the heat of formation, etc.) are the tools of choice in this respect as they only require the molecular structure as input. Quantumchemical descriptors are currently made available in an increasing user-friendly mode via, amongst others, the internet. Basically, information on the chemical structure is the sole requirement for the derivation of quantumchemical descriptors. Recent progress in computational hardware and the development of efficient algorithms has assisted the routine development of molecular quantum-mechanical calculations. Quantumchemical calculations are thus an attractive

source of new molecular descriptors which can in principle express all of the electronic and geometric properties of molecules and their interactions, turning them well-suited to provide the building blocks for modelling fate and effects of data-poor chemicals. So far, QSARs based upon quantumchemical descriptors are scarce.

The main objective of the study reported here is to explore the possibilities of using quantumchemical descriptors as the basis for deriving QSARs for prediction of the aquatic toxicity of data-poor chemicals and their (predicted and/or measured) metabolites. The latter allows assessment of the possibility of formation of more toxic metabolites following environmental release of a chemical. Carbamates and organophosphate-esters (OP-esters) were selected as chemical classes as, despite their large usage volumes and despite their toxic profiles, these compound classes are typical examples of data-poor chemicals that require assessment within REACH. Experiences gained with these chemical classes may serve in setting-up specific modules within ITSs for assessing the risks of data-poor chemicals. Metabolite formation was predicted by means of the application of CATABOL. CATABOL is a model which can be used for quantitative assessment of the biodegradability of chemicals. The model allows for identifying potentially persistent catabolic intermediates, their molar amounts, and solubility (water solubility, log K_{ow}, BCF). Presently, the system simulates the biodegradability in MITI-I OECD 301 C and Ready Sturm OECD 301 B tests. Other simulators will be available in the program upgrades. The latest version of CATABOL (version 5) allows defining the degree of belonging of chemicals into the domain of the biodegradation simulator.

It should be noted that the objective is NOT to develop an ITS for biodegradation or for assessment of toxicity. Instead we intend to develop building blocks for filling in the ITS on these endpoints.

The report is structured as follows: following an overview of the regulatory background as related to REACH (chapter 2), models are derived in chapter 3 for predicting the aquatic toxicity of a selected number of carbamates and OP-esters. Chapter 4 deals with the validation of CATABOL; predicted metabolite formation patterns are compared to experimental observations. Chapter 5 deals with toxicity prediction of the metabolites identified during validation of CATABOL using the models reported in chapter 3. Finally, chapter 6 provides a short overview of the highlights of this study.

2. Regulatory background

2.1 REACH and the need for alternative testing

The implementation of REACH requires demonstration of the safe manufacture of chemicals and their safe use throughout the supply chain. REACH is based on the precautionary principle, but aims to achieve a proper balance between societal, economic and environmental objectives. Both new and existing chemicals will be evaluated within REACH, on the one hand aiming to speed up the slow process of risk assessment and risk management of existing substances whilst on the other hand attempting to efficiently use the scarce and scattered information available on the majority of new and existing substances. REACH thus aims at closing huge gaps of knowledge on physico-chemical properties and adverse effects of large numbers of chemicals. Thereupon REACH aims to reduce animal testing by optimized use of qualitative and quantitative information on related compounds. Detailed information on all aspects of the implementation of REACH can be found on the website: http://ec.europa.eu/enterprise/reach/prep_guidance_en.htm.

The basic elements of REACH are as follows:

Registration - In principle REACH covers all substances, but some classes of substances are exempted (e.g. radioactive substances, polymers and substances for research and development). The safety of substances is the responsibility of industry. Manufacturers and importers of chemicals are therefore required to obtain information on their substances in order to be able to manage them safely. The extent of the obligations depends upon the quantity of the substances manufactured or imported. For quantities of 1 tonne or more per year a complete registration has to be submitted. For substances of 10 tonnes or more per year, a Chemical Safety Report (CSR) has to be included. Since one of the goals of REACH is to limit vertebrate testing and reduce costs, sharing of data derived from *in vivo* testing is mandatory.

The information on hazards and risks and how to manage them is passed up and down the supply chain. The main tool for downstream information is the safety data sheet (SDS), for dangerous substances only. An SDS contains information which is consistent with the chemical safety assessment. Relevant exposure scenarios are annexed to the SDS. The downstream user is required to apply appropriate measures to control risks as identified in the SDS.

Evaluation - Evaluation will be performed on registration dossiers, to check the testing proposals and the compliance with the requirements of registration. In addition, substances which are

suspicious of being a threat to human health or the environment can be evaluated by a Member State.

Authorisation - Authorisation of use and placing on the market is required for all substances of very high concern, regardless of tonnage level.

Restrictions - Restrictions may apply to all substances, regardless of tonnage level.

Classification and labelling inventory - Directives 67/548/EEC on classification and labelling of substances and 1999/45/EC on classification and labelling of preparations will be amended to align them with REACH.

Currently, around 100,000 different substances are registered in the EU, of which around 30,000 are manufactured or imported in quantities above 1 tonne. The existing regulatory system inherent in current EU policy for dealing with the majority of these chemicals - known as 'existing' substances - has been in place since 1993 and has prioritised 140 chemicals of high concern. Although a programme of work has been drawn up, the current EU legislation on chemicals has several drawbacks. Firstly, a substantial number of existing chemicals which are marketed have not been adequately tested. Information related to their hazard potential is minimal (less than base-set), and they may be harmful to human health or the environment. This contrasts sharply with new chemicals which have to be notified and tested starting from volumes as low as 10 kg per year, discouraging research and invention of new substances. Secondly, there is a lack of knowledge on use and exposure. Thirdly, the present process of risk assessment and chemical management in general is relatively slow, and certainly too ineffective and inefficient to take care of the problems raised by the huge data gap in the field of the existing chemicals. And last but not least, the current allocation of responsibilities is not appropriate: public authorities are responsible for the risk assessment of substances, rather than the enterprises that produce or import them (JRC, 2005).

The 30,000 existing substances manufactured or imported in quantities above 1 tonne are to be assessed through the REACH process within a proposed time-window of eleven years. A major topic within the assessments is the availability of data. On the basis of experiences within the US Challenge Program on regulation of High Production Volume Chemicals, it is expected that adequate data are available only for about 50 % of the endpoints to be assessed, various estimation methods and strategies to limit data needs will substitute for the majority of the lacking data, and about 6-7 % of the lacking data are expected to be derived by means of additional testing (Table 1).

Table 1: Experience from the US HPV Challenge Program.

	Human health data	Environmental effects
Adequate studies	50%	58%
Estimation	44%	35%
Testing	6%	7%

The REACH proposals advocate the use of non-animal testing methods for the generation of lacking data, but guidance is needed on how these methods should be used. As an example: the REACH system requires that non-animal methods should be used for the majority of tests in the 1-10 tonne band, even though such methods are not yet available for most of the endpoints relevant at this tonnage.

In an attempt to resolve the issue of lack of guidance, the European Commission made suggestions on how reduction, refinement and replacement strategies could be applied to animal use in the REACH system:

1 – encouragement of the use of validated *in silico* techniques such as (Q)SAR models.

2 – encouragement of the development of new *in vitro* test methods.

3 – minimization of the actual numbers of animals used in the required tests, and replacement of animal tests wherever possible by alternative methods.

4 – formation of Substance Information Exchange Forums (SIEFs) for the obligatory provision of data and cost sharing.

5 - requirement of official sanctioning of proposals for tests for compounds with production volumes of above 100 tonnes to minimize animal testing.

The consequence of REACH is that in a relative short time period the risk of a large group of chemicals has to be assessed, which implies that also a large amount of information on the fate and effects of chemicals has to become available. In principle, this can be achieved by conducting a large number of human toxicity and ecotoxicity studies as well environmental fate and behaviour studies. However, not only in REACH but in OECD as well, there is understanding that for reasons of animal welfare, costs and logistics, it is important to limit the number of tests to be conducted. In line with ANNEX XI of the REACH proposal, the generation of a comprehensive test dataset for every chemical will not be needed if these test data can be replaced by the following methods:

- non testing methods:
  1 - the application of grouping (categories) and read-across
  2 - computational methods (SARs, QSARs and biokinetic models);
- *in vitro* tests;
- existing experimental and historical data;
- substance-tailored exposure driven testing;
- weight-of-evidence reasoning (WOE) based on several independent sources of information.

This means that alternative methods (non-testing methods or *in vitro* tests) have to be developed as well as weight-of-evidence schemes that allow regulatory decisions to be made. These have until now been used but to a varying degree and in different ways for risk assessment, classification & labelling, and PBT assessment of chemicals (EC 2003, EC 2004). The benefits of using such non-testing methods have included:

• avoiding the need for (further) testing, i.e. information from non-testing methods has been used to replace test results;

• filling information gaps, also where no test would be required according to current legislation;

• improving the evaluation of existing test data as regards data quality and for choosing valid and representative test data for regulatory use. Furthermore, use of non-testing data in addition to test data employing weight of evidence could increase the confidence in the assessments;

Thus, the use of non-testing information has improved the basis for taking more appropriate regulatory decisions (as well as for voluntary non-regulatory decisions taken by industry). In effect, use of non-testing information has decreased uncertainty, or even made it possible to conclude on a classification or the need for more information in relation to hazard, risk and PBT assessment.

Alternative methods are in several stages of development, verification and validation, and they therefore cannot yet be used as stand alone. Other information gaps will exist. It is therefore necessary to integrate all available information into a so called integrated or intelligent testing strategy (ITS). In this way, all possible available information on a substance can be optimally used and further testing will only be required where essential information is lacking (Bradbury et al., 2004; Combes et al., 2003; Vermeire et al., 2006). Figure 1 schematically depicts the various approaches that may provide the building blocks for Intelligent Testing Strategies.

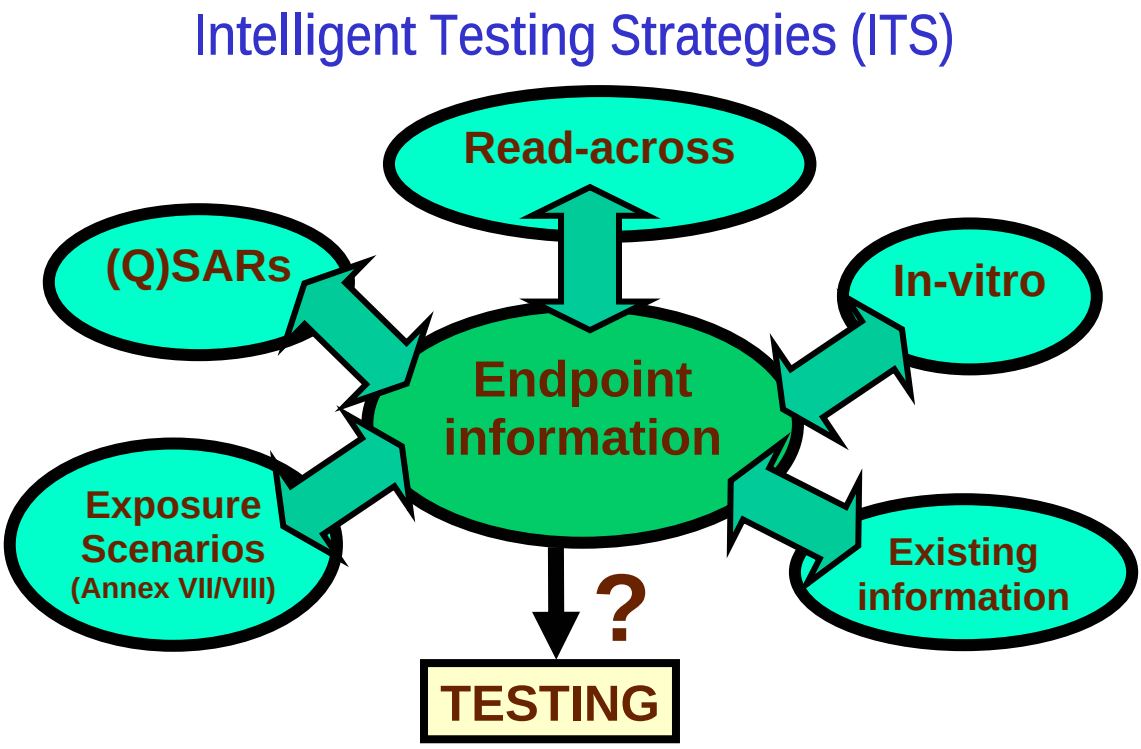

Figure 1: Constituents of an Intelligent (or Integrated) Testing Strategy (ITS). Taken from a presentation of Van Leeuwen and Bradbury (2005).

Six components are currently proposed for inclusion in ITSs (IHCP, 2005):

Exposure-based waiving. In a tiered approach to evaluating the risk associated with chemical substances, derogation of testing requirements can be justified at an early stage if the exposure is known to be negligible in the environmental compartments of interest. To this end, increased realism in regulation-relevant exposure assessment is required. This will include refinement of exposure models targeted for triggering ITSs accordingly, elaboration and harmonization of the meaning of low exposure, and development of procedures for incorporating relevant use pattern information, taking into account the European diversity.

Read-across and chemical categories. Read-across has great potential to reduce animal testing, and contributes to achieve better-informed decisions through evaluating the chemistry-specific context. However, transparent extrapolation from information gained for chemically similar compounds requires specifying how to define chemical similarity covering both the structural domain and property profile, and to develop guidance for the qualitative and quantitative extrapolation of biological testing results in a regulatory context.

Structure-activity relationships and computational chemistry. Qualitative and quantitative structure-activity relationships (QSARs) are used to predict the toxicity and fate of chemicals from molecular structure information, employing different levels of computational chemistry. They are sometimes called '*in* silico' models because they can be applied by using a computer. Both

categorical data and continuous data can be addressed. Besides regulatory endpoints for classification, labelling and hazard evaluation, *in silico* methods may also generate mechanistic knowledge that guides targeted testing and enables informed extrapolation across species as well as between human health and ecotoxicology. *In silico* methods need to be made fit for regulatory purpose, and should include technical issues such as applicability domain (one of the OECD principles for QSAR validation), prediction power, and also metabolism simulators.

Thresholds of toxicological concern. Thresholds of toxicological concern (TTCs) are exposure thresholds below which no significant risk is expected. The TTC concept relies on the assumption that one can identify a concentration threshold below which the risk of any chemical for any harm is acceptably low, as has been proposed for food additives by the FDA (1995).

***In vitro* tests**. *In vitro* systems implicitly require the use of prediction models to extrapolate from *in vitro* data to *in vivo* information. Since one decade there is a focus on the development and validation of alternative test methods within ECVAM and OECD. So far, eight *in vitro* methods have been proposed as scientifically valid for the assessment of chemicals, and a full replacement of animal testing may not be possible for the majority of endpoints because of the reductionistic nature of *in vitro* cell cultures as compared to *in vivo* systems. The mechanistic basis of *in vitro* approaches needs further study. Focus should be on their great potential to contribute significantly to the reduction and refinement of animal tests, particularly when combined in an intelligent manner with other ITS components such as QSARs and genomics.

Optimized *in vivo* tests. Strategies to reduce the use of laboratory animals include elimination of redundant tests, use one sex whenever possible, greater use of screens, and threshold approaches instead of full dose-response. Another route of optimization concerns the refinement of animal testing through introduction of non-lethal endpoints. Procedures for the regulatory acceptability of the various optimization strategies need development, and new opportunities to refine and reduce animal testing through guidance from mechanistic non-testing information provided by QSARs, read-across and *in vitro* data as well as through species-species extrapolation (biological read-across) need further attention.

2.2 Alternatives for testing within REACH

REACH implicitly requires that operational procedures are developed, tested, and disseminated that guide a transparent and scientifically sound evaluation of chemical substances in a risk-driven, context-specific and substance-tailored manner. The procedures include alternative methods such as chemical and biological read-across, *in vitro* results, *in vivo* information on analogues, qualitative and quantitative structure-activity relationships (SARs and QSARs, respectively), thresholds of toxicological concern, and exposure-based waiving. As stated in paragraph 1.1, the concept of 'Intelligent Testing Strategies' (ITSs) for regulatory endpoints has been outlined to facilitate the assessments. The basic idea is to obtain the information needed for carrying out hazard and risk assessments for large numbers of substances by integrating multiple methods and approaches with the aim to minimize testing, costs, and time. The goal is to feed regulatory decision making through a targeted exploitation of exposure, chemical and biological information with minimal additional testing. For example, (Q)SARs have been used in regulatory assessment of chemical safety in some OECD member countries for many years, but universal principles for their regulatory applicability were lacking. The OECD member countries agreed in November 2004 on the principles for validating (Q)SAR models for their use in regulatory assessment of chemical safety. In February 2007, the OECD published a 'Guidance Document on the Validation of (Q)SAR Models' with the aim of providing guidance on how specific (Q)SAR models can be evaluated with respect to the OECD principles (OECD/IPCS, 2005). An OECD Expert Group on (Q)SARs was established for this purpose and a (Q)SAR Application Toolbox is in development. Under the current EU legislation for new and existing chemicals, the regulatory use of estimation models or (Q)SARs is limited and varies considerably among the member states. This is probably due to the fact that there is no agreement in the scientific and regulatory communities over the application of (Q)SARs and the extent to which (Q)SARs estimates can be relied on. It is anticipated that these non-testing methods, in the interests of time- and cost-effectiveness and animal welfare, will be used more extensively under the future REACH system, and especially ECVAM (the European Centre for the Validation of Alternative Methods) is playing an important role in the operationalization of (Q)SARs for regulatory endpoints.

