SAR and QSAR in Environmental Research,
Vol. 18, Nos. 1–2, January–March 2007, 3–20

Taylor & Francis
Taylor & Francis Group

Evolution of the international workshops on quantitative structure-activity relationships (QSARs) in environmental toxicology†

K. L. E. KAISER*

TerraBase Inc., 1063 King St. West, Hamilton, Ontario, Canada

(*Received 12 May 2006; in final form 21 August 2006*)

This presentation will review the evolution of the workshops from a scientific and personal perspective. From their modest beginning in 1983, the workshops have developed into larger international meetings, regularly held every two years. Their initial focus on the aquatic sphere soon expanded to include properties and effects on atmospheric and terrestrial species, including man. Concurrent with this broadening of their scientific scope, the workshops have become an important forum for the early dissemination of all aspects of qualitative and quantitative structure-activity research in ecotoxicology and human health effects. Over the last few decades, the field of quantitative structure/activity relationships (QSARs) has quickly emerged as a major scientific method in understanding the properties and effects of chemicals on the environment and human health. From substances that only affect cell membranes to those that bind strongly to a specific enzyme, QSARs provides insight into the biological effects and chemical and physical properties of substances. QSARs are useful for delineating the quantitative changes in biological effects resulting from minor but systematic variations of the structure of a compound with a specific mode of action. In addition, more holistic approaches are being devised that result in our ability to predict the effects of structurally unrelated compounds with (potentially) different modes of action. Research in QSAR environmental toxicology has led to many improvements in the manufacturing, use, and disposal of chemicals. Furthermore, it has led to national policies and international agreements, from use restrictions or outright bans of compounds, such as polychlorinated biphenyls (PCBs), mirex, and highly chlorinated pesticides (e.g. DDT, dieldrin) for the protection of avian predators, to alternatives for ozone-depleting compounds, to better waste treatment systems, to more powerful and specific acting drugs. Most of the recent advances in drug development could not have been achieved without the use of QSARs in one form or another. The pace of such developments is rapid and QSARs are the keystone to that progress. These workshops have contributed to this progress and will continue to do so in the future.

Keywords: QSAR; Workshops; Environmental toxicology; Research

*Email: mail@terrabase-inc.com
†Presented at the 12th International Workshop on Quantitative Structure-Activity Relationships in Environmental Toxicology (QSAR2006), 8–12 May 2006, Lyon, France.

SAR and QSAR in Environmental Research
ISSN 1062-936X print/ISSN 1029-046X online
DOI: 10.1080/10629360601053927

1. Introduction

Environmental problems rose to widespread public concern in the 1960s. Rachel Carson's book *Silent Spring* was a major event in pushing these problems to the forefront of media coverage [1]. Then, in 1969, Soren Jensen discovered the widespread contamination of wildlife with polychlorinated biphenyls [2]. Governments responded by creating agencies to deal with these findings, such as the *Environmental Protection Agency* in the USA (in 1970), the *Department of the Environment* in Canada (in 1971), and similar agencies in the other countries belonging to the *Organization for International Cooperation and Development* (OECD). Soon, their scientists discovered many other contamination problems in their environs. One of such compounds, very specific to Lake Ontario, was mirex, a perchlorohydrocarbon of highly symmetric structure. Some of these discoveries could be linked directly to biological problems, such as egg shell thinning in avian predators, widely believed to be caused by the insecticide DDT and its major metabolite DDE, resulting in poor hatching success and declines of their populations [3].

A decade later, many reports had been published describing environmental contamination problems at many localities, and in wildlife species. Particularly in higher trophic niches of oligotrophic aquatic systems, such as in the Laurentian Great Lakes and in the Arctic, top predators were found to have high levels of certain contaminants. In turn, scientists began to try to rationalize the findings and to develop ways to predict the effects and environmental pathways of such compounds. The *Great Lakes Water Quality Agreement* of 1978 between Canada and the United States called for an expansion of quantitative structure-activity relationships (QSARs) research in this field. The "*International Workshops on QSAR in Environmental Toxicology*" (at some times also entitled "... *Workshops in Environmental Sciences*", and in 2004 "*International Workshop on Quantitative Structure-Activity Relationships in Human Health and Environmental Sciences*") came into being as a result of this desire. The following gives a brief review of the evolution of these workshops. The reader is also referred to the communication by John Walker [4]. Table 1 gives an overview of the workshop chairs and proceedings' formats.

Table 1. Workshop chairs and proceedings editors, 1983 to 2006.

Workshop Year	*Publisher*	*Format*	*Workshop Chair*[+]*, Proceedings Editor(s)*
1983	Kluwer	Book	**K. Kaiser**
1986	Kluwer	Book	**K. Kaiser**
1988	US NTIS	Proceedings	**T. Schultz**, J. Turner, W. England, N. Kwaak
1990	Elsevier	Journal	**J. Hermens**, A. Opperhuizen
1992	Taylor & Francis	Journal	**G. Veith**, S. Broderius, G. Niemi
1994	Taylor & Francis	Journal	**W. Karcher**, J. Devillers
1996	SETAC	Book	**F. Chen**, G. Schüürmann
1998	SETAC	N/A*	**J. Walker**
2000	Taylor & Francis	Journal	**O. Mekenyan**, T. Schultz
2002	Wiley	Journal	**R. Breton**, R. Purdy, G. Schüürmann
2004	Taylor & Francis	Journal	**M. Cronin**, J. Dearden, J. Duffy, T. Schultz
2006	Taylor & Francis	Journal	**J. Devillers**, A. Carpy, B. T. Fan

[+]Workshop Chair in bold.
*Not available.

2. The workshops in brief

2.1 *1983, Hamilton, Ontario*

The first workshop was held in 1983 at McMaster University, Hamilton, Ontario. At that time, my research director at the National Water Research Institute in Burlington, Dr Roderick J. Allan, encouraged me to organize a workshop in this field which I had worked on for several years. Perhaps, Rod (being a geologist) felt that I had spent enough research dollars on this esoteric QSAR activity without having any *physical* product to show for the investment (as opposed to some mathematical equations) and he may have wanted some assurance that QSAR was not a figment of my imagination but real science. Luckily for me, I could persuade some of my esteemed colleagues to join me in this and we had a very successful meeting. As could be expected, the first workshop was limited in scope. My intention was to bring together people from various disciplines and to explore commonalities and QSAR approaches within this frame. Naturally, the first workshop was dominated by methods to rationalize and predict the aquatic toxicity of compounds. Figure 1, hitherto unpublished, shows some of the participants.

