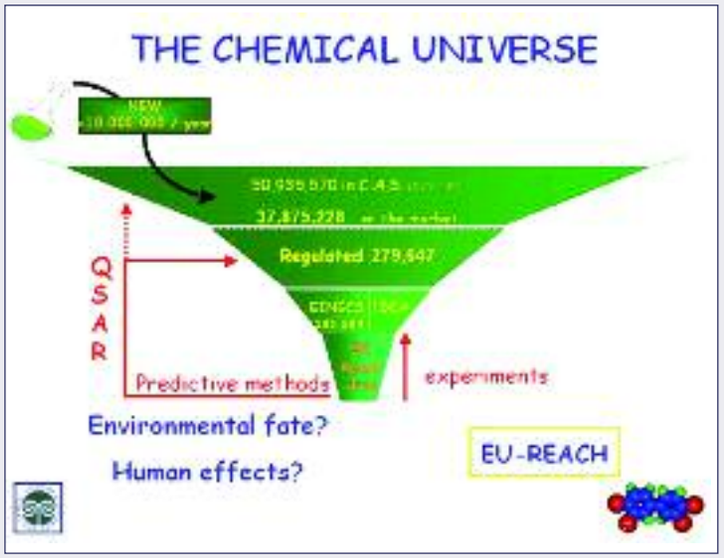


Paola Gramatica
Dipartimento di Biologia Strutturale e Funzionale
Università dell'Insubria
paola.gramatica@uninsubria.it; www.qsar.it

INTRODUCTION TO THE NEED OF ALTERNATIVE METHODS IN REACH

In vitro and in silico methods are foreseen in the EU regulation REACH, to prioritize more dangerous compounds, to focus expensive experiments, by reducing animal tests and to fill the data gaps. A brief historical introduction of QSAR modelling and the importance of model validation according to the OECD principles for reliable predictions is presented.

The chemical universe is huge and is enlarging every day: the CAS registry includes more than 50 million of chemicals, of which nearly 38 million are commercially available and almost 280,000 are regulated and overall listed in the various inventories (for instance the EU-EINECS, US-EPA TSCA, Canada-DSL). New chemicals are being developed continuously (thousands each year), but reports on physico-chemical properties, and biological activities are more slowly produced. While the degree of knowledge for "new" chemicals could be acceptable, the same cannot be stated for the majority of the "existing" chemicals in commerce, even for the High Production Volume (HPV) compounds; thus, there is generally a lack of sufficient information publicly available in order to assess and control these substances effectively. The problem of lack of data and slow assessment procedures is huge: we know a lot about a few chemicals (<5%), but we have very little information on the properties and risks of most (>95%) chemicals.

In Europe the new legislation REACH (Registration Evaluation Authorization and Restriction of Chemicals) [1] creates a single system for the so-called "existing" and "new" substances (respectively, around 100,000 chemicals, put on the market before 1981 and around 5,000 chemicals introduced after 1981) to obtain from industries relevant information on properties and activities of all the commercialised substances and to use that data to manage them safely. The principal aim is to ensure greater safety in the manufacture and use of chemicals, a high level of protection of health and the environment.

REACH reverses the responsibility for providing the necessary information and taking effective risk management measures to industry, both producers and importers of substances, rather than the public authorities, as in the past. REACH requires manufacturers and importers to gather comprehensive information on properties of their substances produced or imported in volumes over 1 tonne per year, into different deadline dates, and to submit the necessary information to demonstrate their safe use in a registration dossier to the European CHemicals Agency (ECHA). Public authorities will examine registration dossiers and substances of concern and they will also scrutinise all proposals for animal testing to keep it to the minimum absolutely necessary. Use-specific authorisations will be required for Substances of Very High Concern (SVHC), that cause cancer, mutations or reproduction problems (CMRs, including endocrine disruptors ED), or that highly accumulate and persist in our bodies and in the environment (vPvB)

or are also toxic (Persistent Bioaccumulative and Toxics, PBT). Authorisation will be granted only to companies that can show that the risks are adequately controlled or if social and economic benefits outweigh the risks and suitable alternative substances do not exist. This will encourage substitution of unsafe substances by safer ones.