As stated above, eight *in vitro* methods have been proposed so far as scientifically valid for the assessment of chemicals (for example: methods for skin absorption, skin corrositivity, genotoxicity and phototoxicity), but many more still need to be developed, validated and accepted for regulatory use. For environmental endpoints, a number of fish cell lines, primary fish cell cultures and fish embryos are currently being studied to assess acute toxicity, a new approach for testing prolonged exposure in fish cells is being developed as alternative for chronic toxicity testing, and metabolically competent fish cell lines and primary cell cultures as gill epithelial cell

cultures are being used to mimic bioaccumulation as fish gills are the first point of contact for water-borne toxicants. *In vitro* systems implicitly require the use of prediction models to extrapolate from *in vitro* data to *in vivo* information, amongst others taking account of the observation that further investigations are required to ascertain the reasons for the reduced sensitivity of fish and mammalian cell lines to aquatic toxicants, compared with *in vivo* fish systems. There is a focus on the development and validation of alternative test methods within ECVAM and OECD. A full replacement of animal testing may not be possible for the majority of endpoints because of the reductionistic nature of *in vitro* cell cultures as compared to *in vivo* systems. Despite all progress achieved and promising future prospects, scientific advisory committees of the Commission (Scientific Committee on Toxicity, Ecotoxicity and the Environment [CSTEE] and Scientific Committee on Cosmetic Products and Non-Food products intended for Consumers [SCCNFP]) raised serious doubts about the potential of *in vitro* methods to fully replace *in vivo* experiments in the near future.

Concerted action and intensive efforts are needed to translate the ITS concept into a workable, consensually acceptable, and scientifically sound strategy. Initial ITS work has been performed in the REACH Implementation Project (RIP) scoping studies, amongst others attempting to develop testing strategies on four specific endpoints (irritation, reproductive toxicity, biodegradation and aquatic toxicity). One of the main conclusions was that existing strategies should be developed further and that the concept of ITS has maximal applicability across the REACH regulatory endpoints. Furthermore, the production of guidance and (web-based) tools was considered essential, and the outcome of the strategies should be applicable for risk assessment, classification and labelling, and PBT assessment. Thereupon, within the sixth Framework Programme the Integrated Project OSIRIS (Optimized Strategies for Risk Assessment of Industrial Chemicals through Integration of Non-Test and Test Information) has been initiated. The goal of OSIRIS is to develop integrated testing strategies fit for REACH that enable to significantly increase the use of non-testing information for regulatory decision making, and thus to minimize the need for animal testing.

So far, the use of non-testing methods in the European regulatory context is quite limited and fragmented. Reasons include the lack of distinct application criteria and guidance, and the fact that uncertainty has not been addressed rigorously. Industry is primarily made responsible for carrying out the risk assessments, and practical guidance is therefore needed on how to apply the elements of the newly derived testing strategies in a consistent manner.

2.3 *In silico* alternatives for application within REACH

2.3.1 General

ITSs are guidelines for the effective testing of the hazards of chemical substances, showing which tests or mathematical methods should be used for a particular substance, and in what order. ITSs should be an answer to the ever increasing demand for testing in regulations for a great number of substances with limited databases. Key for ITSs is the development of strategies on the basis of test methods at cellular level (*in vitro*) and mathematical methods (*in silico*). The mathematical methods are needed for the assessment of exposure and of the relation between adverse effects and chemical structure. In addition, some tests with experimental animals (*in vivo*) will remain necessary. Knowledge on the effects of chemical substances with sufficient certainty should be derived by smart integration of these methods. In this way it is expected that chemical substances can be assessed cheaper and faster and with less experimental animals.

Validated QSARs are powerful tools within ITSs. QSARs are used to predict the toxicity and fate of chemicals from molecular structure information, employing different levels of computational chemistry. Both categorical data (y/n) and continuous data can be addressed. Besides regulatory endpoints for classification, labelling and hazard evaluation, *in silico* methods may also generate mechanistic knowledge that guides targeted testing and enables informed extrapolation across species as well as between human health and ecotoxicology. *In silico* methods need to be optimized for regulatory purposes, and should include technical issues such as applicability domain (one of the OECD principles for QSAR validation), prediction power, and also metabolism simulators.

At the 37th Joint Meeting of the Chemicals Committee and Working Party on Chemicals, Pesticides & Biotechnology, the OECD Member Countries adopted five principles for establishing the validity of (Q)SAR models for use in regulatory assessment of chemical safety. The OECD Principles for (Q)SAR Model Validation state that to facilitate the consideration of a (Q)SAR model for regulatory purposes, the model should be associated with the following information:

1 - a defined endpoint

2 - an unambiguous algorithm

3 - a defined domain of applicability

4 - appropriate measures of goodness-of-fit, robustness and predictivity

5 - a mechanistic interpretation (if possible)

Various QSARs are available and have been operationalized in (Q)SAR software tools like ECOSAR and DEREK, whereas tools like DRAGON, OpenEye (electrostatic descriptors), ChemAxon and many others, are available for calculation of descriptor values needed in the (Q)SAR software.

As risk assessment usually boils down to comparison of predicted environmental concentrations and no effect concentrations, predictive models are needed for both fate-related endpoint and effect-related endpoints. Mackay (2007) and Mackay and Boethling (2000) presented two handbooks which focus on environmental fate prediction and QSAR analysis. The Environmental Fate Data Base of Syracuse Research Corporation (SRC, 1999) comprises a number of models for prediction of fate-related endpoints.

Models for predicting ecological effects were reviewed by Cronin et al. (2003, and references cited therein). Hermens and Verhaar (1996) are among the pioneers in the area of QSAR modelling of aquatic toxicity, based upon assessing the mode of action. These authors developed a framework that is especially applicable for modelling aquatic toxicity of organic compounds acting primarily by the mechanism of apolar narcosis as the basic mechanism of toxicity. QSARs for predicting adverse effects of chemicals acting via different modes of action are scarce. With regard to computational toxicology, a major challenge is the identification of prevalent modes of action of chemical compounds through suitable descriptors encoding local molecular reactivity. Existing parameter motifs for identifying specifically acting compounds such as electrophiles, redox cyclers and endocrine disrupters are not sufficiently specific to apply across chemical classes, and little work has been devoted to parameterizing the bioreactivity of radical intermediates.

Hulzebos and Posthumus (2003), Hulzebos et al. (2005), and Posthumus et al. (2005) provide examples of the outcome of validation efforts of QSAR models and expert systems.

2.3.2 Biodegradation and biodegradation prediction

Introduction

During production and use, organic chemicals can be released into sewers, soil, surface water, sea, and air, dumped or incinerated after use. Their fate and potential environmental hazard is strongly determined by the potential of degradability. Substances that do not degrade rapidly have a higher potential for longer term exposures and may consequently have a higher potential for causing long term adverse effect on biota and human than degradable substances. Prediction and understanding of the fate of the chemicals are therefore essential so that measures can be taken to avoid effects on humans and the environment. For this reason information on the biodegradability is used for different regulatory purposes: (1) environmental hazard classification, (2) PBT and vPvB assessment and (3) exposure assessment for use in the risk characterization.

Transformation of chemicals in the environment involves abiotic degradation and biodegradation. Abiotic degradation includes hydrolysis, oxidation, reduction, and photolysis. Biodegradation is defined as the transformation of substances caused by micro-organisms. Primary biodegradation of a molecule refers to any microbial process which leads to the formation of metabolites and thereby contributes to the degradation of the original substance. Ultimate biodegradation is known as the complete mineralization of a substance into carbon dioxide, water, and mineral salts.

Testing

In order to investigate biodegradation, standardized biodegradation tests have been developed by different organizations (amongst others: OECD, ISO, EU, US-EPA and STM), which can roughly be divided into three groups:

- screening (ready or ultimate biodegradation) tests
- intermediate (inherent or primary biodegradation) tests
- definitive (simulation) tests

Screening studies

A positive result in the screening studies can be considered as indicative of rapid ultimate degradation in most aerobic environments including biological sewage treatment plants (Struijs and Stoltenkamp, 1994) and may take away the necessity for further testing. A negative result in a test for ready biodegradability does not necessarily mean that the chemical will not be degraded under relevant environmental conditions, but it means that it should be considered to progress to the next level of testing, i.e. either an inherent biodegradability test or a simulation test.

Inherent or primary biodegradation tests

Using favourable conditions, the tests of inherent biodegradability have been designed to assess whether the chemical has any potential for biodegradation under aerobic conditions. Compared with the ready biodegradability tests, the inherent biodegradability tests are usually characterized by a high inoculum concentration and a high test substance concentration. A negative result will normally be taken as an indicator of that non-biodegradability (persistence) should be assumed for precautionary reasoning.

Simulation studies

Compared to ready and inherent biodegradability tests, simulation tests are higher tier tests that are more relevant to the real environment. These tests aim at assessing the rate and extent of biodegradation in a laboratory system designed to represent either the aerobic treatment stage of STPs or environmental compartments like surface water, sediment, and soil. They usually employ specific or semi-specific analytical techniques to assess the rate at which a substance undergoes degradation and to provide insight into subsequent metabolite formation and their decay. The fate

of chemicals in STPs can be studied in the laboratory by using the simulation tests: activated sludge units (OECD 303) and biofilms (TG 303B). Simulation tests in soil (OECD 307), in aquatic sediment systems (OECD 308), and in surface water (OECD 309) have been also included in the guidelines of OECD (OECD, 1981-2006). No specific pass-levels have been defined for the simulation tests. Simulation tests are especially useful, if it is known from other tests that the test substance can be mineralized and that the degradation covers the rate determining process.

The tests complexity and the economic consequences bond to the tests increases from the simple screening test for ready biodegradability to the more complex simulation tests. For this reason, the standard information requirements within REACH are based on the tonnage of chemicals. The requirements for these tonnage-driven degradation tests are listed in Table 2.

Table 2: REACH tonnage-driven degradation tests requirements

Tonnage (tpa)	**Degradation tests**
1-10	Ready biodegradation test
10-100	Ready biodegradation test, hydrolysis test
100-1000	Ready biodegradation test, hydrolysis test, and simulation test, identification of the most relevant degradation products
>1000	Ready biodegradation test, hydrolysis test, and simulation test. Further confirmatory testing on rates of biodegradation with specific emphasis on the identification of the most relevant degradation products

Biodegradation estimation

Under the current EU legislation for new and existing chemicals, the regulatory use of estimation models or (Q)SARs is limited and varies considerably among the member states, which is probably due to the fact that there is no agreement in the scientific and regulatory communities over the applications of (Q)SARs and the extent to which (Q)SARs estimates can be relied on. In contrast, it is anticipated that these non-testing methods like (Q)SARs and read-across, in the interests of time- and cost-effectiveness and animal welfare, will be used more extensively under the future REACH system. Below, we briefly review the current status of QSAR application for abiotic degradation and biodegradation.

(Q)SARs for biodegradation can potentially be used to supplement experimental data or to replace testing. The current generation of generally applicable biodegradation models focuses on the estimation of readily and non-readily biodegradability in screening tests. This is because most experimental data are from such tests. In the past decade, the development of QSAR modelling is mainly via three approaches: group contribution approaches, statistical/chemometric approaches, and expert system/Artificial Intelligence (AI) approaches. Table 3 summarizes the most often used QSAR models for biodegradation.

There are six models in BIOWIN. A description of these six BIOWIN models and their application for biodegradation can be found in Posthumus et al. (2005), Hulzebos et al. (2005), and Hulzebos and Posthumus (2003). Briefly, BIOWIN probability models include the linear and non-linear BIODEG and MITI models for estimating the probability of rapid aerobic biodegradation and an expert survey model for primary and ultimate biodegradation estimation. Another model is MultiCASE, which combines a group-contribution model and an expert system to simulate aerobic biodegradation pathways (Klopman and Tu, 1997). This model has also been used by Rorije et al. (1998) to predict anaerobic biodegradation. A promising model which can be used for quantitative assessment of biodegradability in biodegradation pathways of chemicals is CATABOL. The model allows for identifying potentially persistent catabolic intermediates, their molar amounts, solubility (water solubility, log K_{ow}, BCF) and toxic properties (acute toxicity, phototoxicity, mutagenicity, ER/AR binding affinity). Presently, the system simulates the biodegradability in MITI-I OECD 301 C and Ready Sturm OECD 301 B tests. Other simulators will be available in the program upgrades. The latest version of CATABOL (version 5) allows defining the degree to which chemicals belong into the domain of the biodegradation simulator.

Table 3: QSAR models for biodegradability

<table>
<tr><td rowspan="7">Group contribution approaches</td><td rowspan="6">Biodegradability probability program BIOWIN</td><td rowspan="2">BIODEG models</td><td>BIOWIN 1</td></tr>
<tr><td>BIOWIN 2</td></tr>
<tr><td rowspan="2">Expert survey models</td><td>BIOWIN 3</td></tr>
<tr><td>BIOWIN 4</td></tr>
<tr><td rowspan="2">MITI models</td><td>BIOWIN 5</td></tr>
<tr><td>BIOWIN 6</td></tr>
<tr><td>MULTICASE anaerobic biodegradation program</td><td colspan="2">To model anaerobic aquatic biodegradation rates</td></tr>
<tr><td>Statistical or chemometric approaches</td><td colspan="3">Ready biodegradability is modelled more adequately than not-ready biodegradability.</td></tr>
<tr><td rowspan="3">Expert system/AI approaches</td><td>Biodegradability evaluation and simulation system</td><td colspan="2">Qualitative, aerobic biodegradation pathways
The model needs to be validated.</td></tr>
<tr><td>MultiCASE/META</td><td colspan="2">Aerobic biodegradation pathways</td></tr>
<tr><td>CATABOL</td><td colspan="2">Quantitative assessment of biodegradability</td></tr>
</table>

An evaluation of the predictions of the models for the high production volume chemicals showed that all models are highly consistent in their prediction of not-ready biodegradability, but much less consistency is seen in the prediction of ready biodegradability. This complies with the observation that the models show better performance in their predictions of not-ready biodegradability (Rorije et al., 1999).

Degradation products - metabolites

When assessing the biodegradation of organic chemicals, it may also be needed to consider the fate and toxicity of the resulting biodegradation products, especially when they have the potency to persist in the environment. The concentration of these products in the different environmental compartments depends on numerous factors and processes, including how the parent compound is released to the environment; how fast it degrades; the half-lives of the degradation products; partitioning to sludge, soil, and sediment; and subsequent movement to air and water. In general, microbial degradation processes lead to the formation of more polar and more water soluble compounds. Hence, the resulting transport behaviour of degradation products may be different. The available data on pesticides demonstrate that in most cases degradation products are as toxic as or less toxic than the parent compounds. However, in some instances, degradation products can be more toxic. In general, the biggest increases in toxicity from parent to degradation products were observed for parent compounds that had a low toxicity. Possible explanations for an increase in toxicity are: (1) the active moiety of the parent compound is still present in the degradation product; (2) the degradation product is the active component of a pro-compound; (3) the bioaccumulation potential of the degradation product is greater than that of the parent; (4) the transformation pathway results in a compound with a different and more potent mode of action than that of the parent (Boxall et al., 2004).

Although the EU TGD (EC, 2003) highlights that, where degradation occurs, consideration should be given to the properties (including toxic effects) of the products that might arise, that information does not exist for many compounds. REACH will introduce a range of required tests which could lead to metabolite investigations, e.g. hydrolysis is required for substances produced in quantities above 10 ten tonnes per year and biodegradation simulation tests in surface water, sediment and soil are required at production volumes above 100 tonnes per year. There might be some concern that such a requirement will lead to an over emphasis on the behaviour of metabolites and that e.g. for such low production volumes, or in the case of inherently degradable substances such investigations will not be cost effective. As >100,000 chemicals are commonly used worldwide every day, pragmatic approaches are needed to identify the primary degradation products and those that are toxic, persistent, or bioaccumulative and/or which pose a risk to the environment. For this purpose guidance is needed to establish the criteria upon which metabolites

of concern may be identified and to determine when a metabolite would not be of concern (see also Vermeire et al., 2006).

2.4 CATABOL

A promising model which can be used for quantitative assessment of biodegradability in biodegradation pathways of chemicals is CATABOL. This system generates most plausible biodegradation products and provides quantitative assessment for their physicochemical properties and toxic endpoints. The possibilities of QSARs in a framework of intelligent testing strategy have been described in the previous paragraphs.