Ih Chu, George Dixon, Gerry LeBlanc, Bob Lipnick (represented by Larry Newsome), Don Mackay, Terry Schultz, and ourselves presented QSARs for the toxicity of chemicals to various aquatic organisms. Kurt Enslein modeled biological oxygen demand, Dieter Freitag described the ''Environmental Hazard Profile'', Efraim Halfon modeled environmental fate in large lakes, Don Hart described mutagenic effects in amphibians, Barry Oliver explored bioconcentration relationships, Rainer Koch contributed a paper on the effects of several groups of chemicals, Bill Dunn III

Figure 1. Some of the participants at the workshop in 1983, photographer unknown. From left to right: Alice Bobra, Barry Oliver, Peter Wells, Don Hart, Wayne Landis, Dieter Freitag, Vlado Zitko [in back], Don Mackay, (?) [obscured], Kazimiera Kwasniewska, Larry Newsome [in back], Juan Ribo, Terry Schultz, Klaus Kaiser, Lynn McCarthy, Wan Ying Shiu, Peter Hodson, (?) [obscured], Bruce Gray, Ih Chu.

(represented by Dave Stalling) described the use of SIMCA modelling and Vlado Zitko discussed data evaluation and descriptors in QSAR.

The inaugural workshop was much appreciated by all participants and a sequel was desired for the future. One reason for its success was the unique setting and format with ample discussion. We were lucky to have had the use of a sufficiently large lecture room with a novel, tiered and semicircular layout. After a few months of editing and getting it all together, the proceedings were published in 1984 [5].

2.2 *1986, Hamilton, Ontario*

In 1986, I organized a second workshop, again at McMaster University. The experience gained with QSAR-I was quite helpful in organizing it. Also, we had more time in preparing for it. QSAR-II was attended by colleagues from a dozen or so countries and was a really vibrant affair with lots of discussions, at times quite heated, and the proceedings were published in 1987 [6]. Initially, I hoped to publish the discussions along with the papers, but it proved too difficult. At the very least, it would have delayed the publication for several months, possibly up to a year and I did not want to cause such a delay.

QSAR-II attracted contributions from a variety of research fields other than aquatic toxicology. There were presentations on larger sets of chemicals to several aquatic organisms, such as fish by J. Hermens, R. Laughlin, R. Purdy, and A. Sablijc, and bacteria by myself, and J. Ribo. L. Newsome, D. Roberts, T. Schultz, W. Shiu, G. Veith and coworkers explored groups of chemicals, such as phenols or anilines. W. Butte, L. McCarthy, and H. Tadokoro presented results on bioconcentration and bioaccumulation of substances in organisms. F. Darvas, and J. Dearden explored biodegradation routes and kinetics. I. Chu presented findings on tissue distribution; M. Crowley, and D. Passino-Reader discussed administrative utility. R. Brüggemann, and W. Williams looked at the administrative utility of QSARs. It is quite noteworthy, that already then, there were several contributions dealing with inter-species relationships of groups of chemicals, such as by Abernethy, Banerjee, and Enslein. This field is experiencing a new impetus and revival at this time as evident from the session devoted to it here. A visit of the renowned *Royal Botanical Gardens* in Burlington and banquet at the *McMaster University Faculty Club* rounded off the event.

The success of the initial workshops was also based on my stipulation that there would be ample time for discussion of each presentation. In fact, I insisted that the discussion periods would be a minimum of one third of the duration of the presentation. More often than not, after some initial period hesitation, this really got things going, and vigorous discussion evolved, often exceeding the allotted time. This was truly a workshop where the discussions went far beyond the clarification of simple aspects of the article and most participants found this extremely useful.

2.3 *1988, Knoxville, Tennessee*

On my suggestion, Terry Schultz, *University of Tennessee*, readily agreed to host the third workshop in 1988 at the site of the former World's Fair in Knoxville, May 22–26, 1988. The theme for QSAR-88 was "*Interrelationships of QSAR and Mechanisms of Toxic Action*". Some 60 participants contributed a variety of presentations. The need for uniform data notation as log(1/EC) with the concentration expressed in molar form (or milliM, or microM), adequate statistical definition, and more detailed QSAR

descriptors were recognized as important factors. Inter-species toxicity correlation were further explored, and QSAR descriptor inter-correlations were explored. Linear solvation energy (LSER) parameters and electronic state parameters, e.g. the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO) found increased use and acceptance. In terms of mode of action (MOA) at least three principal MOAs had been well established. Terry Schultz edited the proceedings together with his colleagues James Turner, Wendy (Williams) England and Norma Kwaak, published as an US Department of Energy report [7]. A hiking excursion into the *Blue Mountains* was much enjoyed by all participants.

2.4 *1990, Eindhoven, The Netherlands*

The fourth workshop was hosted by Joop Hermens and Anton Opperhuizen of the *University of Utrecht*, 16–20 September, 1990, at the Koningshof Conference Center in Eindhoven, The Netherlands with the proceedings published in the journal *Science of the Total Environment* (Elsevier), also available as the book *QSAR in Environmental Toxicology – IV*, edited by Hermens and Opperhuizen [8]. Held for the first time in Europe, the workshop attracted a large gathering of approximately 150 colleagues, primarily from European countries. Five major themes were discussed, i.e. QSAR techniques and descriptors, biodegradation and environmental fate, accumulation and clearance kinetics, modes of action and QSARs for selected groups of compounds and aquatic species, and applications in practice.

2.5 *1992, Duluth, Minnesota*

The fifth workshop was organized by Gil Veith and Gerry Niemi and coworkers of the *US Environmental Effects Research Laboratory at Duluth*, held at the Holiday Inn, Duluth, Minnesota, July 19–23, 1992. An estimated one hundred colleagues attended from several continents. There were no proceedings as such, but some of the papers were published in volume 2 of the journal *SAR and QSAR in Environmental Research*. The social highlight was an evening cruise in the *Duluth Harbor*, complete with a band and lots of suds.

2.6 *1994, Belgirate, Italy*

Walter Karcher and his colleagues of the European Commission's *Joint Research Centre* (JRC) at Ispra organized the sixth workshop close to the shores of Lake Maggiore at the *Hotel Villa Carlotta*, Belgirate, Italy, September 13–17, 1994. Peer reviewed contributions were published in volume 3 of the journal *SAR and QSAR in Environmental Research*, guest edited by W. Karcher. Visits of the JRC and a banquet at the shore of Lake Maggiore rounded off the workshop.

2.7 *1996, Elsinore, Denmark*

The seventh workshop was hosted by Fei Chen of the National Environmental Research Institute of Denmark at the pastoral setting of the *Scanticon Borupgaard Conference Center* in Elsinore, Denmark, June 24–28, 1996. It was attended by representatives from 18 countries and more than 100 papers were presented

covering a broad aspect of QSAR research. The proceedings, edited by Fei Chen and Gerrit Schüürmann, were published the following year as a volume of the SETAC Special Publication Series [9]. In the foreword, Hugo Kubinyi, then Chair of the *QSAR and Modelling Society*, states that ''The scientific community dealing with QSAR and modelling in drug design had hitherto no contact or overlap with scientists who are involved in environmental sciences''. QSAR-96 was the first to bridge this gap by starting to bring the two solitudes together. QSAR-96 was dominated by modelling techniques, physico-chemical characteristics, and statistical and risk assessments. Visits of the *Viking Ship Museum* in Roskilde and the nearby wind turbine research station were highlights of the social events.