Two levels of actions can be identified in relation to the highest concern chemicals: a) the need of tools for their identification and prioritization, and b) the stimulus for the research and production of safer alternatives. It is immediately evident the enormous efforts required to the industry in order to fill the huge gap in data availability into short terms and also in order to design safer alternative chemicals.

Costs, numbers of test animals, speed of the process, and sharing the data are all important issues. It is clear that there is a need to increase the efficiency, cost effectiveness and focus of the risk assessment process, while reducing the current reliance on animal tests.

It is now widely acknowledged that the most efficient way to carry out hazard and risk assessments of large numbers of chemicals, while reducing costs to industry and minimising animal testing, is to obtain the necessary information by means of Intelligent Testing Strategies (ITS).

Intelligent testing strategies are integrated approaches comprising of multiple elements aimed at speeding up the risk assessment process while reducing costs and animal tests:

1) *in vivo and in vitro* experimental tests;
2) computational methods (SARs, QSARs and biokinetic models, Read-across) and chemical categories;
3) exposure assessment.

The ultimate aim of all these approaches is to obtain reliable information on the (toxic) properties of chemicals with minimal use of animals.

To ensure that animal testing is kept to the strict minimum general rules are also set out:

a) for the use of existing information (*data sharing*);
b) for waiving of tests (omitting them if they are not required because of their use or it is not technically possible to carry them out);
c) for the development of alternative *in vitro* tests;
d) for *in silico* techniques such as (Q)SARs (Quantitative Structure-Activity Relationship) and read across.

New animal tests are only required when it is not possible to provide the information in any other permitted way. The underlying strategy is the 3R approach: Reduce, Refine, Replace.

The development and use of alternative methods (e.g. QSAR, etc.) for the assessment of hazards of substances is expressly promoted and inserted in REACH articles (1, 40 and 47). In particular, QSAR models are cited in REACH text in Art. 13, 138, in Annex III, VI-XI.

The use of predictive QSAR models is suggested:

1) to highlight dangerous chemicals;
2) to support priority setting of chemicals and to focus the experimental tests;
3) to fill in data gaps for classification and labelling and for risk assessment;
4) to design safer alternative compounds.

Computational methods: QSAR approach

It is now widely acknowledged that the most efficient way to carry out hazard and risk assessments of large numbers of chemicals, while reducing costs to industry and minimising animal testing, is to obtain the necessary information by means of computational or “in silico” methods like QSARs. QSARs are mathematical models that relate a numerical measure of chemical structure (a molecular descriptor) to a chemico-physical property or to a biological effect (e.g. a toxicological endpoint) and that can be used to predict the unknown activities and properties of molecules.

Computational QSAR models validated according to the recent OECD principles [2] can be employed to pre-screen the compounds, by predicting their activity, and to examine with laboratory tests only those compounds that possess biological activity according to the QSAR models. Moreover, QSAR models can also be employed by chemical industries in the development of new safer chemicals, without dangerous properties, as it is widely applied since long time in pharmaceutical companies for drug design.

QSARs are based on the assumption that the structure of a molecule (i.e. its geometric, steric and electronic properties) must contain the features responsible for its physical, chemical, and biological properties, and on the ability to represent the chemical by one, or more, numerical descriptor(s). By QSAR models, the biological activity (or property, reactivity, etc.) of a new or untested chemical can be inferred from the molecular structure of similar compounds whose activities (properties, reactivities, etc.) have already been assessed. The QSPR (Quantitative Structure-Property relationship) acronymous is used when a property is modeled.

It has been nearly 40 years since the QSAR modelling firstly was used into the practice of agrochemistry, drug design, toxicology, industrial and environmental chemistry . Its growing power in the following years may be attributed also to the rapid and extensive development in methodologies and computational techniques that have

allowed to delineate and refine the many variables and approaches used in this modelling approach.