CATABOL was created to predict the most probable biodegradation pathway, the distribution of stable metabolites and the extent of biological oxygen demand or CO_2 production compared to theoretical limits. It can be considered as a hybrid system, containing a knowledge-based expert system for predicting biotransformation pathway combined with a probabilistic model that calculates probabilities of the individual transformation and overall BOD and/or extent of CO_2 production. The CATABOL system is trained to predict biodegradation within 28 days on the basis of 743 chemicals from MITI database and another training set of 109 proprietary chemicals from Procter & Gamble (P&G) obtained with the OECD 301C and OECD 301B tests, respectively. In the first database biodegradation is expressed as the oxygen uptake relative to theoretical uptake, while in the P&G database biodegradation is measured by CO_2 production.
Version 5.097 used in this study only contains information of the MITI dataset.

CATABOL is based on two sources of information:
1 - a training set containing 743 substances with measured BOD values in a MITI test.
2 - a library with transformations of chemicals fragments and their degradation products. Each transformation has a corresponding probability, which is the likelihood that a particular reaction step will be initiated.

For substances in the training set a measured BOD (y) is available, their transformation steps are based on an observed transformation scheme (for approximately 90 out of 743 substances) or on a pathway estimated by experts.
Probabilities of particular transformation steps have been derived from the training set (e.g. for a sequential pathway):

$$y = \frac{\Delta k_1}{k_{TOD}} P_1 + \frac{\Delta k_2}{k_{TOD}} P_1 P_2 + + \frac{\Delta k_3}{k_{TOD}} P_1 P_2 P_3 + \ldots + \frac{\Delta k_1}{k_{TOD}} P_1 P_2 P_3 .. P_1$$

The TOD is defined as $k_{TOD} = \sum_{i=1}^{I} \Delta k_i$

For a branched pathway:

$$y = \frac{\Delta k_1}{k_{TOD}} P_1 + \frac{\Delta k_2}{k_{TOD}} P_1 P_2 \ldots + \frac{\Delta k_I}{k_{TOD}} P_1 P_2 P_3 \ldots P_I + \frac{\Delta k_3^{'}}{k_{TOD}} P_1 P_2^{'} P_3^{'} + \\ + \ldots \frac{\Delta k_j}{k_{TOD}} P_1 P_2 P_3^{'} \ldots . P_j^{'}$$

The TOD is defined for a branched pathway as:

$k_{TOD} = \Delta k_1 + \Delta k_2 + \Delta k_3 + \ldots + \Delta k_I + \Delta k_3^{'} + \Delta k_J^{'}$, where P_i is the probability of the i^{th} transformation to be initiated.

The probabilities are subsequently used to create a hierarchy of most probable pathways and to predict BOD values for the training set. However, some transformations can be grouped because they have the same BOD and the same probability. Within these groups the hierarchy is established by expert judgement, were the effect of neighbouring groups is taken into account. For some transformations, fragments called ‘masks’ are attached to a source fragment. These inactivating fragments prevent the performance of a specific transformation. With the fitted probabilities it was possible to compute ‘predicted BOD’ for the training set. The correlation between these predicted BODs and the observed BODs was 0.9.

For some of the substances in the training set the predicted BOD did not agree with the observed BOD. These structures are ‘out of domain’. The criteria for a good prediction have been connected to the reliability for a correct prediction of readily or not readily biodegradable. The areas for false positives (wrongly predicted as ‘readily biodegradable’) and false negatives (wrongly predicted as ‘not readily biodegradable’) represent the limitations of the applicability domain.

The properties of substances in the training set are crucial in the determination of the applicability domain. The applicability domain is defined as the group of chemicals for which the model is valid.
CATABOL distinguishes three types of domains:

1 - the general parametric requirements domain
2 - the structure domain
3 - the metabolisation domain

The general parametric requirements restrict the applicability domain based upon variation of log K_{ow} and the molecular weight of the training set.

The structure domain defines the structural similarity with chemicals that are correctly predicted by the model. It is based upon the principle that the properties of a substance depend on the nature

of their atoms as well as of their arrangement. In order to check whether a new substance is in the structure domain its fragments are compared with those substances in the training set that had good BOD predictions. When the fragments of the substance of interest are not found in this group within the training set, the substance is considered 'out of domain'. The limitations in the structure domain are very dependent upon the variety of structures in the training set; substances with unknown structural fragments are by definition 'out of domain'. A technical description about how these molecular fragments are determined is described by Dimitrov et al. (2004). For substances that are 'out of structure domain' this does not mean that the structure is unknown to transformation library. A new substance, although 'out of structure domain' will be degraded according to the hierarchy and probabilities in the transformation library. However, the predicted BOD should be considered less reliable, because 'out of structure domain' only refers to the dissimilarity with substances in the training set that had a good BOD prediction.

A third domain is the 'metabolisation domain'. A list of reactions included in the library is given in Table 4. The BOD is based on those pathways that can happen on familiar fragments of the molecule and unknown fragments will remain as recalcitrant residues. Spontaneous reactions obtain a probability of 1, the probabilities of microbial reactions have been derived statistically from the training set. When a substance is 'out of the metabolisation domain', then there is no pathway (and no probability) available for a particular (sub)structure. Structures that are unknown to the library are ignored and do not contribute to the result. Consequently, CATABOL is unable to mineralize the target substance - part of degradation pathway is not generated and predicted BOD could be very wrong.

The most severe violation of the applicability domain is Metabolism Domain, followed by Structural Domain and finally the General Requirements.

Another measure of the quality of generated pathways is the reliability which is expressed in a value between 0 and 1. It is determined by making use of the reliability of transformations (their successive use versus their total use within the training chemicals). Reliability close to 1 means that all transformations used to generate a certain pathway were used correctly within the training set. The Reliability is close to 0 should be interpreted as a warning message that some of the used transformations may generate not realistic (not documented within the training set) pathways.

The interpretation of the combinations of 'high reliability and out of domain' or 'low reliability and in domain' needs some expert knowledge and should be solved case by case analysing causality for such a combination. Generally for BOD prediction: 'high reliability and out of domain (General or Structural)' is an indication that the prediction could be correct if the target chemical does not contain very 'strange' functionalities. Substances with a 'low reliability and in

domain' requires an analysis of the effect of the used transformations with low reliability on the predicted BOD.

Table 4: List of metabolic steps in the CATABOL library

1.	Spontaneous reactions	2.	Microbially catalyzed reaction
3.	Addition to ketenes and isocyanates	4.	Alkyne hydrogenation
5.	Alkaline salt hydrolysis	6.	Aromatic ring cleavage
7.	Aldehyde oxidation	8.	Acetone degradation
9.	Acyl halide hydrolysis	10.	Aromatic ring oxidation
11.	Alpha-pinene oxidation	12.	Ammonium and iminium salt
13.	Anhydride hydrolysis	14.	Alkylammonium salt
15.	Ammonium and iminium salt	17.	Alkoxysilane hydrolysis
18.	Alkoxide hydrolysis	19.	Alkylphosphinite hydrolysis
20.	Aromatic ring cleavage	21.	Azo compounds reduction
22.	Aziridine hydrolysis	23.	Oxidative deamination and N-
24.	Benzotriazole tautomerism	25.	Beta-oxidation
26.	Carbamate hydrolysis	27.	Baeyer-Villiger oxidation
28.	Cyclopropane oxidative	30.	Beckmann rearrangement
31.	Cyanuric acid isomerization	32.	Bisphenol A cleavage
33.	Diketone and unsaturated ketone	34.	Carboxylation
35.	Geminal derivatives decomposition	36.	Carbodiimide hydrolytic
38.	Hydrazine oxidation	39.	Cycloalkadiene oxidative ring
41.	Hydroxylation of substituted	43.	Diketone and unsaturated
45.	Hydroperoxide decomposition	46.	Decarboxylation
47.	Keto-enol tautomerism	48.	Dehalogenation
49.	Lactone hydrolysis and formation	50.	Diarylketone oxidation
51.	N-nitrosoamine hydrolysis	53.	Dibenzofuran oxidative
55.	Nitrate ester denitration	56.	Epoxidation
57.	Oxidative denitrification of azides	59.	Ester hydrolysis
60.	Oxirane hydration	61.	Furans oxidation
62.	Primary hydroxyl group oxidation	63.	Hexahydrotriazine hydrolytic
65.	Phosphine oxidation	66.	Imine reduction
67.	Polyphosphate decomposition	68.	Imidazole and triazole C-
69.	Quinone reduction	70.	Lactone hydrolysis
72.	Reductive deamination	73.	Methyl group oxidation
74.	Thiophosphate oxidative	76.	Nitrogroup reduction and nitrite
77.	Thiol-thion tautomerism	78.	Nitrile and amide hydrolysis
79.	Tetrahydrofuran oxidation	80.	Omega oxidation
81.	Thiol oxidation and reduction	82.	Organotin compound oxidation
83.	Thiolic acid and thioester	85.	Oxidative desulfonation
86.		87.	Oxidative thion desulfuration
88.		89.	Oxidative S-dealkylation
90.		91.	Organic sulfide S-oxidation
92.		93.	Oxidative desulfuration
94.		95.	Oxidative O-dealkylation
96.		97.	Perfluoroketone degradation
98.		99.	Pyridinium salt decomposition
100.		101.	Phosphate hydrolysis
102.		103.	Pyridine and azine ring
105.		106.	Reductive deamination
107.		108.	Sulfate hydrolysis
109.		110.	Subterminal oxidation
111.		112.	Sulfoxide reduction
113.		114.	Sulfonyl derivative hydrolysis
115.		116.	Thiol oxidation and reduction
117.		118.	Tin and lead carboxylate

2.5 Data-poor chemicals and ITSs

Data-poor chemicals cannot only be defined in a literal sense as chemicals for which few experimental fate and effect data are available, but subsequently also as chemicals for which inherently also no or only very few predictive models have been developed. As a matter of course, data-poorness hinders proper risk assessment of chemicals and necessitates the optimum use of the scarce data available by means of the tools exemplified in Figure 1. ITSs for data-poor chemicals are not existent yet but the contents of the building blocks of ITSs for data-poor substances start to surface. An example of an assessment strategy is given in Figure 2. The example deals with a framework used by Health Canada to evaluate whether metabolites are 'a cause for concern'. Metabolites are a special class of compounds as they are not the primary substances of focus, which implies that in general even fewer data are available for metabolites as for the parent compounds. The framework given in Figure 2 integrates QSAR models, read-across and testing for both the parent compound and the metabolites to end up with an assessment of the properties of the metabolite in terms of Persistence, Bioaccumulation and Toxicity.

(QSAR) models require as input one or more chemical-structure related properties. In numerous cases, this information is not available for data-poor substances. Data-poor chemicals do however share one communality, this being the availability of the chemical structure. Quantumchemical descriptors (i.e. descriptors based upon the basic properties of a chemical, like the charge of atoms in the molecule, the energy of molecular orbitals, dipole moment, polarity, the total energy of the chemical, the heat of formation, etc.) are currently made available in an increasing user-friendly mode via, amongst others, the internet. Basically, information on the chemical structure is the sole requirement for the derivation of quantumchemical descriptors. Recent progress in computational hardware and the development of efficient algorithms has assisted the routine development of molecular quantum mechanical calculations. Novel semi-empirical methods supply realistic quantum-chemical molecular quantities in a relatively short computational time frame. Quantum chemical calculations are thus an attractive source of new molecular descriptors, which can in principle express all of the electronic and geometric properties of molecules and their interactions. Indeed, many recent QSAR/QSPR studies have employed quantum chemical descriptors alone or in combination with conventional descriptors. Quantum chemistry provides a more accurate and detailed description of electronic effects than empirical methods, turning them well-suited to provide the building blocks for modelling fate and effects of data-poor chemicals.

Amongst various other chemical classes, carbamates and organophosphates are typical examples of chemicals that are widely used in a variety of applications ranging from pesticides/herbicides to flame retardants. Nevertheless, data on the fate and effects of carbamates and organophosphates

are scarce, despite their large number of applications. With regard to their aquatic toxicity, carbamates and organophosphates are known to act by a specific mode of action. No QSARs have been developed to predict their toxicity to aquatic species, let alone that estimation methods are available based on quantumchemical structure properties.

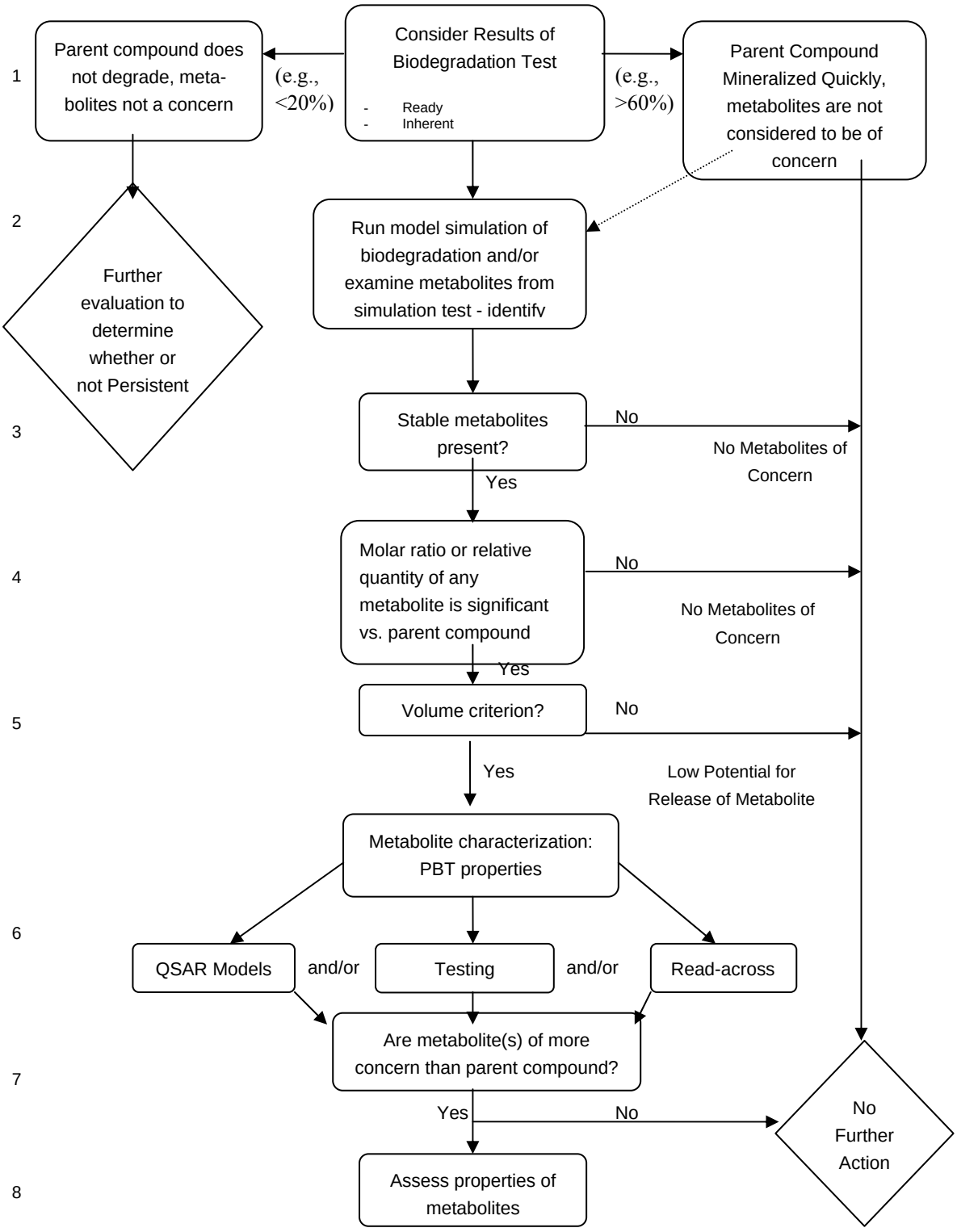

Figure 2: A framework used by Health Canada for evaluating whether metabolites are a cause for concern (slightly modified from Vermeire et al., 2006).

Carbamates or urethanes are a group of organic compounds sharing a common functional group with the general structure -NH(CO)O-. Carbamates are esters of carbamic acid, NH_2COOH, an unstable compound. Since carbamic acid contains a nitrogen atom attached to a carboxyl group it is also an amide. Therefore, carbamate esters may have alkyl or aryl groups substituted on the nitrogen, or the amide function. For example, urethane or ethyl carbamate is unsubstituted, while ethyl N–methylcarbamate has a methyl group attached to the nitrogen.

A group of insecticides also contain the carbamate functional group: for example Aldicarb, Carbofuran, Furadan, Fenoxycarb, Carbaryl, Sevin, Ethienocarb and 2-(1-Methylpropyl)phenyl N-methylcarbamate. These insecticides can cause cholinesterase inhibition poisoning by reversibly inactivating the enzyme acetylcholinesterase. The organophosphate pesticides also inhibit this enzyme, though irreversibly, and cause a more severe form of cholinergic poisoning.