2.8 *1998, Baltimore, Maryland*

John Walker of the *US Environmental Protection Agency* organized the eighth workshop, QSAR-98, at the historic *Lord Baltimore Hotel*, downtown Baltimore, Maryland, 16–20 May, 1998. The workshop was well attended by approximately 250 delegates. The program was well rounded but, unfortunately, publication of the proceedings somehow went astray. After two rounds of peer review, and other delays, the intended publisher (SETAC Press) came recently up with other, rather strange reasons for delay, and at least some of the original submissions are still ''in press'' with a different publisher. While regrettable, this event should not distract from the very successful workshop itself and the novel results presented there. The social highlight was a visit of the *National Aquarium in Baltimore* with banquet.

2.9 *2000, Bourgas, Bulgaria*

The ninth workshop was organized by Ovanes Mekenyan and coworkers of the *Prof. Zlatov University*, Bourgas and convened at the modern Black Sea resort *Dyuni Conference Center*, Bourgas, Bulgaria, September 16–20, 2000. It was well attended with over 100 participants and there were lively discussions both during and after the sessions. The peer reviewed contributions were published in volume 13 of the journal *SAR and QSAR in Environmental Research*, guest edited by O. Mekenyan and T.W. Schultz. An after-workshop excursion by bus to Istanbul with its imposing downtown architecture was much enjoyed by the participants.

2.10 *2002, Ottawa, Ontario*

Roger Breton of *Environment Canada* organized and chaired the tenth workshop at the downtown setting of the *Chateau Laurier Hotel*, Ottawa, Ontario, May 25–29, 2002. Some 30 contributed papers were peer reviewed and published in volume 22 of the journal *QSAR & Combinatorial Science*, guest edited by R. Breton, R. Purdy and G. Schüürmann. The social highlight was the banquet at the unique *Museum of Civilization*, on the shores of the Ottawa River.

2.11 *2004, Liverpool, England*

Mark Cronin and John Dearden of the *Liverpool John Moores University* organized the ''*11th International Workshop on Quantitative Structure-Activity Relationships in Human Health and Environmental Sciences*'' at the venerable *Brittania Adelphi Hotel*, downtown Liverpool, England. Some 170 colleagues attended from 28 countries. The workshop

was complemented by an exhibition of a dozen or so software providers in the field as well as by a training course on "Building better QSARs". Selected peer-reviewed papers were published in volumes 15 and 16 of the journal SAR and *QSAR in Environmental Research*, guest edited by M.T.D. Cronin, J.C. Dearden, J.C. Duffy and T.W. Schultz. The social highlight was the banquet at a prestigious private club in Liverpool.

3. Discussion

3.1 *Toxicity data*

At the beginning of the 1980s, QSAR research in environmental toxicology was very much hampered by a lack of available compatible data for any aquatic species. Recognizing this problem, several groups set out to systematically measure the acute effects of many individual chemicals to several important aquatic species. In particular, the 96 h LC50 to the fish fathead minnow (*Pimephales promelas*), the 40 h IGC50 for the ciliate *Tetrahymena pyriformis*, and 5–30 min EC50 for the bacterium *Vibrio fischeri* (formerly *Photobacterium phosphoreum*). At present, the published data for these assays comprise approximately 900, 1600, and 2300 compounds, respectively. In addition, smaller data sets are available for other widely used species, including the fishes rainbow trout (*Oncorhynchus mykiss*), red killifish (*Oryzias latipes*), guppy (*Poecilia reticulata*) and bluegill (*Lepomis macrochirus*), as well as the waterflea (*Daphnia* sp.), and the algae *Chlorella* sp. and *Selenastrum* sp., and others. These data now form much of the experimental backbone for the QSAR work in environmental sciences. Regrettably though, further testing and expansion of these experimental efforts has all but ceased several years ago. Therefore, there is a lack of such data for newer substances of environmental concern, such as certain fluorinated compounds, modern pesticides, widely used prescription and non-prescription drugs, major components of fragrances and cosmetic products, and many of the surfactants and detergent constituents which are in large-scale use, and the predominant degradation products and metabolites of such materials. Obviously, this void precludes efforts to assess the performance of existing QSAR models towards many of these newer chemicals.

3.2 *Octanol/water partition coefficients*

Prior to the early 1980s, measured octanol/water partition coefficients (commonly known as $\log P$ or $\log K_{ow}$) of most compounds of relevance in the environment were very sparse. Moreover, the available data were of questionable quality. The reason for that was that they were difficult to measure with the common testing procedures and the available analytical methods, due to their generally low solubility. Our ability to measure such values with more confidence came with development of the slow-stirring method by Hermens and coworkers in 1989 [10]. This can be shown on the examples of available $\log P$ values for some chlorinated compounds as found in the compilation by Hansch *et al*. in 1979 [11] and its later edition in 1995 [12], table 2. However, there are still substantial voids in good measurements for environmental contaminants at the present time. In particular, I am referring to surfactants, such as non-ionic polyethoxylates, which I will show in more detail later in this overview. Unfortunately, such experimental work is widely considered as less glamorous than

Table 2. Available measured octanol/water partition coefficients for some environmental contaminants in 1979 [11] and in 1995 [12].

Compound	*Formula*	*log P, 1979*	*log P, 1995*
Aldrin	$C_{12}H_8Cl_6$	3.01	6.50
p,p′-DDT	$C_{14}H_9Cl_5$	3.98–6.19	6.91
2,4,5,2′,4′,5′-C_{16}-PCB	$C_{12}H_4Cl_6$	6.72	7.16

developing QSAR models using substitute log *P* methods, such HPLC, or calculation with computer programs even though some may give results which are demonstrably far from reality. For a good introduction into the field of physico-chemical property estimations, the handbook by Boethling and Mackay is recommended [13].

While log *P* is a highly useful parameter in optimizing effects of novel lead compounds in drug development (*vide infra*), from a macro perspective, it does not help in finding problematic compounds which may act by an enzyme-specific mode. This can be demonstrated with a plot of log *P* *versus* the rat/mouse oral log(LD_{50}) values for approximately 6000 compounds, as shown in figure 2. The complete absence of any relationship whatsoever is quite apparent.