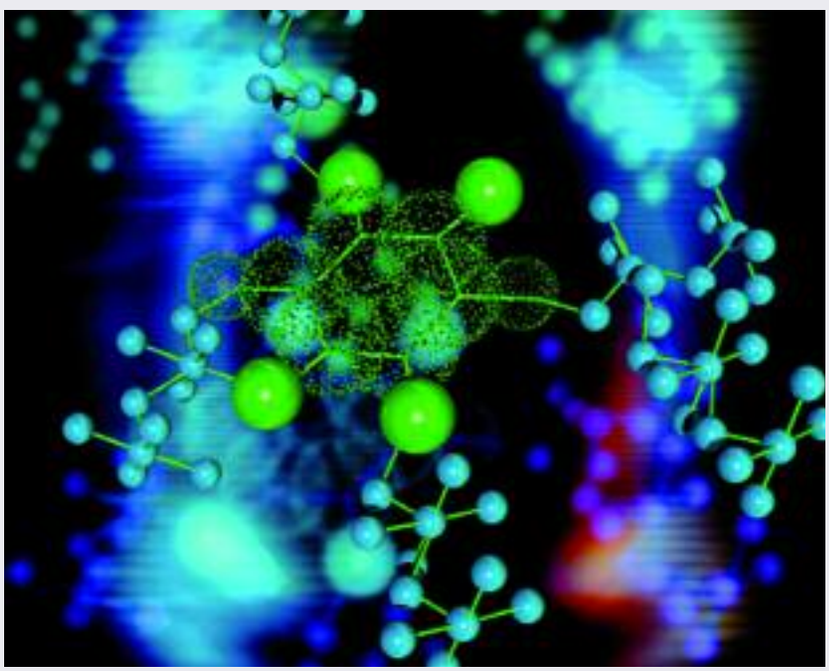
QSAR modelling is born in toxicology field. In 1863, Cros noted that a relationship existed between the toxicity of primary aliphatic alcohols and their water solubility. This relationship demonstrated the central axiom of structure-toxicity modelling - the toxicity of substances is governed by their properties, which in turn are determined by their chemical structure. Therefore, there are interrelationships between chemical structure, properties, and toxicity.

At the turn of the 20th century, Meyer and Overton [3, 4] independently suggested that the narcotic (depressant) action of a group of organic compounds paralleled their olive oil/water partition coefficients. In following years on the physical organic front, the seminal work of Hammett gave rise to the "σ-ρ" culture [5,6] in the delineation of substituent effects on organic reactions, while Taft devised a way for separating polar, steric, and resonance effects and introducing the first steric parameter, *ES* [7] .

There is a consensus among current predictive toxicologists that Corwin Hansch is the founder of modern QSAR. In the classic article [8] it was illustrated that, in general, biological activity for a group of 'congeneric' chemicals can be described by a comprehensive model:

$$\text{Log } 1/C_{50} = a\pi + b\varepsilon + cS + d \qquad (1)$$

in which C, the toxicant concentration at which an endpoint is manifested (e.g. 50% mortality or effect), is related to a hydrophobicity term, π, (this is a substituent constant denoting the difference in hydrophobicity between a parent compound and a substituted analog, it has been replaced with the more general molecular term the log of the 1-octanol/water partition coefficient, log Kow), an electronic term, ε, (originally the Hammett substituent constant, σ) and a steric term, S, (typically Taft's substituent constant, ES).



At present, the QSAR science, founded on the systematic use of mathematical models and on the multivariate point of view, is one of the basic tools of modern drug and pesticide design, has an increasing role in environmental sciences and is suggested in REACH.

QSAR models exist at the intersection of chemistry, statistics and biology, in toxicological studies. The development of a new QSAR model requires these three components: 1) a data set that provides experimental measures of a biological activity or property for a group of chemicals; 2) molecular structure and/or property data (i.e. the descriptors, variables, or predictors) for this group of chemicals; and 3) statistical methods, to find the relationship between these two data sets.

The limiting factor in the development of QSARs is the availability of high quality experimental data. In QSAR analysis, it is imperative that the input data be both accurate and precise to develop a meaningful model. In fact, it must be realized that any resulting QSAR model that is developed is only as valid statistically as the data that led to its development.