Organophosphate (sometimes abbreviated OP) is the general name for esters of phosphoric acid. Phosphates are probably the most pervasive organophosphorus compounds. Many of the most important biochemicals are organophosphates, including DNA and RNA as well as many cofactors that are essential for life. Organophosphates are also the basis of many insecticides, herbicides, and nerve gases. Organophosphates are widely used as solvents, plasticizers, and EP additives. Organophosphates are widely employed both in natural and synthetic applications because of the ease with which organic groups can be linked together. In health, agriculture, and government, the word ‘organophosphates’ refers to a group of insecticides or nerve agents acting on the enzyme acetylcholinesterase (the pesticide group Carbamates also act on this enzyme, but through a different mechanism). The term is used often to describe virtually any organic phosphorus(V)-containing compound, especially when dealing with neurotoxins. Many of the so called organophosphates contain C-P bonds. For instance, sarin is O-isopropyl methylphosphonofluoridate, which is formally derived from HP(O)(OH)2, not phosphoric acid. Also many compounds which are derivatives of phosphinic acid are used as organic phosphorus containing neurotoxin.

Organophosphate pesticides (as well as Sarin and VX nerve gas) irreversibly inactivate acetylcholinesterase, which is essential to nerve function in insects, humans, and many other animals. Organophosphate pesticides affect this enzyme in varied ways, and thus in their potential for poisoning. For instance, parathion, one of the first OPs commercialized, is many times more potent than malathion, an insecticide used in combating the Mediterranean fruit fly (Med-fly) and West Nile Virus-transmitting mosquitoes.

The carbamates and OP esters used in this study are given in Table 5.

Table 5: Selected carbamates and O-P and S-P esters. Pesticides not registered in the Netherlands are indicated in bold.

O-P and S-P esters	Carbamates
Chlorfenvinphos	Aldicarb
Diazinon	**Propoxur**
Parathion, methyl	1-Naphthalenol, methylcarbamate
Fenitrothion	**Nabam**
Malathion	Carbofuran
Methylazinphos	Methiocarb
Phosphoric acid, 2,2-dichloroethenyl, dimethyl	**Phenol,2-(1-methylethyl)-**
Dimethoate	**Trimethacarb**
Parathion	**Phenol, 2-(1-methylpropyl)-,**
Fenthion	Methomyl
Dipterex	**1,3-Benzodioxol-4-ol,2,2-dimethyl-**
Ethoprophos	Pirimicarb
Profenofos	**Oxamyl**
Phosphamidon	**Butoxycarboxim**
Methamidphos	Thiodicarb - symmetrical carbamate
Demeton	**Benfuracarb**
Fonophos	
Phosmet	

3. Prediction of aquatic toxicity

3.1 General

QSAR-models were derived for predicting the aquatic toxicity of carbamates and OP-esters. The models are based on quantumchemical structure descriptors and use the molecular structure as the sole input, taking advantage of the premises that the molecular structure is the minimum amount of information always available on a chemical. Quantumchemical descriptors are calculated on the basis of the optimized geometry (i.e. energy-minimized three-dimensional structure) of the chemical. Actually, the numerical values of quantumchemical descriptors are strongly dependent on this energy-minimized three-dimensional structure. Software-packages are freely available on the internet to optimize the geometrical structure of the chemical of interest and to actually calculate descriptor values. To gain experience with this software and to make sure that the software used is indeed capable of reproducing the optimal 3-D structure, first of all an initial study was carried out in which a QSAR (based on quantumchemical descriptors) was reproduced, that has already been reported in literature. Subsequently, QSARs were developed for predicting acetyl cholinesterase inhibition of a test set of carbamates and OP-esters in fish. The basic assumption is that variances in the toxic interactions between the chemical and the fish species tested are proportional to variations in the chemical structure of the tested carbamate or OP-ester.

3.2 Operationalization of software

3.2.1 – Methods used

Before performing quantumchemical calculations, a geometry optimization was carried out with the aid of CHEMFINDER. The optimized geometry was subsequently used as input for the software package MOPAC (version 6.0: JJP Stewart, Frank J. Seiler Research Laboratory, US Air Force Academy, Co 80840) used to calculate descriptor values. MOPAC further optimizes the geometry. Chen et al. (2001) report on a QSAR based on quantumchemical descriptors. These descriptors were selected for this part of the project. The publication of Chen et al. is on a study on quantumchemical parameters to predict the rate of photolysis of dioxins and furans.

3.2.2 – Results

The results of the calculations following structure-optimization were compared to the results reported by Chen et al. (2001).

Table 6 provides an overview of the results of the calculations as carried out within this project (referred to as 'RIVM') and the results reported by Chen et al. in 2001. The following parameters were calculated for this purpose:

1 – heat of formation of the chemical (HOF – Heat of Formation)

2 – total energy of the chemical (TE - Total Energy)

3 – electronic energy (EE)

4 – the energy level of the occupied molecular orbital (Ehomo)

5 - the energy level of the lowest unoccupied molecular orbital (Elumo)

6 – the highest positive charge on a chlorine-atom (Q_{Cl})

7 – the most negative charge on the carbon atom to which the chlorine-atom indicated under 6 is connected ($Q_{Cl\text{-}C}$)

8 – the most positive charge on a H-atom (Q_{H+})

9 – the most negative charge on a carbon atom ($Q_{C\text{-}}$)

10 – the most negative charge on an O-atom ($Q_{O\text{-}}$)

11 – the average polarisability of the molecule (α)

These 11 parameters jointly give a good description of the impact of molecular structure on rates of photolysis of dioxins and furans. It is, however, evident that for the purpose of modelling the toxicity of carbamates and OP-esters, other descriptors are to be used, especially to better reflect the specific interactions between these chemicals and the receptors for acetyl-cholinesterase inhibition.

Table 6 shows that the results published by Chen et al. are well reproducible using the approach described above. The average deviation between calculated descriptor values and the values reported by Chen et al. is no more than 0.6 %. It may thus be concluded that it is not to be expected that large deviations will pop up between calculated descriptor values and their 'real value' in case of carbamates and OP-esters.

Table 6: Overview of the differences between quantumchemical descriptors for the photolysis of six dioxins and furans, as calculated by Chen et al., and as calculated in the study reported here. See body text for explanation of the symbols used. RIVM = this study, Chen = results of Chen et al., 2001.

Chemical	HOF (KJ)			TE (eV)			EE (eV)			E_{homo} (eV)			E_{lumo} (eV)			Q_{Cl}		
	RIVM	Chen	Δ (%)	RIVM	Chen	Δ (%)	RIVM	Chen	Δ (%)	RIVM	Chen	Δ (%)	RIVM	Chen	Δ (%)	RIVM	Chen	Δ (%)
2378TCDD	-32.8	-32.6	**-0.7**	-3336.1	-3336.1	**0.0**	-18441.7	-18441.7	**0.00**	-8.799	-8.800	**0.01**	-0.784	-0.785	**0.1**	0.105	0.106	**1.2**
OCDD	-47.9	-47.7	**-0.5**	-4541.2	-4541.2	**0.0**	-27151.7	-27151.5	**0.00**	-9.005	-9.009	**0.04**	-1.037	-1.037	**0.0**	0.146	0.146	**-0.1**
2378TCDF	1.6	1.8	**9.7**	-3042.7	-3042.7	**0.0**	-16417.6	-16417.1	**0.00**	-9.035	-9.033	**-0.02**	-1.074	-1.076	**0.2**	0.103	0.104	**1.0**
123678HCDF	-6.8	-6.6	**-2.7**	-3645.3	-3645.3	**0.0**	-20439.6	-20438.5	**-0.01**	-9.099	-9.099	**0.00**	-1.262	-1.262	**0.0**	0.137	0.136	**-0.7**
12347PCDD	-35.2	-35.0	**-0.6**	-3637.3	-3637.3	**0.0**	-20734.3	-20736.4	**0.01**	-8.895	-8.898	**0.03**	-0.823	-0.822	**-0.1**	0.141	0.141	**0.3**
27DCDD	-22.4	-22.2	**-0.9**	-2733.4	-2733.4	**0.0**	-14957.2	-14957.4	**0.00**	-8.764	-8.768	**0.05**	-0.523	-0.522	**-0.2**	0.076	0.076	**0.6**

Chemical	Q_{Cl-C}			Q_{H+}			Q_{C-}			Q_{O-}			α (atomic units)		
	RIVM	Chen	Δ (%)	RIVM	Chen	Δ (%)	RIVM	Chen	Δ (%)	RIVM	Chen	Δ (%)	RIVM	Chen	Δ (%)
2378TCDD	-0.136	-0.136	**0.1**	0.138	0.138	**0.3**	-0.136	-0.136	**0.1**	-0.092	-0.094	**1.9**	159.6	159.5	**-0.01**
OCDD	-0.138	-0.138	**-0.2**				-0.140	-0.140	**-0.1**	-0.075	-0.075	**0.2**	200.7	200.6	**-0.02**
2378TCDF	-0.112	-0.112	**0.0**	0.138	0.138	**0.2**	-0.161	-0.161	**0.1**	-0.073	-0.074	**1.2**	157.2	157.2	**0.02**
123678HCDF	-0.143	-0.142	**-0.4**	0.140	0.140	**-0.2**	-0.174	-0.174	**-0.2**	-0.061	-0.061	**-0.4**	177.3	177.3	**0.01**
12347PCDD	-0.140	-0.140	**0.2**	0.138	0.127	**-8.5**	-0.141	-0.141	**0.1**	-0.085	-0.085	**0.3**	167.9	167.9	**-0.02**
27DCDD	-0.124	-0.124	**-0.3**	0.136	0.136	**0.2**	-0.124	-0.124	**-0.3**	-0.095	-0.095	**-0.3**	135.0	134.9	**-0.04**

3.3 Development of QSARs for acetylcholinesterase inhibition of carbamates and organophosphate esters

1 - Introduction

Carbamates and OP-esters are classes of chemicals for which the toxicity stems from their acetylcholinesterase-inhibiting mode of action. Acetylcholinesterase is an enzyme that makes sure that stimules of the central nerves caused by any stress factor and that are passed on to the brain, is stopped following termination of the stress. Acetylcholinesterase inhibitors like carbamates and OP-esters limit the functioning of the enzyme, causing the nerves to remain stimulated. Adverse effects are observable at low doses. Interactions between the receptors for acetylcholinesterase inhibition and the carbamates and OP-esters result from specific polar, hydrophobic and steric interactions. Thus, the acetylcholinesterase inhibiting action of carbamates and OP-esters is more specific than the general mechanism of polar narcosis. It is to be expected that molecular descriptors that are directly or indirectly related to the interactions between the active substance and the receptor, are adequate to model experimentally observed differences in toxicity between the various substances tested, i.e. descriptors that provide information on the charge distribution in the toxicant and differences in the energy content of the test compound. In this study it was investigated to which extent these types of descriptors can be used to predict the toxicity of carbamates and OP-esters. Sterical descriptors were not taken into account, but these could provide additional possibilities of increasing the predictive capabilities of the QSARs reported here.

The research was split up in three steps:

1 – Selection of toxicity data. The main source used was the database collected within the Laboratory for Ecological Risk assessment RIVM on aquatic toxicity data. The database contains a total of 595 toxicity data on carbamates and 2369 data on OP-esters. The toxicity of both chemical classes was assessed for a large number of species but the number of chemicals tested per species, is limited. This limits the possibility of selecting a sufficiently large test set of uniformly determined toxicity data, suited for QSAR-modelling. For four species, relatively many data are available in the database: *Crassostrea virginica* (Oyster), *Aedes aegypti* (midge larvae), *Cyprinus carpio* (carp) and the fish *Brachydanio rerio*. This implies that toxicity data are available for three trophic levels.

2 – For all carbamates and OP-esters for which one or more toxicity data were available, the quantumchemical properties explained above were calculated using the methods described in this paragraph.

3 – For each organism, the experimental toxicity data were correlated to the calculated quantumchemical descriptors. This was done by means of linear regression, using the statistical software package SYSTAT. Principle Component Analysis (PCA) was used apart from linear regression. PCA is a more advanced data processing technique, capable of visualizing non linear (multivariate) links between datasets. The number of toxicity data for each organism was however too small to retrieve statistically significant relationships by means of PCA. Multivariate data processing will therefore not be discussed any further in this report. Please note that this does not mean that multivariate techniques do not have an added value over regression analysis.

Note ad 1 and 2:
The procedure commonly employed in developing QSARs, involves a reverse order of the steps 1 and 2, i.e. first a suited test set of chemicals to be tested is selected, and subsequently the toxicity data (or data related to other endpoints of interest) are collected or generated. The main advantage of this latter procedure is that the final QSAR to be developed is applicable to a larger range of structurally well-defined chemicals (the so-called chemical domain). The chemical domain is limited in this study due to the initial approach taken here as the domain is determined by the properties of the carbamates and OP-esters for which experimental toxicity data were available. The lack of toxicity data for these important classes of chemicals was the main limitation in this respect. This limitation on the other hand clearly shows the need of applying predictive methods to enable proper estimation of the risk of chemicals within these classes of chemicals.

Appendix 1 contains an overview of the toxicity data used. Thereupon, this appendix contains the numerical values of the quantumchemical descriptors calculated. Appendix 1 shows that the set of toxicity data for each chemical differs for each organism considered, both with regard to the number of chemicals as with regard to the variance in structure properties. Appendix 1 shows that in case of OP-esters, toxicity data are available for both O-substituted and S-substituted phosphate esters. It is to be expected that separate models are to be developed for both substance classes with clearly distinct properties.

2 – Results

Available toxicity data

Appendix 1 shows that toxicity data are available for a limited number of chemicals. Apart from the obvious necessity of having available a sufficient test set of toxicity, derivation of predictive models also requires variance in the observed effect data.

Most data are available for *Cyprinus carpio*: for 8 carbamates and 18 (7 O-P esters, 10 S-P esters, while 1 compound was both an O-P and an S-P ester). EC50-values range in between 0.5 – 3.5 mg L^{-1} for the carbamates, from 0.0015 tot 69.01 mg L^{-1} for the O-P esters and from 0.002 - 4.65

mg L^{-1} for the S-P esters. This implies a range of EC50-values of a factor of 7, 46007 and 2325 respectively, which in turn implies *a priori* that the impact of chemical structure on the aquatic toxicity of carbamates is limited.

EC50-data for the larvae of the midge *Aedes aegypti* are available for only 4 carbamates and 5 organophosphate esters (all S-P esters). Apart from the limited number of toxicity data, also the variance in measured toxicity is limited for the carbamates tested (0.09 – 0.38 mg L^{-1}). The toxicity of the S-P esters varies in between 0.00265 and 5.35 mg L^{-1} (factor 2019), which implies a clear impact of molecular structure on toxicity.

No data on carbamate toxicity are available for the oyster *Crassostrea virginica.* The toxicity of 10 OP-esters (3 O-P esters and 7 S-P-esters) varied in between 0.33 and 9.07 mg L^{-1} (factor 27).

Finally, also for the fish *Brachydanio rerio* no data were found on carbamate toxicity. EC50-values of 6 OP-esters (5 of which were S-P-esters) vary in between 0.0012 and 100 mg L^{-1} (factor 83333).

It may be concluded from these data that the number of toxicity data is limited for especially carbamates. Also, the variance in toxicity is limited for this class of chemicals. Although the number of toxicity data is limited for OP-esters, there is a clear impact of molecular structure on the toxicity of these compounds.

Modelling of toxicity

Log transformed values of the toxicity data collected and log-transformed values of the quantumchemical descriptors were used as the basis for obtaining relationships between the structure of the carbamates and the OP-esters, and the measured toxicity. Log-transformation is needed to meet one of the primary requirements of linear regression, i.e. a linear distribution of the datasets used.

A - *Cyprinus carpio*

A1 - Carbamates

Toxicity data are available for 8 carbamates. As indicated in the previous paragraph, EC50 values range in between 0.5 and 3.5 mg L^{-1}. In part related to this limited variance in cholinesterase inhibition, no statistically significant relationship was found between individual quantumchemical descriptors and carbamate-toxicity. A closer analysis reveals a significant relationship between (log-transformed) toxicity and a combination of two descriptors: the (log-transformed) dipole moment of the toxicant and the (log-transformed) sum of the electronic charges on the atoms of the carbamate-moiety. Jointly, these descriptors explain about 55 % of the variance in the measured toxicity. As a rule of thumb it has been postulated that a QSAR is suited for risk assessment when the QSAR is capable of explaining over 50 % of the variance in the data. The

correlation between on the one hand the measured toxicity of eight carbamates and on the other hand the dipole moment of the toxicant and the sum of the electronic charges of the carbamate moiety, thus satisfies this requirement. The observation of dipole moment and charge distribution being of dominant importance in explaining toxicity, confirms previous findings that electronic interactions play an important role in the mechanisms of toxicity of carbamates for *Cyprinus carpio*.

Equation 1 gives both the multiple regression equation obtained as well as the most relevant statistical parameters. Figure 3 graphically displays the relationship between observed and predicted toxicity for *Cyprinus carpio* for the eight carbamates investigated.