3.3 *QSARs for acute toxicity*

Linear QSARs for selected groups of chemicals, primarily compounds acting on aquatic species by non-polar and polar narcotic MOAs, were developed early on, primarily as functions of the octanol/water partition coefficient. Indeed, Hans Könemann of the Netherlands received the first *International QSAR Award* (at Eindhoven, 1990) in recognition for his contributions to the development of such models. Many related models along the same lines have been described since then and are still being developed today (e.g. Pavan *et al*. [14]). While their statistics and predictive power are high, they are mostly applicable only to narrowly MOA-defined sets of compounds, such as for non-polar narcosis. Their application to more complex chemicals, i.e. those with more than one functional group and particularly those with strong electrophilic substituents, such as $–NO_2$, –CN, etc., is not possible at this time. Moreover, this limitation is not likely to be overcome by further research along this line. More efforts should be directed towards the development of more holistic approaches, i.e. non-MOA-based models, such as the approach described by Devillers [15], Raevsky and Dearden [16], and our *TerraQSAR – FHM* acute fish toxicity estimation program [17].

3.4 *QSARs for sub-acute toxicity*

The field of QSAR research in sub-acute toxicity could benefit from some work. Unfortunately, there are not many data to choose from. The AQUIRE database lists sub-acute data for some 30 compounds only for the fathead minnow. Lowest Effect Concentrations (LOECs) and No Effect Concentrations (NOECs) are widely used in risk assessment procedures. While there are some measurements reported for such, more often than not, they are calculated as the 1/10th or 1/100th concentration of acute effects, often with an additional safety factor of 0.1. As a result, many LOEC and NOEC data are highly conservative estimates based on little experimental evidence. Obviously, more experimental data and model development would be desirable.

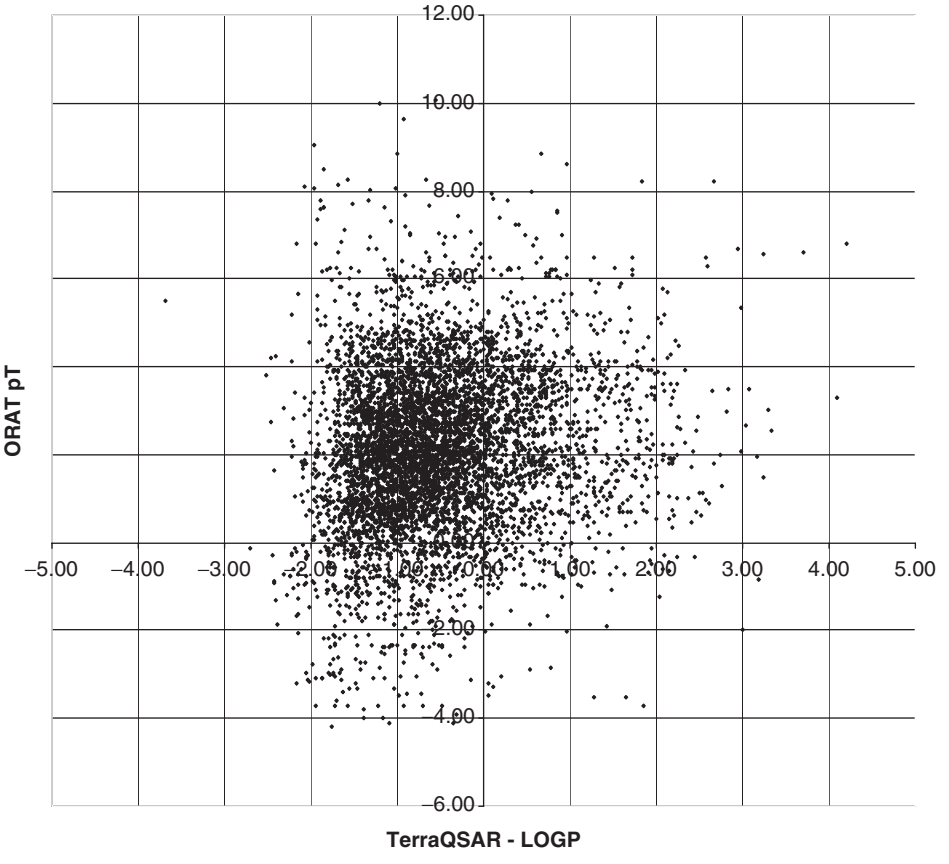

Figure 2. Plot of mouse/rat oral LD_{50} values ($pT = -\log LD_{50}$) *vs.* computed octanol/water partition coefficients for approximately 6000 compounds, demonstrating the visual absence of any relationship.

3.5 *Interspecies toxicity correlations*

This area of research is experiencing a revival. When comparing bioassay results of individual chemicals between closely related species, differing results can often be ascribed to experimental variances, differences in environmental conditions, e.g. water hardness, pH, or alkalinity, differences in species acclimation, feeding, light and dark cycles, differences in exposure times, differences in the chemicals' purities, and so forth. Therefore, biologically closely related species will generally have similar susceptibilities to a variety of toxicants. This effect can be made use of in estimating the effects of one chemical from the experimental results for one organism to that on another, as shown with the *Vibrio fischeri* bacteria and a variety of fish species several years ago [18], and more recently with *Tetrahymena pyriformis* [19]. While not part of the "classical" QSAR research, this effort is often quite successful and generally accepted to be part of the field of "*QSAR in environmental research*". For some recent results, see, for example, Lessigiarska *et al.* [20] and the program *ICE* by Asfaw *et al.* [21].

3.6 *QSARs for bioconcentration*

This field has not seen much work since the early days of environmental QSAR research. Basically, bioconcentration and the food-chain dependant biomagnification of hydrophobic substances follow their lipophilicity as expressed by log *P*. In a laboratory setting with pristine water, bioconcentration of a lipophilic contaminant onto fish can be well demonstrated. In contrast, in most natural systems, the surface area of the plankton, primarily algae, is several orders of magnitudes higher than that of the fish present and, therefore, the primary target of bioconcentration. Consequently food-chain biomagnification is the predominant mechanism by which compounds accumulate in higher trophic level organisms in the aquatic environment. The recent work by Dearden and Shinnawei [22] found an additional dependence on molecular shape which, of course, also affects chemicals' partitioning between the aqueous phase and lipids.

3.7 *QSARs for biodegradation*

Degradation of compounds finding their way into waste water streams and the receiving environment is an important property. Any significant resistance of materials to degradation in waste treatment systems will result in their release to the environment. Once there, the compounds can accumulate in the aqueous phase or, if they have high lipophilicity, adsorb onto particulates from which they enter the food chain and can become biomagnified at higher trophic levels. DTT, mirex and PCBs are well known examples of that. Therefore, the biotic and abiotic degradation of compounds is an importance consideration when intending to use substances in large quantities and delocalized discharge to the environment. Alternatively, volatile compounds, such as CFCs can volatilize and migrate to the troposphere when they interact with the ozone layer, thus causing higher UV radiation incidence at the earth surface. Several commercial or free programs exist, e.g. *BIOWINx*, recently evaluated by Posthumus *et al*. [23], *MetabolExpert* [24], *CATABOL* [25], as well as algorithms for developing categories of persistence/biodegradability, see, for example, Mekenyan and coworkers [26].

A few years ago, I reviewed the existing testing methods proposed by various agencies, including the OECD, the *American Society for Testing and Materials* (ASTM), the International Standards Organization (ISO), and the US Environmental Protection Agency (EPA) [27]. The most widely used test for 'Ready Biodegradability' is the OECD 301C test, which is a modified version of Japan's MITI test #1. It requires practically complete degradation within a period of 28 days. This period is comprised of a lag time and a biodegradation phase. The latter is further limited to a maximum of 10 days. Measurements performed on the same chemical but with different testing protocols do not always result in the same findings as the acclimation requirements of the bacteria to a novel substrate differs between regions. Obviously, that can create problems when trying to assess biodegradability. One example concerns a series of sulfonamides [28] which are degraded within the 28 day period of the 'Ready Biodegradability' test, albeit at a high sludge concentration, but are assessed as 'not ready biodegradable' (see in ref. [29]). Without trying to express an opinion about which assessment is correct or not, I feel that such differences in interpretation ought not to exist. Perhaps, this problem could be resolved by devising a different set of testing parameters which would give less equivocal results.

3.8 *Modelling techniques*

Classical linear and multilinear QSAR models were supplanted with newer modelling techniques, especially neural networks. Among the latter, a variety of methods emerged and were applied to all of the major aquatic toxicity data sets or subsets thereof, as described in more detail elsewhere [30]. The primary advantage and use of neural networks (NNs) lies in their ability to process non-linear relationships. This is of importance when modelling large data sets of compounds, especially when they include substances acting by different mechanisms or mode of action. There are some twenty or so different types of NNs with substantially different mathematical structures and methodologies. Some of these may be more suited to common problems than others. Overtraining and the resulting lack of predictability is a common problem with some NNs. However, with proper training and optimization settings, most NNs can provide highly reliable predictions. At the workshop in Bourgas, Johann Gasteiger vehemently opposed the view, expressed by some, that NNs are somewhat of a ''black box''. Unfortunately, there is still a considerable resistance to the widespread acceptance of NNs for predictive QSAR modelling, mainly because of their widely conceived ''lack of transparency''. Nevertheless, up to 2002, prior to the workshop in Liverpool, 2004, the number of papers using NNs was generally quite small. In 2004, at Liverpool, there was a noticeable increase in the application of NNs to a variety of modelling problems. Perhaps the time is ripe for a wider acceptance of such modern techniques. The NN applied in our *TerraQSAR* models is based on the principle of 'estimation of the conditional average'. That methodology has a number of significant advantages over the other types of NNs, the two most important being quick and unequivocal training optimization. For another recent development in this field, see Devillers [15].

3.9 *Interaction with human health research*

The year 2004 was a watershed in terms of broadening the scope of the workshop to include adsorption and metabolism (ADMET) properties as important aspects of human health as it relates to environmental effects. Specifically, the absorption and irritancy of chemicals used in large scale in fragrances and toiletries, which therefore, also find their way into the environment, became a focal point for discussion. The list of such materials in cosmetics is substantial and covers easily more than 2000 compounds. Moreover, some of our common household products contain a variety of substances that affect both humans and aquatic species. Good models with a good predictive power in this field are highly desirable, both from the development (of products) side and the environmental protection side. Also in recent years, a variety of over-the counter and prescription drugs have been found in the rivers and lakes. Their presence has caused considerable concern and their long-term effects on aquatic organisms is largely unknown. Much more work is needed in this field.

The convergence of the human health (i.e. drug development) field with environmental research is also very apparent in the area of endocrine disrupting chemicals, as outlined below. As welcome as this interaction may be, one should not loose sight of the fundamental differences between these two branches of the QSAR field. Basically, the drug development is interested in finding new lead compounds and in optimizing their efficacy for a certain effect. The environmental research goal is to recognize existing or potential problem chemicals and to prevent them from entering our ecosphere in quantities causing negative effects on non-target organisms.

3.10 *QSARs for steroid receptor binding*

In recent years, the area of endocrine disrupting chemicals (EDCs) has become of much interest in environmental research. Some field observations have shown changes in the normally balanced distribution of genders in both aquatic and terrestrial species. This has been ascribed to chemicals interacting disrupting the endocrine function organisms. Primary targets for modelling the effects of substances on the endocrine system are the relative binding affinities (RBAs) of compounds to the estrogen (ER) and androgen (AR) receptor molecules.

The *National Center for Toxicological Research*, USA, initiated a large research program several years ago, with the structures and affinity data available on the internet as the *Endocrine Disruptor Knowledge Base*. It lists ER binding data for approximately 300 individual chemicals [31]. In contrast, the commercially available *TerraTox – Steroids* database has binding affinities for nearly 4500 individual chemicals [32]. Several models for estrogen receptor binding affinity (E2-RBA), e.g. including the multidimensional *Quasar* and *Raptor* modelling technologies [33], a partial least squares (PLS) model [34], and the neural network *TerraQSAR – E2-RBA* program [35] were developed. The latter is based on absolute canonical SMILES strings and measured binding affinity data of over 2000 compounds. Figure 3 shows the predicted *versus* measured E2-RBAs for this data set [35].

In terms of modelling any specific enzyme interaction, such as the ER-RBA, it is important to use absolute 3D-structure coordinates. This can easily be demonstrated on the examples of 17*beta*- and 17*alpha*-estradiol and their 17-ethinyl analogues. The RBA of both of the 17*alpha*-OH stereoisomers is much lower (RBA $\sim$ 5%) than those of the corresponding 17*beta*-OH compounds (RBA 100(+)%) [32], as shown in figure 4. From an environmental point of view, EDCs of high binding affinity are generally of little concern, except for those which are widely used pharmaceuticals, such as 17*alpha*-ethynylestradiol. In contrast, a variety of high production volume chemicals which are used in many different products and which are discharged into every waste stream, such as phthalates could be of concern. Therefore, the recent measurements of the RBAs of a series of low-affinity phthalates and related compounds by Akahori *et al.* [36] are important, as summarized in table 3. However, in my opinion, we need reliable measurements at even lower log(E2-RBA) values, down to at least -5, in order to be able to assess the potential for environmental impact of many of these compounds with confidence.

In other areas of compound – enzyme interactions, the activity of optical antipodes often varies by factors of 1000 or more, as shown for the R- and S-isomers of a phenyl derivative, shown in figure 5 [37]. In terms of acute toxicity to aquatic organisms, there are very few measurements of stereo isomers reported, and those data differ only within their experimental uncertainties.