Log EC50 (mg L^{-1}) = 3.56 – 0.16 * Log Dipole moment (Debeye) + 8.2 * Log (Sum of the charges of the atoms of the carbamate functionality) (1)

R^2 = 0.55, p-value = 0.08, F-value = 2.41

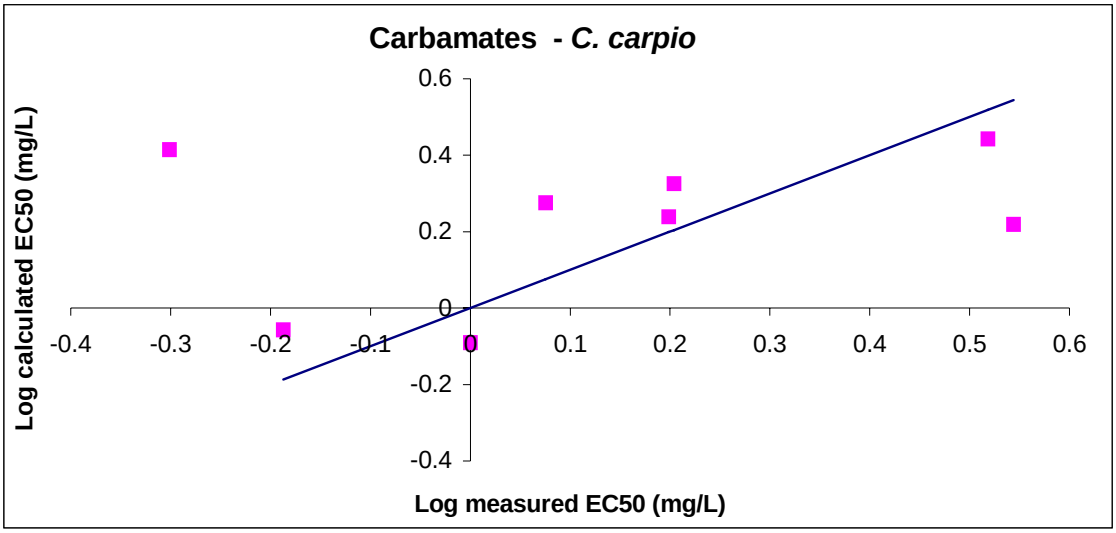

Figure 3: Relationship between measured cholinesterase inhibition in *Cyprinus carpio* by carbamates (x-axis: log-transformed EC50-values) and predicted log-transformed EC50-values (y-axis), using equation 1.

A2 – Organophosphate esters

Again, following log-transformation of the data no statistically significant linear relationship was found between any of the individual descriptors and the experimental toxicity data for the whole dataset of organophosphate esters. As it is to be expected that due to intrinsic differences between the properties of S-P and O-P esters, there is an intrinsic difference in the impact of structure on toxicity, the whole set of toxicity data was split up in data for these two categories.

The best prediction of toxicity in case of O-P esters is obtained by a combination of the (log-transformed) heat of formation and the (log-transformed) electronic energy of the O-P esters as descriptors. Equation 2 shows the relationship observed as well as the relevant statistical

parameters, figure 4 provides a graphical representation of the relationship between measured and calculated toxicity when using equation 2.

Log EC50 (mg L^{-1}) = 23.95 – 1.22 * Log Heat of Formation (KJoule) –
4.93 * Log electronic energy (eV) (2)

R^2 = 0.79, p-value = 0.02, F-value = 7.38

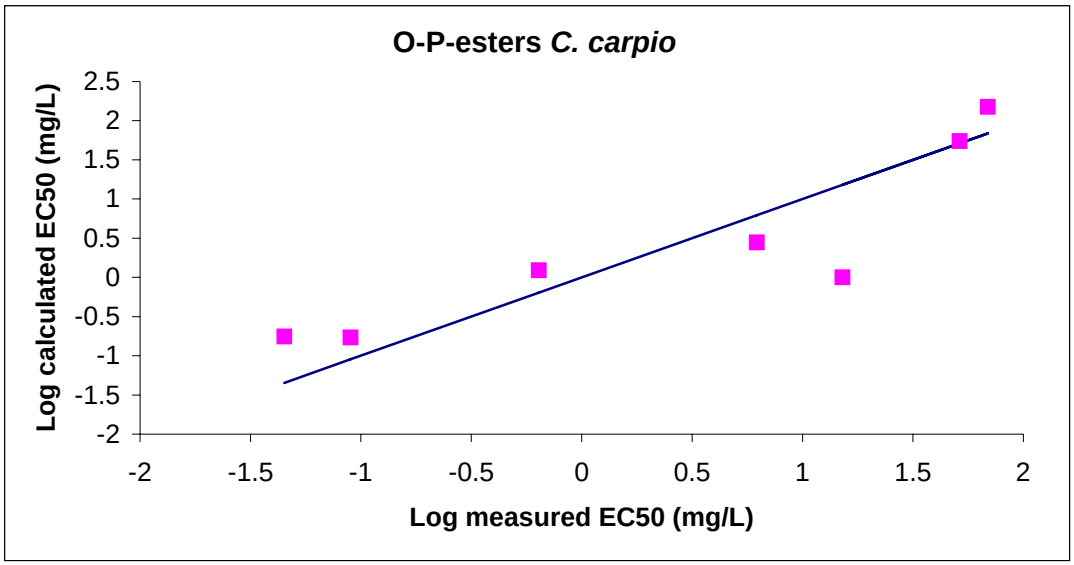

Figure 4: Relationship between measured cholinesterase inhibition in *Cyprinus carpio* by O-P esters (x-axis: log-transformed EC50-values) and predicted log-transformed EC50-values (y-axis), using equation 2.

The descriptors that best describe toxicity of S-P esters are electronic energy and the sum of the atomic charges at the S- and P-atoms of the esters. Equation 3 provides the relationship found as well as the relevant statistical parameters, figure 5 graphically displays this relationship by comparing observed and predicted toxicity.

Log EC50 (mg L^{-1}) = 29.71 – 7.34 * Log electronic energy (eV) + 8.98 * Log (sum of the charges of the atoms at the S-P moiety) (3)

R^2 = 0.55, p-value = 0.04, F-value = 4.26

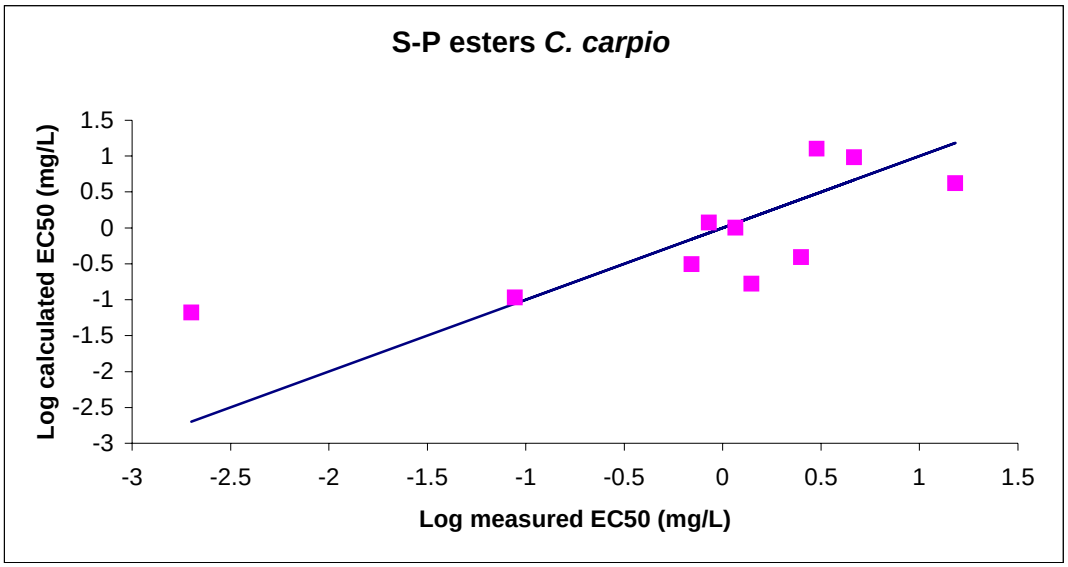

Figure 5: Relationship between measured cholinesterase inhibition in *Cyprinus carpio* by S-P-esters (x-axis: log-transformed EC50-values) and predicted log-transformed EC50-values (y-axis), using equation 3.

B - *Aedes aegypti*

B1 - Carbamates

Toxicity data are available for four compounds. These data correlate quite well ($R^2 > 0.8$) with various descriptors. It should be noted however that the significance of relationships is strongly reduced by the limited number of data. Equation 4 illustrates the relationships between individual descriptors and toxicity, in this case exemplified using the log-transformed Heat of Formation of the carbamates as molecular descriptor. Figure 6 shows the predictive capability of this descriptor, by comparing measured and predicted toxicity on the basis of this descriptor. Given the limited number of data, no significant relationships were found when applying multiple linear regression.

$$\text{Log EC50 (mg L}^{-1}) = 29.71 - 7.34 * \text{Log Heat of Formation (kJ)} \quad (4)$$

$R^2 = 0.88$, p-value = 0.06, F-value = 15.0

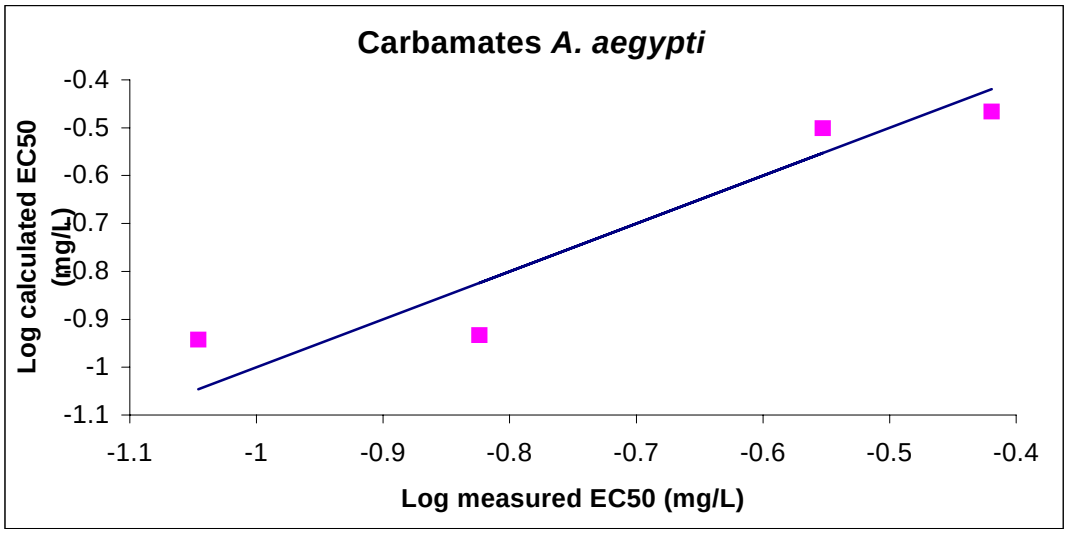

Figure 6: Relationship between measured cholinesterase inhibition in *Aedes aegypti* by carbamates (x-axis: log-transformed EC50-values) and predicted log-transformed EC50-values (y-axis), using equation 4.

B2 – Organophosphate esters

Toxicity data are available for five S-P esters. Similar to the carbamates, these data correlate well with various individual descriptors. Equation 5 illustrates the relationship for the example of Dipole moment as predictive molecular property. Figure 7 shows the predictive capability of this descriptor.

$$\text{Log EC50 (mg L}^{-1}\text{)} = 8.00 - 9.27 * \text{Log Dipole moment (Debye)} \quad (5)$$

$R^2 = 0.83$, p-value = 0.03, F-value = 14.3

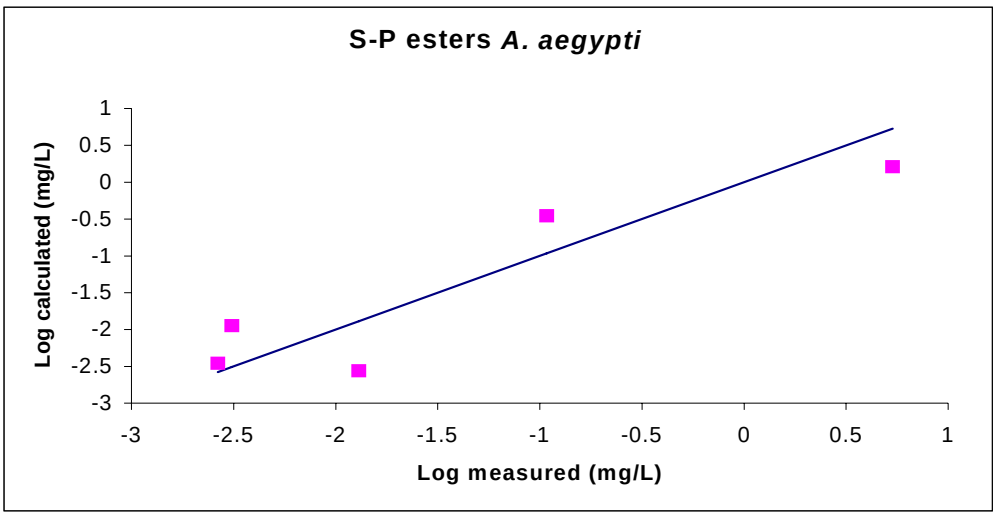

Figure 7: Relationship between measured cholinesterase inhibition in *Aedes aegypti* by organophosphate esters (x-axis: log-transformed EC50-values) and predicted log-transformed EC50-values (y-axis), using equation 5.

C - *Crassostrea virginica*

C1 – Carbamates

No toxicity data are available in the database.

C2 – Organophosphate esters

Toxicity data are available for 3 O-P-esters and 7 S-P-esters. The toxicity data for the whole dataset are not significantly correlated to any single molecular descriptor. A combination of three descriptors yields the most significant relationship:

Log EC50 (mg L^{-1}) = -50.16 + 14.7 * Log Ionisation potential – 3.21 * Log Dipole moment + 8.34 * Log Electronic energy (6)

R^2 = 0.66, p-value = 0.12, F-value = 3.2

Figure 8 shows the relationship between measured and calculated toxicity when using equation 6 for the whole dataset of organophosphate esters.

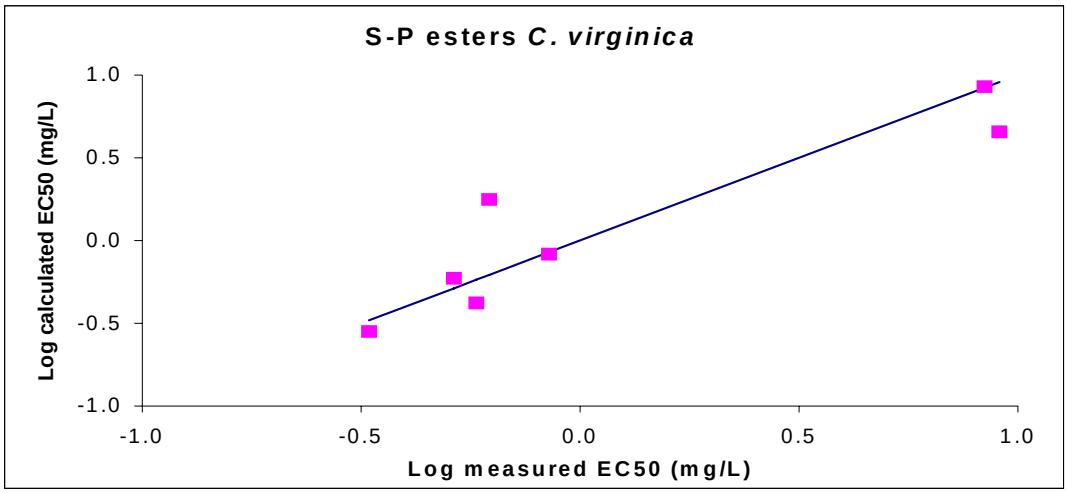

Figure 8: Relationship between measured cholinesterase inhibition in *Crassostrea virginica* by organophosphate esters (x-axis: log-transformed EC50-values) and predicted log-transformed EC50-values (y-axis), using equation 6.