3.11 *QSARs for regulatory purposes*

Much of the recent developments are in response to the European regulatory REACH (Registration, Evaluation and Authorisation of Chemicals) program. The European Centre for Alternative Methods (ECVAM), primarily interested in replacing animal testing with other methods, is in the forefront of assessing and validating QSAR models of potential use for regulatory purposes. Similar developments are taking place in Japan. With these recent legislative initiatives and desires to assess QSAR

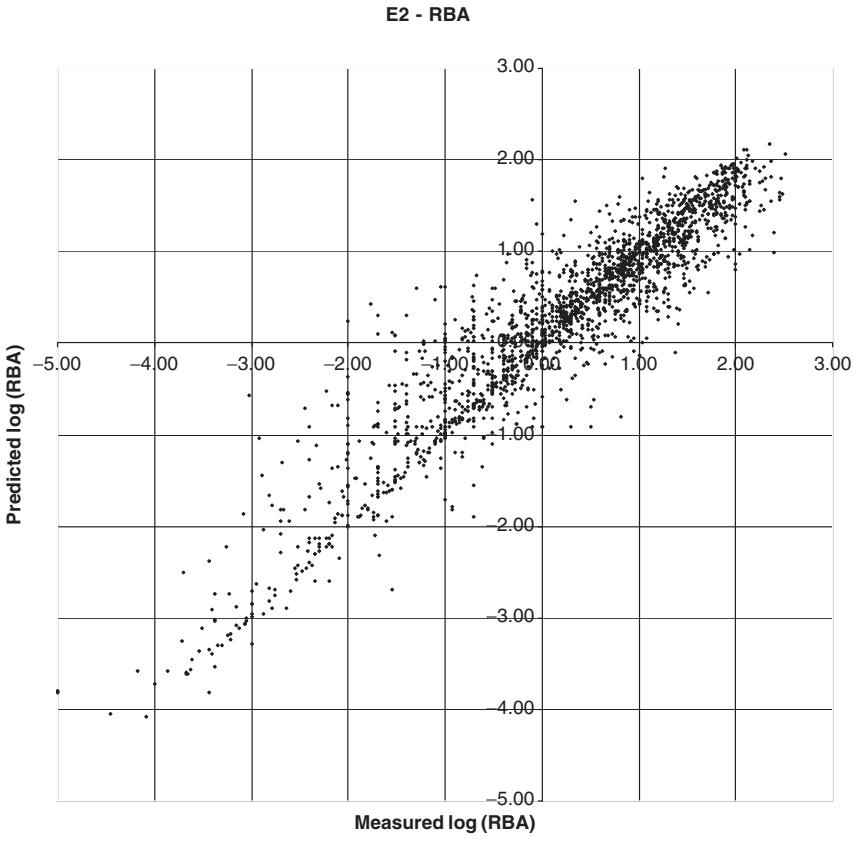

Figure 3. Predicted *vs*. measured estrogen receptor binding affinity (log[E2-RBA]) for approximately 2000 compounds [35].

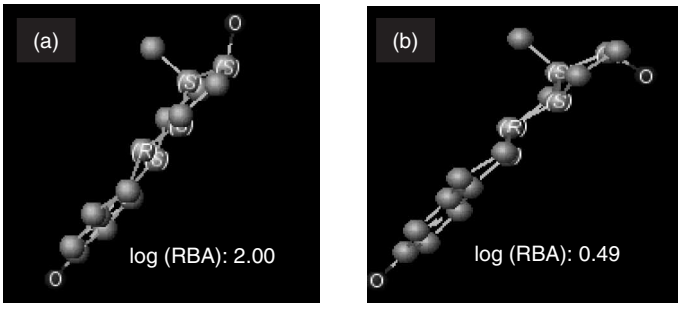

Figure 4. Structures and estrogen receptor binding affinities (RBA) of 17*beta*-estradiol (a) and 17*alpha*-estradiol (b) [32].

methodologies acceptable for regulatory purposes the terms ''applicability domain'', ''chemical category'', or ''classification'' have been coined to describe the types of compounds for which a given QSAR relationship may be applied. Efforts to assess and validate various QSAR algorithms and computer programs rely heavily on the

Table 3. Estrogen receptor binding affinities (E2-RBAs) of four series of low-affinity compounds relative to 17*beta*-estradiol, log(E2-RBA) = 2.00 [36].

Structure group	*log (E2-RBA)*
Phenols	−3.41 to +0.23
Phthalates	−3.49 to −1.15
1,2-Diphenylethanes	−2.96 to +1.57
Benzophenones	−2.09 to −1.03

conditions of "transparency" of such methods and programs [38]. Other methods rely on "rules", "structural alerts" [39], or on a calculated "confidence index" [40]. Unfortunately, even with the limitations of acceptability of an algorithm coerced by such notations, they do not provide any guarantee for the ability of such programs to make correct predictions. This can be demonstrated on a variety of examples. In fact, it has been argued that "the difficulties of articulating the QSAR [applicability] domain will not be overcome by a validation process for QSAR models", as stated by Veith [41].

The question then becomes: are there any "transparent" QSARs which can be used without question to allow the calculation of properties of relevant compounds with certainty. Unfortunately, at least at this time, the answer remains "No". Therefore, I have previously argued that the terms "applicability domain", or "chemical category", and similar nebulous terms are nice concepts in theory, but of little practical use. QSAR models which do not require an *a priori* knowledge of "applicability domain" or anything similar and which rely only on a set of training compounds with various MOAs, etc., despite their "lack of transparency", etc., are the only way out to overcome this conundrum. I do expect this debate to continue for some while as to which type of methodology is to be preferred, particularly for regulatory purposes, but eventually it needs a resolution.

3.12 *Descriptors*

The octanol/water partition coefficient (commonly noted as $\log K_{ow}$ or $\log P$) has been used for several decades to describe the ability of compounds to enter biological systems, such as permeation through cell membranes. There are probably in the order of 10,000+ equations which use $\log P$ as the primary or sole physico-chemical parameter to correlate its variation with the biological effects or other physico-chemical properties of a series of compounds. Many of these QSARs had an enormous impact on the development of novel compounds with desirable medicinal properties. Conversely, linear correlations with $\log P$ brought about the basic principle of assessing the environmental bioconcentration and bioaccumulation potentials of compounds. In addition, the toxicity to aquatic species of chemicals which are acting solely by narcosis can be modeled well by the octanol/water partition coefficient.

As the experimental measurement of $\log P$ can be difficult, frequently the computation with any of the $\log P$ estimation programs is used instead. Unfortunately, the computed $\log P$ values for several types of compounds is not without significant problems. As shown above, even the experimental values for highly chlorinated pesticides have seen revisions by several orders of magnitude [42]. Therefore, it may not surprise that the

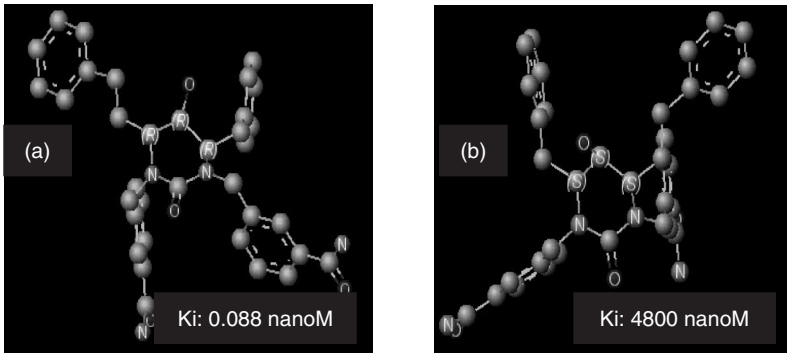

Figure 5. Structures and HIV-1 activities of the R,R,R-isomer (a) and S,S,S-isomer (b) of a cyclic urea derivative [37].

Table 4. Computed octanol/water partition coefficients (CLOGPs) for five *n*-alkanols, computed by six programs, i.e. CLOGP (CLOGP), KOWWIN (KOWWIN), miLogP (MILOGP), SciLOGP (SCILOGP), ALOGPS (ALOGPS), and TerraQSAR - LOGP (TQ-LOGP), the maximum difference in estimated values for each compound (DELTA); N/A: not available; ALOGPS values added here [43].

Compound	*CLOGP*	*KOWWIN*	*MILOGP*	*SCILOGP*	*ALOGPS*	*TQ-LOGP*	*DELTA*
Methanol	−0.76	−0.63	−0.27	−0.66	−1.38	−0.75	0.49
Butanol	0.82	0.84	1.00	0.80	0.84	0.92	0.20
Decanol	4.00	3.79	3.61	4.07	4.24	4.25	0.54
Triacontanol	14.58	13.61	10.83	8.57	10.47	10.11	6.01
Hexacontanol	30.45	28.34	12.73	6.52	10.97	7.36	23.93

computed results may also differ by orders of magnitude. For some types of compounds, such variations are nothing short of astronomical, i.e. well over 20 orders of magnitude, as shown in table 4 [43]. Particularly problematic compounds are many surfactants, especially those with extended alkyl and/or alkoxy chains. Their experimental determination in the water/octanol system is extremely difficult due to micelle formation and their measurement by surrogate methods, such as liquid chromatography [44] also provide poor values due to molecular folding.

It would be beyond the scope of this overview to go into great details of the field of structural descriptors. For a comprehensive review of molecular descriptors, see the handbook by Todeschini and Consonni [45]. Suffice it say that they range far and wide, from bulk parameters, such as $\log P$, molar volume and parachor, to molecular descriptors of many different types. Among the latter, descriptors of physical molecular size, e.g. the Sterimol parameters pioneered by Verloop [46] and applied to a variety of modelling studies by Magee [47], electronic energy levels (e.g. HOMO, LUMO), indicators of certain structural fragments (e.g. –cyano, –nitro), are in common use. With increasing computational capabilities, also parmacophore-specific components and three-dimensional chirality sensitive descriptors find increasing use, e.g. [48]. Of course, our *TerraQSAR – E2-RBA* estrogen receptor binding affinity estimation program is also using such absolute structure information and we have made some 10,000 of these structures freely available for download from our web site in form of

their SMILES strings and associated information [35]. The efforts to develop 3D-SMILES strings for highly chiral compounds, such as the many anti HIV-1 virus active protein-like structures were somewhat hampered by a faulty Accelrys molecular structure viewer product, which switched the chirality under certain circumstances. This problem has recently been corrected [49] and we also recommend the MarvinView program by ChemAxon [50] as a great product.

In terms of database searching tools, for quite some time now, various software providers offer "similarity" as a database searching tool. The user typically can adjust this function to find all compounds with a similar structure to anywhere between 100% and 0% similarity to the entered structure. Most of these programs use the "Tanimoto similarity coefficient", also referred to as the "binary Rogers & Tanimoto similarity coefficient". It is a purely two-dimensional measure and 3D-similarity methods would be welcome. Some recent advances in the field of similarity functions include the "tailored similarity methods" developed by Basak and coworkers [51] and a two-dimensional reduction of three-dimensional structures by Albrecht *et al.* [52].

4. Outlook

The current workshop, the 12th in the series, convened by James Devillers and Alain Carpy is now underway. Once again, it has attracted a substantial crowd of colleagues from many countries. Obviously, the importance the scientific endeavor of finding structure-activity relationships is increasing in importance. The prediction of all kinds of effects and properties of the myriad of chemicals which are known for many years or those which have been developed quite recently, is greatly enhanced by more powerful computers, better models, novel techniques, and so forth. The *Chemical Abstracts* now list over 80 million compounds known to man. Obviously, only a fraction thereof is of potential widespread use. However, even a fraction of, say 1/1000th still leaves some 80,000 compounds for which we need much more knowledge on their effects. This cannot be achieved within a reasonable time by bioassay testing, but only by *in-silico* computation. And for that, non-MOA-based methods are the only practical solution. The scientific debate between the developers of QSAR models using traditional linear-algorithm-based techniques and those using sophisticated new methods, such neural networks, fuzzy theory, etc., will continue for some time to come. In the end, I believe that the twain will meet and more powerful models will emerge. The future for QSAR is bright, better models are emerging every day and a wider regulatory acceptance is on the horizon.

References

[1] R. Carson. *Silent Spring*, The Riverside Press, Cambridge, MA (1962).
[2] S. Jensen, A.G. Johnels, M. Olssen, G. Otterlind. *Nature (Lond.)*, **224**, 247 (1969).
[3] J.J. Hickey, D.W. Anderson. *Science*, **162**, 271 (1968).
[4] J.D. Walker. *QSAR Comb. Sci.*, **22**, 415 (2003).
[5] K.L.E. Kaiser (Ed.). *QSAR in Environmental Toxicology*, D. Reidel Publishing Company, Dordrecht, Holland (1984).
[6] K.L.E. Kaiser (Ed.). *QSAR in Environmental Toxicology – II*, D. Reidel Publ. Co., Dordrecht, Holland (1987).