Splitting of the dataset yields a subset of three O-P esters, which is too small for the derivation of QSARs. On the other hand, the toxicity of the seven S-P esters is well correlated to the dipole moment of the esters and their electronic energy (equation 7 and figure 9):

Log EC50 (mg L^{-1}) = -25.85 – 1.07 * Log Dipole moment + 6.25 * Log Electronic energy (7)

R^2 = 0.85, p-value = 0.02, F-value = 11.1

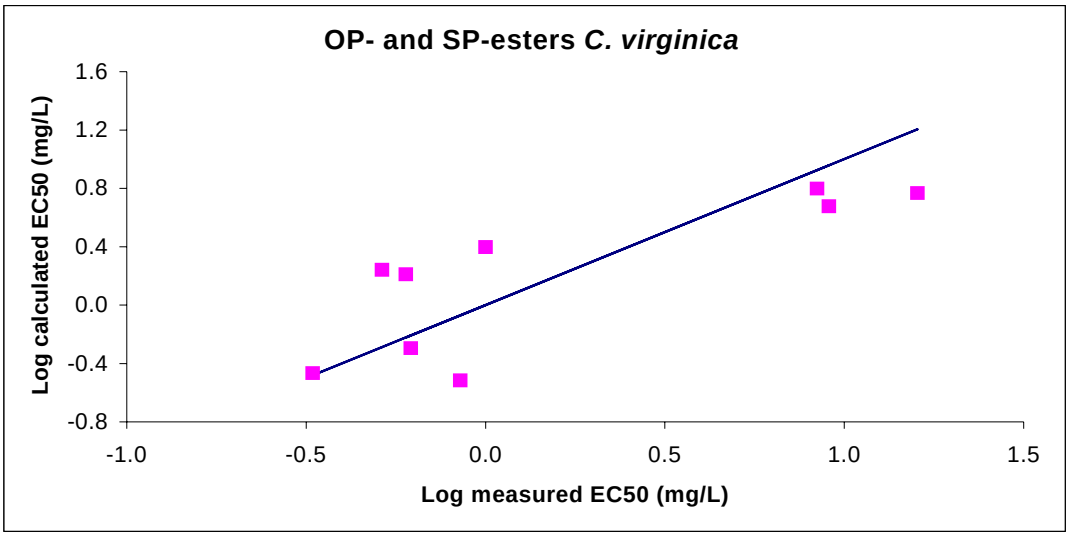

Figure 9: Relationship between measured cholinesterase inhibition in *Crassostrea virginica* by S-P esters (x-axis: log-transformed EC50-values) and predicted log-transformed EC50-values (y-axis), using equation 7.

D - *Brachydanio rerio*

D1 – Carbamates

No toxicity data are available in the database.

D2 – Organophosphate esters

A limited dataset of six data was found. The descriptors calculated in this study correlate to a varying extent of statistical significance with the observed toxicity of the chemicals in the dataset. The energy level of the highest occupied molecular orbital, heat of formation and dipole moment are examples of descriptors that are capable of explaining the variance in the toxicity data for *Brachydanio rerio* to a large extent. The most significant relationship is obtained by a combination of the energy level of the highest occupied molecular orbital (Ehomo) and the difference in electronic charges between the P- and either the O- or the S-atoms of the organophosphate ester (equation 8 and figure 10):

$$\text{Log EC50 (mg L}^{-1}) = 124.7 + 11.7 * \text{Log Ehomo} - 14.4 * \text{Log } (\Delta \text{ electronic charges at P and O/S-atoms}) \quad (8)$$

R^2 = 0.95, p-value = 0.01, F-value = 30.9

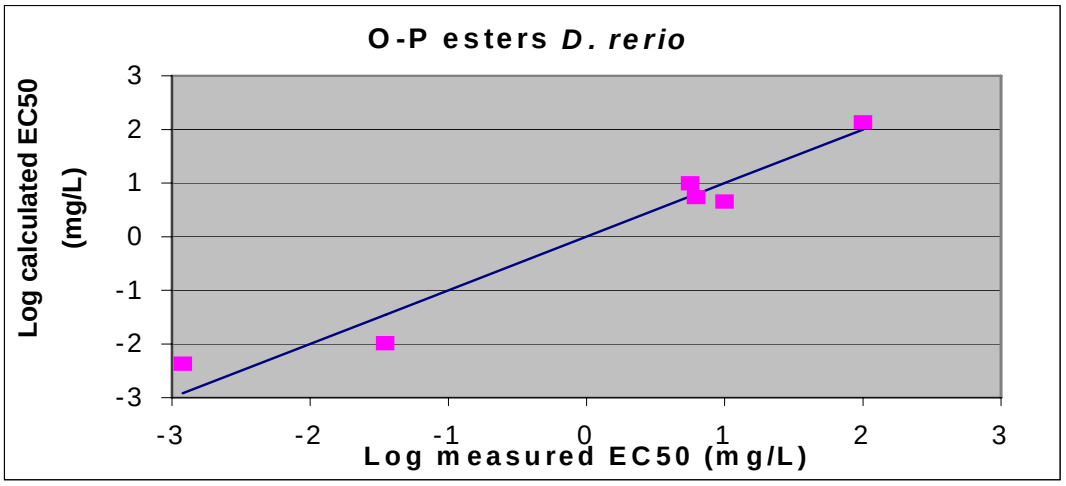

Figure 10: Relationship between measured cholinesterase inhibition in *Crassostrea virginica* by six organophosphate esters (x-axis: log-transformed EC50-values) and predicted log-transformed EC50-values (y-axis), using equation 8.

3 – Conclusions and recommendations

Conclusions

The database containing toxicity data that is used as the basis for this study, is to be considered as being representative for the number of aquatic toxicity data reported in literature up till now. This is despite the obvious fact that the database does not contain 'all' toxicity that have ever been measured. The inventory does show, however, that the number of toxicity data available for both carbamates and organophosphate esters, is quite limited. This is worrisome given the fact that we are dealing here with two classes of specifically acting compounds that are toxic to large numbers of aquatic species.

Within REACH the necessity is proclaimed of assessing the risk associated with the production and use of large numbers of chemicals, given the pre-condition of minimizing the use of test animals and optimizing the testing strategies. The lack of sufficient toxicity data for chemicals like carbamates and OP-esters clearly shows the necessity within the REACH-legislation of having methods available for estimating toxicity. Preferably, these estimation routines are based on easily deducible descriptors. Quantumchemical parameters in this respect have the advantage of being easy retrievable without any additional experimental effort: the software needed is increasingly getting more user-friendly, it is freely available on the internet, the precision and robustness of the parameter values calculated is increasing, and the number of parameters that is to be calculated, is increasing. Potentially, quantumchemical descriptors are thus suited to quickly and reliably estimate the toxicity of large numbers of chemicals.

The number of predictive methods based upon quantumchemical descriptors is currently still limited. It is shown in this study that it is well feasible to generate predictive routines for substances belonging to the chemical classes of carbamates and OP-esters. These QSARs are

aquatic species-specific, whilst their robustness and predictive capabilities are strongly dependent on the number of toxicity data available. The relationships are also chemical class-specific, albeit that descriptors based upon the energy-content of the toxicants and descriptors related to the charge distribution within the functional moiety in general are determining toxicity for both chemical classes investigated. This finding confirms the general similarity (specific interactions with the receptors of toxicity) of carbamates and OP-esters.

Recommendations

Quantumchemical descriptors are shown in this study to have an added value on top of existing (partly experimentally derived) descriptors for calculating the toxicity of non-tested analogue chemicals. Combining quantumchemical descriptors with descriptors of other natures (like information on the size of the toxicant, the three-dimensional structure of the chemical, and more specific indicators for the interactions between the toxicant and the receptors of toxicity) would allow for further optimization of the predictive capabilities. Thus, the specific compounds class-related interactions of chemicals like carbamates and organophosphate esters with the receptors of toxicity would be modelled more optimally.

Apart from further optimisation of the QSARs for predicting toxicity of carbamates and OP-esters, a similar approach may be taken to predict toxicity of other chemicals with a specific mode of action. This initiates the construction of a module within ITSs for assessing the risks of data-poor chemicals to be used within REACH for estimating toxicity of specifically acting toxicants. This will in general involve polar and ionogenic chemicals like metabolites of pesticides, endocrine disrupting chemicals, and pharmaceuticals and veterinary products, each of which constitutes a diverse set of chemical classes for which limited numbers of toxicity data are available.

4. Validation of CATABOL

4.1 General

CATABOL is in principle applicable to a wide array of chemical substances that on forehand do not share structural features. Given the current status of CATABOL as being in the later stages of development and still prone to thorough validation, application of the tool on a diverse set of chemicals is unlikely to increase the insight in the predictive capabilities. With one of the aims of testing its applicability, we therefore applied CATABOL to a well-defined domain of chemical structures and used CATABOL to predict metabolite formation of 16 carbamates and 18 organophosphate esters. These two classes of chemicals were selected as they constitute two classes of chemicals that, given their emission patterns, and given their fate and effect profiles, may pose a risk for the environment. Thereupon, as described in Chapter 3, QSARs based on quantum chemical descriptors were developed to predict their toxicity for aquatic organisms. The dataset used for the development of these QSARs was also used in the CATABOL-study reported here. In addition to prediction of metabolite-formation by CATABOL, databases were searched to find experimental data on metabolite formation in soil of the two compound classes, and the toxicity of the measured and/or predicted metabolites was assessed using the previously developed QSARS. The database used as the source for experimental data on metabolite formation was: http://www.herts.ac.uk/aeru/footprint/. The website provides the following indication of its content: 'The best sources of information currently available for pesticide properties are the monographs produced as part of the EU review process and these documents have been used in priority for putting together the FOOTPRINT PPDB. Where EU documents were not available, alternative sources were used:

- databases and documents from various national government departments including the UK's PSD, Germany's Federal Environment Bureau, the EPA in the USA and the French Agritox database;
- on-line databases including ARIS, EXTOXNET, ARS/OSU, PAN, GLEAMS, etc.;
- manufacturers safety datasheets and environmental fact sheets, on- and off-line;
- publications such as the Pesticide Manual;
- data derived from research projects such as the Pandora data set;
- peer reviewed scientific publications.

In a very limited of instances, data had to be retrieved from miscellaneous on-line sources. All data held in the FOOTPRINT PPDB are 'tagged' with a code so that their source and quality can be identified (see below).'

Evaluation of the toxicity profile of the metabolites formed (or predicted to be formed) is an essential part of the proposed framework for evaluating whether metabolites are a cause for concern. This framework is depicted in Figure 2 of paragraph 2.4.

4.2 Results

Experimental data

Experimental data on metabolite formation were available for ten out of the sixteen (63 %) carbamates, which implies that experimental data were lacking for the remaining six compounds (37 %). In case of organophosphates, experimental data were lacking for twelve of the eighteen compounds considered (67 %), clearly highlighting the need of having available estimation methods for predicting metabolite formation.

CATABOL predictions

Twelve of the sixteen carbamates studied (75 %) were out of the structure domain of the CATABOL model and hence also out of the total domain. Experimental BOD-data from the MITI-test were available for three of the four carbamates that were in the structure domain. Fourteen of the eighteen organophosphorus esters (78 %) were out of the structure domain of the model and hence also out of the total domain. Experimental BOD-data from the MITI-test were available for three of the four organophosphates that were in the structure domain.

Comparison between CATABOL predictions and experimental data

CATABOL in general predicts the formation of a large number of stable and less stable intermediates. This is done according to a hierarchic systematic. When comparing predictions and experimental data, it is important to realize that reports on metabolites observed during biodegradation testing in soil, are not always encompassing 'all' metabolites formed. Apart from analytical constraints (detection limits), there often is bias with regard to the expected metabolites formed. The latter implies that not all metabolites predicted to be formed by CATABOL were taken into consideration at the time of deciding on which metabolites to focus on during the experimental degradation study. When comparing predicted and observed metabolites, it is important to take note of these (and additional) considerations. In this study we therefore did not focus on matching individual metabolites. Instead the focus was on deducing from the molecular structure of the metabolites reported to be formed during biodegradation in soil, whether these metabolites could be formed as the outcome of the pathways predicted by CATABOL. Or, in other words, the focus was not on individual compounds but on degradation pathways, investigating the possibility of the reported metabolites having the possibility of being formed along one of the degradation pathways predicted by CATABOL to take place. To provide an example: a common transformation step for S-P-esters containing a thiobenzene-group is the

oxidation of the S-atom attached to the benzene-moiety. This transformation leaves the S=P linkage intact and yields the corresponding sulfoxide as a stable metabolite. CATABOL, however, predicts the oxidation of the P=S moiety to yield the corresponding P=O ester, with a probability of 1. In this case, the experimentally observed S-P esters, oxidized at the benzenic S-atom cannot be formed along the CATABOL-predicted pathway of S=P oxidation. This is graphically illustrated in Figure 11.

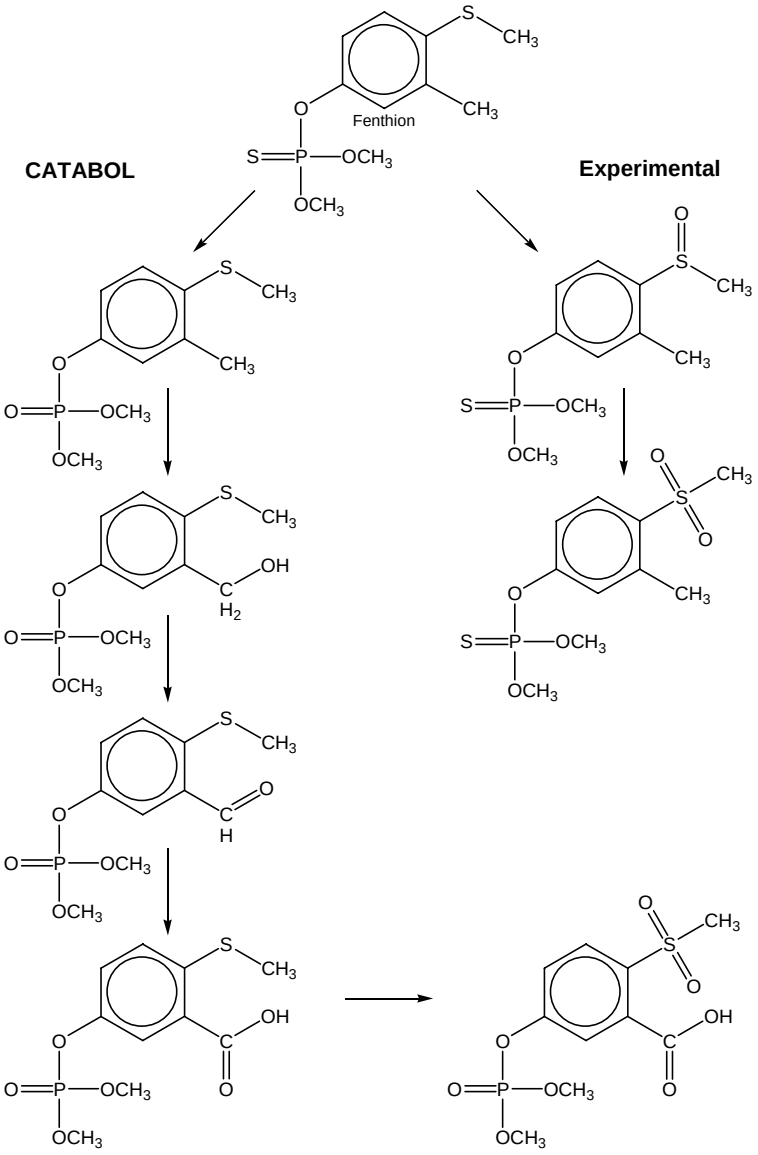

Figure 11: Illustration of the difference between the CATABOL-predicted metabolites of fenthion and the metabolites found in soil.

Carbamates

CATABOL-predictions matched the experimental findings for three out of the ten compounds for which experimental data were available: for seven of the carbamates studied with experimental data on metabolite formation, the experimentally observed metabolites cannot be formed along

the pathways predicted by CATABOL. Of the three compounds for which the CATABOL predictions were correct, one carbamate was in the structural domain and two were out of domain.

Organophosphates

CATABOL-predictions matched the experimental findings for two out of the six compounds for which experimental data were available. This was not the case for four organophosphates studied (67 %). Of the two compounds for which the CATABOL predictions were correct, one was in the structural domain and one was out of domain.

Overall

In Table 7 a more detailed comparison is given of the metabolites predicted to be formed by CATABOL (indicated as 'est.'), and the metabolites actually found in soil (indicated as 'exp.').

Table 7: Comparison between CATABOL predicted metabolites and metabolites found in soil. Est. = Predicted by CATABOL, exp. = observed metabolite in soil.