[7] J.E. Turner, M.W. Williams, T.W. Schultz, N.J. Kwaak (Eds). In *Proceedings of the Third International Workshop on Quantitative Structure-Activity Relationships (QSAR) in Environmental Toxicology*, National Technical Information Service, Springfield, Virginia (1989).
[8] J.L.M. Hermens, A. Opperhuizen (Eds). QSAR in *Environmental Toxicology – IV*, Elsevier, Amsterdam (1991).
[9] F. Chen, G. Schüürmann (Eds). *Quantitative Structure-Activity Relationships in Environmental Sciences – VII*, SETAC Press, Pensacola, Florida (1997).
[10] J. De Bruijn, F. Busser, W. Seinen, J.L. Hermens. *Environ. Toxicol. Chem.*, **8**, 499 (1989).
[11] C. Hansch, A. Leo. *Substituents Constants for Correlation Analysis in Chemistry and Biology*, John Wiley & Sons, New York (1979).
[12] C. Hansch, A. Leo, D. Hoekman. *Exploring QSAR, Hydrophobic, Electronic, and Steric Constraints*, American Chemical Society, Washington, DC (1995).
[13] R.S. Boethling, D. Mackay (Eds). *Handbook of Property Estimation Methods for Chemicals*, Environmental and Health Sciences, Lewis Publishers, Boca Raton (2000).
[14] M. Pavan, A.P. Worth, T.I. Netzeva. Preliminary analysis of an aquatic toxicity dataset and assessment of QSAR models for narcosis. *European Commission* EUR 21749 (2005).
[15] J. Devillers. *SAR QSAR Environ. Res.*, **16**, 433 (2005).
[16] O.A. Raevsky, J.C. Dearden. *SAR QSAR Environ. Res.*, **15**, 433 (2004).
[17] *TerraQSAR^{TM} – FHM*. TerraBase Inc., http://www.terrabase-inc.com (2004).
[18] K.L.E. Kaiser. *Environ. Health Persp.*, **106**, 583 (1998).
[19] S. Dimitrov, Y. Koleva, T.W. Schultz, J.D. Walker, O. Mekenyan. *Environ. Toxicol. Chem.*, **23**, 463 (2004).
[20] I. Lessigiarska, A.P. Worth, B. Sokull-Kluttgen, S. Jeram, J.C. Dearden, T.I. Netzeva, M.T.D. Cronin. *SAR QSAR Environ. Res.*, **15**, 413 (2004).
[21] A. Asfaw, M.R. Ellersieck, F.L. Mayer. *Interspecies Correlation Estimations (ICE) for Acute Toxicity to Aquatic Organisms and Wildlife*, US EPA, EPA/600/C-03/106 (2004).
[22] J.C. Dearden, N.M. Shinnawei. *SAR QSAR Environ. Res.*, **15**, 449 (2004).
[23] R. Posthumus, T.P. Traas, W.J.G.M. Peijnenburg, E.M. Hulzebos. *SAR QSAR Environ. Res.*, **16**, 135 (2005).
[24] *MetabolExpert*. CompuDrug International Inc., http://www.compudrug.com (2006).
[25] Y. Sakuratani, J. Yamada, K. Kasai, Y. Noguchi, T. Nishihara. *SAR QSAR Environ. Res.*, **16**, 403 (2005).
[26] O.G. Mekenyan, S.D. Dimitrov, T.S. Pavlov, G.D. Veith. *SAR QSAR Environ. Res.*, **16**, 103 (2005).
[27] K.L.E. Kaiser. *Water Qual. Res. J. Canada*, **33**, 185 (1998).
[28] F. Ingerslev, B. Halling-Sorensen. *Environ. Toxicol. Chem.*, **19**, 2467 (2000).
[29] R.S. Boethling, D.G. Lynch, J.S. Jaworska, J.L. Tunkel, G.C. Thom, S. Webb. *Environ. Toxicol. Chem.*, **23**, 911 (2004).
[30] K.L.E. Kaiser. *QSAR Comb. Sci.*, **22**, 185 (2003).
[31] *Endocrine Disruptor Knowledge Base*. US Food and Drug Administration, http://edkb.fda.gov/webstart/edkb/index.html (2006).
[32] *TerraTox^{TM} – Steroids* database. TerraBase Inc., http://www.terrabase-inc.com (2006).
[33] M.A. Lill, M. Dobler, A. Vedani. *SAR QSAR Environ. Res.*, **16**, 149 (2005).
[34] T. Ghafourian, M.T.D. Cronin. *SAR QSAR Environ. Res.*, **16**, 171 (2005).
[35] *TerraQSAR^{TM} – E2-RBA*. TerraBase Inc., http://www.terrabase-inc.com (2004).
[36] Y. Akahori, M. Nakai, Y. Yakabe, M. Takatsuki, M. Mizutani, M. Matsuo, Y. Shimohigashi. *SAR QSAR Environ. Res.*, **16**, 323 (2005).
[37] G.V. De Lucca, J. Liang, P.E. Aldrich, J. Calabrese, B. Cordova, R.M. Klabe, M.M. Rayner, C.-H. Chang. *J. Med. Chem.*, **40**, 1707 (1997).
[38] A.P. Worth, C.J. van Leeuwen, T. Hartung. *SAR QSAR Environ. Res.*, **15,**, 331 (2004).
[39] I. Gerner, H. Spielmann, T. Hoefer, M. Liebsch, M. Herzler. *SAR QSAR Environ. Res.*, **15**, 359 (2004).
[40] T.W. Schultz, T.I. Netzeva, M.T.D. Cronin. *SAR QSAR Environ. Res.*, **15**, 385 (2004).
[41] G.D. Veith. *SAR QSAR Environ. Res.*, **15**, 323 (2004).
[42] J. Pontolillo, R.P. Eganhouse. The search for reliable aqueous solubility (Sw) and Octanol-Water partition coefficient (Kow) data for hydrophobic organic compounds: DDT and DDE as a case Study, U.S. Department of the Interior, Water Resources Investigations Report (2001).
[43] K.L.E. Kaiser. *The TerraQSAR Advantage*. Available online at http://www/terrabase-inc.com/tbr-.htm (2004).
[44] C. Leeke, J.R. Dean, W.R. Tomlinson, M.H.I. Comber. In *Quantitative Structure-Activity Relationships in Environmental Sciences – VII*, F. Chen, G. Schüürmann (Eds), pp. 207–218, SETAC Press, Pensacola, Florida (1997).
[45] R. Todeschini, V. Consonni. *Handbook of Molecular Descriptors*, Wiley-VCH, Weinheim (2000).
[46] A. Verloop, W. Hogenstraaten, J. Tipker. In *Drug Design*, E.J. Ariens (Ed.), Vol. 7, pp. 165–207, Academic Press, New York (1976).
[47] P.S. Magee. In *Comparative QSAR*, J. Devillers (Ed.), pp. 137–168, Taylor & Francis, London (1997).

[48] A. Kovatcheva, A. Golbraikh, S. Oloff, J. Feng, W. Zheng, A. Tropsha. *SAR QSAR Environ. Res*., **16**, 93 (2005).
[49] *DS ViewerPro*, vers. 6. Accelrys Inc., http://www.accelrys.com (2005).
[50] *MarvinView*. ChemAxon Ltd., http://www.chemaxon.com (2005).
[51] S.C. Basak, B.D. Gute, D. Mills, D.M. Hawkins. *Theochem*., **622**, 127 (2003).
[52] B. Albrecht, G.H. Grant, W.G. Richards. *Protein Eng. Des. Select*., **17**, 425 (2004).