Carbamates	**Transformation**	**Remarks**
1-Naphthalenol, Methylcarbamate	1-Naphthalenol, methylcarbamate	
Metabolite 1 - est.	1-Naphthalenol	This is no carbamate: cleavage of the carbamate functionality
Metabolite 1 - exp.	1-Naphthalenol	Prediction correct
2-(1-Methylethoxy) phenol, Methyl carbamate	Propoxur	
Metabolite 1 - est.	Propoxur, hydroxylated at the 2- and 3-position next to the carbamate moiety	Prediction incorrect, this metabolite cannot be formed
Metabolite 1 - exp.	2-isopropoxyphenol - formed by cleavage of the carbamate moiety	out of the CATABOL-predicted metabolites
2-Methyl-2-(methylthio)propional dehyde, O-(Methyl-carbamoyl)oxime	Aldicarb	
Metabolite 1 - est.	Hydrolysis of the C=N bond: C=N becomes OH-C-N-H	Prediction incorrect

Metabolite 2 - est.	Cleavage of the carbamate functionality of metabolite 1	Prediction incorrect. This is the metabolite in which the S-atom is oxidized once: (S=O) bond
Metabolite 1 - exp.	Aldicarb sulfoxide	
Metabolite 2 - exp.	Aldoxycarb	This is the metabolite in which the S-atom is oxidized twice: (O=S=O) bond
1,2-Ethanediylbis-carbamothioic acid, disodium salt	Nabam	
Metabolite 1 - est.	H-S-C=S replaced by H-S-C=O	
Metabolite 2 - est.	H-S-C=O replaced by H-O-C=O	
Metabolite 3 - est.	Cleavage of the carboxylic group	
Metabolite 1 - exp.	No information available	
2,3-Dihydro-2,2-dimethyl-7-benzofuranol, Methylcarbamate	Carbofuran	
Metabolite 1 - est.	Cleavage of the carbamate functionality of carbofuran; yields the alcohol	Prediction incorrect
Metabolite 1 - exp.	3-hydroxycarbofuran	
Metabolite 2 - exp.	3-ketocarbofuran	
3,5-Dimethyl-4-(methylthio)phenol, Methylcarbamate	Methiocarb	
Metabolite 1 - est.	Cleavage of the carbamate functionality of methiocarb; yields the alcohol	Prediction incorrect
Metabolite 1 - exp.	3,5-dimethyl-4-(methylsulfinyl)phenol - cleavage carbamate functionality and oxidation C-S bond	
Metabolite 2 - exp.	p-(methylsulfonyl)phenol - oxidation C-S bond and cleavage of 2 CH_3-groups	
Methylcarbamate 2-(1-methylethyl) phenol	Phenol,2-(1-methylethyl)-, methylcarbamate	
Metabolite 1 - est.	Hydroxylation of the aromatic ring yields the diol	

Metabolite 2 - est.	Further hydroxylation of the aromatic ring yields the triol	
Metabolite 1 - exp.	No info in database	
Trimethacarb	N-Me-3,4,5-triMePhenyl carbamate	
Metabolite 1 - est.	Cleavage of the carbamate functionality	
Metabolite 1 - exp.	Cleavage of the carbamate functionality	Prediction correct
2-(1-Methylpropyl) phenol, Methyl-carbamate	Phenol, 2-(1-methylpropyl)-, methylcarbamate	
Metabolite 1 - est.	Cleavage of the carbamate functionality	
Metabolite 1 - exp.	No info in database	
N-[[(Methylamino) carbonyl]oxy]ethanim idothioic acid methyl ester	Methomyl	
Metabolite 1 - est.	Cleavage of the carbamate functionality	
Metabolite 1 - exp.	No info in database	
2,2-Dimethyl-1,3-benzodioxal-4-ol methylcarbamate	1,3-Benzodioxol-4-ol,2,2-dimethyl-, methylcarbamate (bendiocarb of isoprocarb)	
Metabolite 1 - est.	Hydroxylation of the aromatic ring yields the diol	Prediction incorrect
Metabolite 1 - exp.	2,2-dimethyl-1,3-benodioxol-4-ol: hydrolysis carbamate moiety	
Dimethylcarbamic acid, 2-(Dimethylamino)-5,6-dimethyl-4-pyrimidinyl ester	Pirimicarb	
Metabolite 1 - est.	Hydrolysis of the carbamate moiety	
Metabolite 1 - exp.	Hydrolysis of the carbamate moiety as primary transformation	Prediction correct
2-(Dimethylamino)-N-[[methylamino-carboxy]oxy]-2-oxo, Methyl ester ethan-imidothioic acid	Oxamyl	
Metabolite 1 - est.	Hydrolysis of the carbamate moiety	
Metabolite 1 - exp.	No info in database	
butoxycarboxim	Butoxycarboxim	

Metabolite 1 - est. Metabolite 1 - exp.	Hydrolysis of the C=N bond: C=N becomes OH-C-N-H No info in database	
N,N'-[Thiobis [(methylimino)carbonyloxy]]bisethanimidothioic acid, Dipentyl ester	Thiodicarb - symmetrical carbamate	
Metabolite 1 - est.	Hydrolysis of one of the carbamate groups - cleavage of O-N-binding	Prediction incorrect
Metabolite 1 - exp.	Methomyl - formed by cleavage of N-S-bond in middle of the molecule	
2-Methyl-4-(1-methylethyl)-7-oxo-8-oxa-3-thia-2,4-diazadecanoic acid, 2,3-Dihydro-2,2-dimethyl-7-benzofuranyl ester	Benfuracarb	
Metabolite 1 - est.	Hydrolysis of the ester functionality	
Metabolite 2 - est.	Additional hydrolysis of the carbonyl group thus formed: cleavage acetic acid	
Metabolite 1 - exp.	Carbofuran	Prediction incorrect – site of carbamate hydrolysis wrongly predicted

O-P and S-P esters	**Transformation**	**Remarks**
(2,2,2-Trichloro-1-hydroxyethyl)phosphonic acid, Dimethyl ester	Dipterex	
Metabolite 1 - est.	Hydroxylation of the CCl_3 moiety: CCl_3 transformed into the carboxylic acid	
Metabolite 2 - est.	Hydrolysis of the carboxylic acid, followed by hydrolysis of the P-O-CH_3 moiety	
Metabolite 1 - exp.	Dichloroethanol and acetic acid - dichloroethanol is not predicted.	Prediction incorrect

O,O-Dimethyl O-[3-Methyl-4-(methyl thio)phenyl]ester phosphorothioic acid	Fenthion	
Metabolite 1 - est.	Transformation of SP-ester into the OP-ester	
Metabolite 2 - est.	Hydroxylation of the aromatic CH_3-moiety!	
Metabolite 1 - exp.	Fenthion sulfoxide and fenthion sulfone	Prediction incorrect
Phosphorothioic acid, O,O-Diethyl O-(4-nitrophenyl) ester	Parathion	
Metabolite 1 - est.	Transformation of SP-ester into the OP-ester	
	4-nitrofenol en dimethylfosfaat-ester: splitting of the moiety attached to the	
Metabolite 2 - est.	phosphate functionality	
Metabolite 1 - exp.	No info in database	
Phosphorodithioic acid, O,O-Dimethyl S-[2-methylamino)-2-oxoethyl]ester	Dimethoate	
Metabolite 1 - est.	Transformation of SP-ester into the OP-ester	
	Removal of the moeity attached to the	
Metabolite 2 - est.	phosphate-functionality	
Metabolite 1 - exp.	No info in database	
Phosphoric acid, 2,2-Dichloroethenyl di-methyl ester (Dichlorvos)	Phosphoric acid, 2,2-dichloroethenyl, dimethyl ester	
	Demethyldichlorvos (one of the O-CH_3	
Metabolite 1 - est.	methyl-groups is removed)	
Metabolite 2 - est.	Also the 2nd methylgroep is removed	
	Demethyldichlorvos (one of the O-CH_3	
Metabolite 1 - exp.	methyl-groups is removed)	Prediction correct
O,O,-Dimethyl S-[(4-oxo-1,2,3-benzotriazin-3(4H)-yl)methyl] ester, Phosphorodithioic acid	Methylazinphos	
Metabolite 1 - est.	Transformation of SP-ester into the OP-ester	
Metabolite 2 - est.	Ring opening of the N-C=O-binding:	

	substituent to the phosphorus atom	
Metabolite 1 - exp.	No info in database	
((Dimethoxyphosphinothioyl)thio)butanedioic acid, Diethyl ester	Malathion	
Metabolite 1 - est.	Transformation of SP-ester into the OP-ester	
Metabolite 2 - est.	Cleavage substituent	Prediction incorrect
Metabolite 1 - exp.	Malathion monocarboxylic acid and malathion dicarboxylic acid.	
O,O-Dimethyl O-(3-methyl-4-nitrophenyl ester phosphorothioic acid	Fenitrothion	
Metabolite 1 - est.	Transformation of SP-ester into the OP-ester	
Metabolite 2 - est.	Cleavage substituent	Prediction correct
Metabolite 1 - exp.	3-methyl-4-nitrophenol, cleavage substituent	
Phosphorothioic acid, O,O-Dimethyl-O-(p-nitrophenyl)ester	Parathion, methyl	
Metabolite 1 - est.	Transformation of SP-ester into the OP-ester	
Metabolite 2 - est.	Cleavage substituent	
Metabolite 1 - exp.	No info in database	
Phosphorothioic acid, O,O-Diethyl O-(2-isopropyl-6-methyl-4-pyrimidinyl) ester	Diazinon	
Metabolite 1 - est.	Transformation of SP-ester into OP-ester	Prediction incorrect
Metabolite 2 - est.	Hydrolysis CH3-functionality of the aromatic ring of the substituent to yield the carboxylic acid	
Metabolite 1 - exp.	2-isopropyl-4-methyl-6-hydroxypyrimidine (cleavage substituent)	
Phosphoric acid, 2-chloro-1-(2,4-dichlorophe-nyl)ethenyl diethyl ester	Chlorfenvinphos	
Metabolite 1 - est.	Cleavage 1 Me group of P: P-O-C becomes P-O-H	
Metabolite 2 - est.	Cleavage 2nd Me group of P: P-O-C	

	becomes P-O-H	
Metabolite 1 - exp.	No info in database	
S-[(1,3-Dihydro-1,3-dioxo-2H-isoindol-2-yl)methyl]O,O-dimethyl ester, phosphorodithioic acid	Phosmet	
Metabolite 1 - est.	Transformation of SP-ester into the OP-ester	
Metabolite 2 - est.	Ring opening of the substituent	
Metabolite 1 - exp.	No info in database	
Ethylphosphonodithioic acid, O-Ethyl S-phenyl ester	Fonophos	
Metabolite 1 - est.	Transformation of SP-ester into the OP-ester	
Metabolite 2 - est.	Cleavage substituent	
Metabolite 1 - exp.	No info in database	
Phosphorothioic acid, O,O-Diethyl O-(2-(ethylthio) ethyl)ester, Mixt. with O,O-diethyl S-(2-(ethylthio) ethyl)ester	Demeton	
Metabolite 1 - est.	Transformation of SP-ester into the OP-ester	
Metabolite 2 - est.	Oxidation S atom of the substituent. This yields the sulfoxide and the sulfon	
Metabolite 1 - exp.	No info in database	
Phosphoramidothioic acid, O,S-Dimethyl ester	Methamidphos	
Metabolite 1 - est.	Hydrolysis of the S-Me group: O-S-Me becomes OH	
Metabolite 1 - exp.	No info in database	
Phosphoric acid, 2-Chloro-3-(diethyl amino)-1-methyl-3-oxo-1-propenyl dimethyl ester	Phosphamidon	
Metabolite 1 - est.	Cleavage of a part of the substituent, yielding the carboxylic acid	

Metabolite 1 - exp.	No info in database	
Phosphorodithioic acid, O-Ethyl S,S-dipropyl ester	Ethoprophos	
	Cleavage of one of the S-CH_2-CH_2-CH_3 groups	
Metabolite 1 - est.		
Metabolite 2 - est.	Cleavage of the 2nd S-CH_2-CH_2-CH_3 group	
Metabolite 1 - exp.	No info in database	
Phosphorothioic acid, O-(4-Bromo-2-chloro-phenyl)-O-ethyl-S-propyl ester	Profenofos	
Metabolite 1 - est.	Cleavage of the S-CH_2-CH_2-CH_3-group	
Metabolite 2 - est.	Cleavage of the aromatic ring	
Metabolite 1 - exp.	No info in database	

Overall, CATABOL predictions matched the experimental pathways for five of the sixteen compounds for which experimental data were available (31 %). In general it may be concluded that for specific classes of compounds like carbamates and organophosphorus esters, the predictive capabilities of CATABOL are limited. This is independent of the chemicals being in the structural domain. On the other hand it should be noted that more accurate predictions might be obtained when studying less 'exotic' substance classes (think of aliphatic alcohols, ketones, and similar chemicals with limited and non-diverse functionalities).

5. Aquatic toxicity of metabolites formed

5.1 Carbamates

The QSAR that is most widely applicable for the prediction of aquatic toxicity of carbamates is equation 1 (*Cyprinus carpio*): amongst other considerations as this equation is based on the highest number of chemicals. The QSAR highlights the electronic interactions responsible for the induction of toxicity by carbamates. As a matter of course, equation 1 is applicable only to metabolites that themselves contain the carbamate moiety. In the dataset used in this study, this was the case for seven out of the sixteen carbamates studied. In Table 8 a comparison is made between the EC50-values of the parent compound and the CATABOL-predicted metabolites for the carbamates for which the metabolites themselves are carbamates too. In all cases, equation 1 was used to predict the EC50-values for *Cyprinus carpio*.

Table 8: Predicted EC50-values (mg L^{-1}) of carbamates and their metabolites for *Cyprinus carpi* (equation 1). The predicted EC50-value of the parent compound is compared with the predicted EC50-value of the metabolites. The predicted metabolite of the compounds in bold is more toxic than the parent compound.

Parent compound	**EC50 (mg L^{-1})**	**EC50 (mg L^{-1})**
	Parent	**CATABOL**
	compound	**predicted**
		metabolite
Propoxur	0.33	8.50
Aldicarb	0.48	0.65
Nabam	**0.03**	**0.02**
Phenol, 2-(1-methylethyl)-methylcarbamate	**0.34**	**0.09**
2,2-Dimethyl-1,3-benzodioxal-4-ol methylcarbamate	**0.31**	**0.03**
Thiodicarb	0.30	0.52
Benfuracarb	**1.65**	**0.05**

As can be deduced from Table 8 the carbamate-metabolites for four out of the seven compounds are more toxic than the parent compound, with increases in EC50 values of up till a factor of 33.

5.2 Organophosphates

In view of inherent differences in toxicity, a distinction needs to be made between O-P esters and S-P esters. Equation 2 is suited to predict EC50-values for the cholinesterase inhibition of O-P esters for *Cyprinus carpio*, whereas equation 3 is applicable for S-P esters. In Table 9 the predicted toxicity of the parent compounds is compared to the predicted toxicity of the metabolites.

Table 9: Predicted EC50-values (mg L^{-1}) of O-P- and S-P-esters for cholinesterase inhibition of *Cyprinus carpio.* The EC50-value of the parent compounds is compared with the EC50 of their metabolites. The predicted metabolite of the compounds in bold is more toxic than the parent compound.

Parent compound	**EC50 (mg L^{-1}) Parent compound**	**EC50 (mg L^{-1}) CATABOL predicted metabolite**
O-P-esters		
Dipterex	10.16	19.44
Dichlorvos	14.38	135.05
Chlorfenvinphos	0.42	8.76
Methamidphos	>10000	>10000
Phosphamidon	0.15	4.23
Ethoprophos	>200	1930.83
Profenofos	>25	25.50
S-P-esters		
Fenthion	**1.05**	**0.91**
Parathion	1.24	1.56
Dimethoate	9.98	14.32
Methylazinphos	**0.33**	**0.29**
Malathion	0.07	0.09
Fenitrothion	**10.65**	**0.59**
Parathion, methyl	**13.17**	**1.16**
Diazinon	**0.17**	**0.05**
Phosmet	**0.41**	**0.07**
Fonophos	0.11	2.03
Demeton	**4.36**	**0.40**

Table 9 shows that in all cases the metabolites predicted to be formed upon biodegradation of O-P-esters are less toxic than their parent compound. In case of S-P-esters, however, seven out of the

eleven esters form more toxic intermediates. In general, these are the compounds in which the P=S-moiety is transformed towards the P=O-moiety. This transformation usually constitutes the first step in the mechanism of toxicity of the S=P-esters studied.

5.3 Conclusion

Tables 8 and 9 show that overall, eleven of the compounds studied are expected to yield more toxic metabolites than their parent compounds: this is the case for four out of the seven carbamates and for seven out of the eleven S-P-esters studied. None of the O-P-esters are expected to yield more toxic metabolites, based upon the application of CATABOL in combination with a QSAR for prediction of toxicity of the parent compounds and the metabolites.

6. Main findings

As stated in the introductory chapter, the main objective of the study reported here was to design of the building blocks of a future ITS-module for dealing with data-poor chemicals. This design was performed for two chemical classes that despite their widespread use are to be considered as being data-poor. Within this objective, the use of quantumchemical descriptors was explored given the assumption that the chemical structure is the minimum piece of information always available. Apart from prediction of aquatic toxicity, prediction of biodegradation by means of CATABOL was explored and the findings on toxicity prediction and biodegradation modelling were combined to explore the possibilities of formation of more toxic metabolites.

It is shown that quantumchemical descriptors for the example of carbamates and organophosphates are suited as a starting point for deriving organism-specific QSARs that are indicative of the supposed mechanism of toxicity. The predictive capabilities of the models are depended on the size of the available test set of toxicity data.

CATABOL is shown to be a poor predictor of biodegradation pathways for the classes of chemicals investigated here. This is due to the fact that degradation of these chemicals is to a large extent via specific pathways that are not (yet?) included in CATABOL. This feature in itself leads one to conclude that predictions generated with CATABOL need to be considered with care. Despite favourable statistics, metabolites predicted by CATABOL as well as the likelihood may be highly inaccurate and on a case by case basis it needs to be considered whether the chemicals of interest are truly within the chemical domain of CATABOL.

Finally, it is shown that metabolites in case of carbamates and OP esters often are more toxic than their parent compounds. This calls for explicit consideration of metabolite formation for data-poor as well as for relatively data-rich compounds. A major conclusion to be drawn is that typically for data-poor chemicals there is a need of well-validated models capable of reliably predicting fate and effects. At the same time, however, the inherent lack of data precludes the generation of such models. This observation in itself shows that the postulate of chemical risk assessment being possible without any (animal) testing is invalid: reliable test data remain the basis for a thorough risk assessment.

In a subsequent step, the building blocks reported here need to be combined with similar methods for predicting the remaining endpoints that are essential for risk assessment of carbamates and OP

esters, as well as for other classes of data-poor chemicals. Combined with available experimental data, thus a module can be derived for the risk assessment of data-poor toxicants that does justice to the objectives of REACH of efficient safety management of chemicals whilst minimizing the use of test animals.

References

Boxall, A.B.A.; Sinclair, C.J.; Fenner, K.; Kolpin, D.; Maund, S.J. 2004. Environ. Sci.Technol. 38, 369a-375a.

Bradbury, S., Feijtel, T., van Leeuwen, K. 2004. Meeting the scientific needs of ecological risk assessment in a regulatory context. Environ. Sci.Technol. 38, 463a-470a.

Chen, J., Quan, X., Yang, F., Peijnenburg, W. 2001. Quantitative structure-property relationships on photodegradation of PCDD/Fs in cuticular waxes of laurel cherry (*Prunus Laurocerasus*). Sci. Total Environ. 269, 163-170.

Combes, R., Barratt, M., Balls, M. 2003. An overall strategy for the testing of chemicals for human hazard and risk assessment under the EU REACH system. ATLA 31: 7-19.

Cronin, M.T.D., Walker, J.D., Jaworska, J.S., Comber, M.H.I., Watts, C.D., Worth A.P. 2003. Use of QSARs in international decision-making frameworks to predict ecologic effects and environmental fate of chemical substances. Environ. Health Perspect. 111: 1376–1390.

Dimitrov, S., Kamenska, V., Walker, J.D., Windle, W., Purdy, R., Lewis, M., Mekenyan, O. 2004. Predicting the biodegradation products of perfluorinated chemicals using Catabol. SAR QSAR Environ. Res. 15: 69-82.

EC. 2003. Technical Guidance Document in support of the Commission Directive 93/67/EEC on risk assessment for new substances, Commission Regulation (EEC) No.1488/94 on risk assessment for existing substances, and Directive 98/8/EC of the European Parliament and of the Council concerning the placing of biocidal products on the market, Part 3. Ispra, Italy, European Chemicals Bureau, Joint research Centre, EUR 20416 EN/1.

EC. 2004. Commission Directive 2004/73/EC of 29 April 2004 adapting to technical progress for the twenty-ninth time Council Directive 67/548/EEC on the approximation of the laws, regulations and administrative provisions relating to the classification, packaging and labelling of dangerous substances. O.J. No L152 2004.

Hermens, J.L.M., Verhaar, H.J.M. 1996. QSARs in predictive environmental toxicology. Environ. Sci. Poll. Res. 3: 96-98.

Hulzebos, E.M., Posthumus, R. (2003). (Q)SARs: gatekeepers against risk on chemicals? SAR QSAR Environ. Res. 14: 285-316.

Hulzebos, E., Sijm, D., Traas, T., Posthumus, R. and Maslankiewicz, L. 2005. Validity and validation of expert (Q)SAR systems. SAR QSAR Environ. Res. 16: 385-401.

IHCP (Institute for Health and Consumer Protection). 2005. REACH and the Need for Intelligent Testing Strategies. European Commission, Directorate General Joint Research Centre, Institute for Health and Consumer Protection, Ispra, Italy.

JRC. 2005. REACH and the need for Intelligent Testing Strategies. Joint Research Centre, Institute of Health and Consumer Protection, Ispra, Italy. EUR 2154 EN.

Klopman, G., Tu, M. 1997 Structure-biodegradability study and computer-automated prediction of aerobic biodegradation of chemicals. Environ. Toxicol. Chem. 16: 1829-1835.

Mackay, D. 2007. Illustrated Handbook of Physical-Chemical Properties and Environmental Fate for Organic Chemicals, Volume II. CRC, Baton Rouge, USA.

Mackay, D., Boethling, R.S. (Editors). 2000. Handbook of Property Estimation Methods for Chemicals: Environmental Health Sciences. CRC, Baton Rouge, USA. OECD. 1981-2006. Guidelines for testing of chemicals. Paris, France. Organization for Economic Co-operation and Development.

OECD/IPCS. 2005. Report of the OECD/IPCS Workshop on Toxicogemomics, Kyoto, 13-15 Oct. 2004. Paris, Organisation of Economic Cooperation and Development, Environment Directorate, Joint Meeting of the Chemical Committee of the Working Party on Chemicals, Pesticides and Biotechnology, ENV/JM/MONO(2005)10, OECD Series on Testing and Assessment No. 50. http://www.oecd.org/document/23/0,2340%20and%202649_34365_33957015_1_1_1_1,00.html.

Posthumus, R., Traas, T.P., Peijnenburg, W.J.G.M., Hulzebos, E.M. 2005. External validation of Epiwin biodegradation models. SAR QSAR Environ. Res. 16: 135-148.

Rorije, E., Peijenburg, W.J.G.M., Klopman, G. 1998. Structural requirements for anaerobic biodegradation of organic chemicals: a fragment model analysis. Environ. Toxicol. Chem. 17: 1943-1950.

Rorije, E., Loonen H., Muller, M., Klopman, G., Peijenburg, W.J.G.M. 1999. Evaluation and application of models for the prediction of ready biodegradability in the MITI-test. Chemosphere 38: 1409-1417.

SRC – Syracuse Research Corporation. 1999. Environmental Fate Data Base: www.syrres.com.

Struijs, J. and Stoltenkamp, J. 1994. Testing surfactants for ultimate biodegradability.

Chemosphere 28: 1503-1523.

Van Leeuwen, K. and Bradbury, S. 2005 REGULATORY RISK ASSESSMENT: Trends and paradigm shifts are needed - European Commission & USEPA.

Vermeire, T., Hakkert, B., Aldenberg, T., Dang, Z., Hulzebos, E., Janer, G., de Knecht, J., van Loveren, H., Peijnenburg, W., Piersma, A., Traas, T., Verschoor, A., van Zijverden, M. 2006. Selected Integrated Testing Strategies (ITSs) for the risk assessment of chemicals. RIVM report no. 601050001. National Institute of Public Health and the Environment, Bilthoven, The Netherlands.

Appendix I

Overview of the chemicals used in this study, the test set of toxicity data, and the molecular properties of each chemical. Symbols used: MW = Molecular weight, HOF = Heat of Formation (KiloJoule), EE = Electronic Energy (eV), Ehomo = Energy level of the highest occupied molecular orbital (eV), Elumo = Energy level of the lowest unoccupied molecular orbital (eV), IP = Ionization Potential (eV), Q_C, Q_O, Q_N en Q_S are the electronic charges of the theses atoms in the carbamate-moiety of the chemical, MU = Dipole Moment (Debye). EC_{50} = the experimentally determined concentration at which an adverse effect is observable on 50 % of the organisms tested (mg L^{-1}). In case of OP-esters are Q_P, Q_O en Q_S the electronic charges of the atoms of the phosphate-moiety. A distinction is made between S-P and O-P esters: Q_O is equal to 0 in case of S-P esters, Q_S is equal to 0 in case of OP esters.

A - *Cyprinus carpio* (karper)

A1 – Carbamates

Chemical	MW	HOF	TE	EE	Ehomo	Elumo	IP	Q_C	Q_O	Q_N	Q_S	MU	EC_{50}
Carbofuran	221.255	-410.599	-2696.3	-17042.5	-8.862	0.252	8.862	0.383	-0.434	-0.095	-0.181	2.812	0.5
Benfuracarb	410.527	-864.048	-4812.7	-43356.0	-9.039	-0.733	9.039	0.416	-0.400	-0.236	-0.177	2.167	0.65
Aldicarb	190.26	-139.273	-2141.8	-11618.4	-9.168	-0.162	9.168	0.399	-0.385	-0.110	-0.251	4.880	1
1-Naphthalenol, methylcarbamate	201.224	-127.956	-2340.1	-14158.9	-9.015	-0.791	9.015	0.360	-0.390	-0.106	-0.170	4.702	1.19
Methiocarb	225.305	-278.197	-2470.7	-15001.3	-8.792	-0.304	8.792	0.380	-0.408	-0.093	-0.234	2.451	1.58
Phenol,2-(1-methylpropyl)-, methylcarbamate	207.272	-295.287	-2433.8	-16116.6	-9.634	-0.118	9.634	0.359	-0.390	-0.105	-0.166	4.543	1.6
Propoxur	209.244	-401.802	-2577.5	-15892.2	-9.204	0.167	9.204	0.382	-0.432	-0.095	-0.182	2.604	3.3
Phenol,2-(1-methylethyl)-,methylcarbamate	193.245	-279.179	-2284.3	-14060.5	-9.467	0.091	9.467	0.375	-0.391	-0.102	-0.213	3.791	3.5

A2 – Organophosphates

Chemical	MW	HOF	TE	EE	Ehomo	Elumo	IP	QP	QO	QS	MU	EC_{50}
Phosphamidon	299.691	-426.687	-3542.0	-23567.4	-12.626	-4.19	13.309	2.397	-0.823	0.000	4.162	51.5
Phosphoric acid, 2,2-dichloro-ethenyl, dimethyl ester	220.977	-211.489	-2505.9	-11162.0	-14.48	-4.408	13.199	2.373	-0.848	0.000	1.801	0.0015
Dipterex	257.438	-114.733	-2836.7	-13934.6	-14.46	-4.881	13.486	2.292	-0.834	0.000	1.599	6.2
Chlorfenvinphos	359.573	-145.188	-3878.1	-25026.4	-12.283	-4.19	13.212	2.398	-0.793	0.000	7.574	0.045
Profenofos	373.628	153.952	-3386.9	-21058.5	-13.322	-5.425	12.882	2.164	-0.757	0.000	4.504	0.09
Ethoprophos	242.331	278.119	-2315.3	-13453.0	-13.84	-6.578	15.527	2.075	-0.755	0.000	2.135	0.64
Malathion	330.35	-266.330	-3725.9	-24807.1	-13.491	-7.055	14.518	1.780	0.000	-0.352	8.170	0.002
Diazinon	304.343	-42.824	-3254.5	-23251.6	-13.367	-5.304	12.426	1.753	0.000	-0.252	3.107	1.4
Fenitrothion	277.231	268.504	-3178.7	-19077.9	-13.617	-7.102	17.063	2.116	0.000	-0.099	11.833	0.006
Fonophos	246.322	621.596	-2195.8	-12620.2	-13.27	-6.649	13.247	1.097	0.000	-0.231	2.113	0.088
Methylazinphos	317.317	617.162	-3221.2	-20612.7	-12.562	-6.956	13.247	1.884	0.000	-0.426	6.667	0.695
Parathion	291.258	217.201	-3328.5	-20578.5	-13.665	-6.859	14.931	2.085	0.000	-0.396	13.433	0.85
Fenthion	278.32	174.169	-2784.3	-17172.7	-12.616	-4.978	12.592	2.043	0.000	-0.613	8.827	1.16
Phosmet	317.314	303.826	-3310.0	-20473.5	-13.054	-6.948	15.050	1.840	0.000	-0.353	5.715	2.5
Parathion, methyl	263.204	326.380	-3028.8	-16980.3	-13.88	-7.131	16.906	2.143	0.000	-0.265	13.779	3

Dimethoate	229.248	348.068	-2306.1	-11778.1	-12.926	-7.463	13.358	1.607	0.000	-0.257	6.922	4.65
Demeton***	258.33	10.978	-2610.0	-15699.6	-13.886	-5.978	12.566	2.097	0.000	-0.540	2.773	15.2
Methamidphos	141.124	70.534	-1410.8	-5535.7	-14.478	-5.826	16.192	2.196	-0.760	0.000	2.472	69.01
Demeton***	258.33	0.812	-2610.1	-15879.0	-13.816	-5.223	12.642	2.317	-0.783	0.000	2.037	15.2

*** Demeton consists of a mixture of an S-P ester and an O-P ester.

B - *Aedes aegypti* (muggelarve)

B1 – Carbamates

Chemical	MW	HOF	TE	EE	Ehomo	Elumo	IP	Q_C	Q_O	Q_N	Q_S	MU	EC_{50}
Carbofuran	221.255	-410.599	-2696.3	-17042.5	-8.862	0.252	8.862	0.383	-0.434	-0.095	-0.181	2.812	0.09
Aldicarb	190.26	-139.273	-2141.8	-11618.4	-9.168	-0.162	9.168	0.399	-0.385	-0.110	-0.251	4.880	0.28
1-Naphthalenol, methylcarbamate	201.224	-127.956	-2340.1	-14158.9	-9.015	-0.791	9.015	0.360	-0.390	-0.106	-0.170	4.702	0.38
Propoxur	209.244	-401.802	-2577.5	-15892.2	-9.204	0.167	9.204	0.382	-0.432	-0.095	-0.182	2.604	0.15

B2 – Organophosphates

Chemical	MW	HOF	TE	EE	Ehomo	Elumo	IP	QP	QO	QS	MU	EC_{50}
Malathion	330.35	-266.330	-3725.9	-24807.1	-13.491	-7.055	14.518	1.780	0.000	-0.352	8.170	0.108
Fenitrothion	277.231	268.504	-3178.7	-19077.9	-13.617	-7.102	17.063	2.116	0.000	-0.099	11.833	0.0031
Parathion	291.258	217.201	-3328.5	-20578.5	-13.665	-6.859	14.931	2.085	0.000	-0.396	13.433	0.00265
Parathion, methyl	263.204	326.380	-3028.8	-16980.3	-13.88	-7.131	16.906	2.143	0.000	-0.265	13.779	0.013
Dimethoate	229.248	348.068	-2306.1	-11778.1	-12.926	-7.463	13.358	1.607	0.000	-0.257	6.922	5.35

C - *Crassostrea virginica* (Oester)

C1 – Carbamates

No toxicity data present in the database.

C2 – Organophosphates

Chemical	MW	HOF	TE	EE	Ehomo	Elumo	IP	QP	QO	QS	MU	EC_{50}
Dipterex	257.438	-114.733	-2836.7	-13934.6	-14.46	-4.881	13.486	2.292	-0.834	0.000	1.599	1
Chlorfenvinphos	359.573	-145.188	-3878.1	-25026.4	-12.283	-4.19	13.212	2.398	-0.793	0.000	7.574	0.6
Ethoprophos	242.331	278.119	-2315.3	-13453.0	-13.84	-6.578	15.527	2.075	-0.755	0.000	2.135	16
Malathion	330.35	-266.330	-3725.9	-24807.1	-13.491	-7.055	14.518	1.780	0.000	-0.352	8.170	9.07
Diazinon	304.343	-42.824	-3254.5	-23251.6	-13.367	-5.304	12.426	1.753	0.000	-0.252	3.107	8.4
Fenitrothion	277.231	268.504	-3178.7	-19077.9	-13.617	-7.102	17.063	2.116	0.000	-0.099	11.833	0.52
Fonophos	246.322	621.596	-2195.8	-12620.2	-13.27	-6.649	13.247	1.097	0.000	-0.231	2.113	0.33
Methylazinphos	317.317	617.162	-3221.2	-20612.7	-12.562	-6.956	13.247	1.884	0.000	-0.426	6.667	0.62
Parathion	291.258	217.201	-3328.5	-20578.5	-13.665	-6.859	14.931	2.085	0.000	-0.396	13.433	0.85
Fenthion	278.32	174.169	-2784.3	-17172.7	-12.616	-4.978	12.592	2.043	0.000	-0.613	8.827	0.58

D - *Brachydanio rerio* (Vis)

D1 – Carbamates

No toxicity data present in the database

D2 – Organophosphates

Chemical	MW	HOF	TE	EE	Ehomo	Elumo	IP	QP	QO	QS	MU	EC_{50}
Phosphoric acid, 2,2-dichloro-ethenyl, dimethyl ester	220.977	-211.489	-2505.9	-11162.0	-14.48	-4.408	13.199	2.373	-0.848	0.000	1.801	100
Malathion	330.35	-266.330	-3725.9	-24807.1	-13.491	-7.055	14.518	1.780	0.000	-0.352	8.170	0.035
Diazinon	304.343	-42.824	-3254.5	-23251.6	-13.367	-5.304	12.426	1.753	0.000	-0.252	3.107	0.0012
Parathion	291.258	217.201	-3328.5	-20578.5	-13.665	-6.859	14.931	2.085	0.000	-0.396	13.433	5.6
Dimethoate	229.248	348.068	-2306.1	-11778.1	-12.926	-7.463	13.358	1.607	0.000	-0.257	6.922	6.2
Demeton***	258.33	10.978	-2610.0	-15699.6	-13.886	-5.978	12.566	2.097	0.000	-0.540	2.773	10

*** Demeton consists of a mixture of an S-P ester and an O-P ester. The data reported here reflect the S-P ester.

RIVM

National Institute
for Public Health
and the Environment

P.O. Box 1
3720 BA Bilthoven
The Netherlands
www.rivm.com