Data used in QSAR evaluations are obtained either from the literature or generated specifically for QSAR-type analyses. These data can consist of congeneric series of chemicals (local QSAR models) or assure structural diversity even within a chemical class (general QSAR models). This diversity has allowed the generalization of more robust QSARs, applicable in an extended way. A structure- activity model is defined and limited by the nature and quality of the data used in model development and should be applied only within the model's applicability domain.

The ideal QSAR should: (1) consider an adequate number of molecules for sufficient statistical representation, (2) have a wide range of quantified end-point potency (i.e. several orders of magnitude) for regression models or adequate distribution of molecules in each class (i.e. active and inactive) for classification models, (3) be applicable for reliable predictions of new chemicals (validation and applicability domain) and (4) allow to obtain mechanistic information on the modelled end-point.

Chemical descriptor(s) include empirical, quantum chemical, or non-empirical parameters. Empirical descriptors may be measured or estimated and include physico-chemical properties (such as for instance logP). Non-empirical descriptors can be based on individual atoms, substituents, or the whole molecule, they are typically structural features. They can be based on topology or graph theory and, as such, they are developed from the knowledge of 2D structure, or they can be calculated from the 3D structural conformations of a molecule.

A variety of properties have been also used in QSAR modelling, these include physico-chemical, quantum chemical,

and binding properties. Examples of molecular properties are electron distribution, spatial disposition (conformation, geometry, and shape), and molecular volume. Physicochemical properties include descriptors for the hydrophobic, electronic, and steric properties of a molecule as well as other properties including solubility and ionization constants. Quantum chemical properties include charge and energy values. Binding properties involve biological macromolecules and are important in receptor-mediated responses.

In modern QSAR approaches, it is becoming quite common to use a wide set of theoretical molecular descriptors of different kinds, able to capture all the structural aspects of a chemical to translate the molecular structure into numbers. Different descriptors are different ways or perspectives to view a molecule, taking into account the various features of its chemical structure, not only mono-dimensional as the simple counts of atoms and groups, but also bi-dimensional from the topological graph or three-dimensional from a minimum energy conformation. A lot of software calculates wide sets of different theoretical descriptors, from SMILES, 2D-graphs to 3D-x,y,z-coordinates, obtained by geometry optimization methods (MM+, AM1, ecc). Some of the more used are mentioned here: CODESSA [9], MolConnZ [10], and DRAGON [11]. It has been estimated that more than 3,000 molecular descriptors are now available, and most of them have been summarized and explained [12-14]. The great advantage of theoretical descriptors is that they can be calculated homogeneously by a defined software for all chemicals, even those not yet synthesized, the only need being a hypothesized chemical structure, thus they are reproducible.

Modelling methods used in the development of QSARs are of two types in relation to the modelled response: a potency of an end-point (a defined value of EC50) or a category/class (for instance Mutagen/Not mutagen).

For the potency modelling, the most widely used mathematical technique is multiple regression analysis (MRA). Regression analysis is a simple approach that leads to a result that is easy to understand and, for this reason, most QSARs are derived using regression analysis.

Regression analysis is a powerful means for establishing a correlation between independent variables (molecular descriptors X) and a dependent variable Y, such as biological activity:

$$Y = b + aX_1 + cX_2 + \ldots \quad (2)$$

For the modelling of categories, different quantitative models of classification can be applied. A wide range of classification methods exists, including: discriminant analysis (DA; linear , quadratic, and regularized DA), SIMCA (Soft Independent Modelling of Class Analogy), k-NN (k-Nearest Neighbours), CART (Classification And Regression Tree), Artificial Neural Network, Support Vector Machine, etc. In these techniques, the term “quantitative” is referred to the numerical value of the variables (the molecular descriptors) necessary to classify the chemicals in the qualitative classes.

It is evident from the literature analysis that the QSAR world has undergone profound changes since the pioneering work of Corwin Hansch, considered the founder of modern QSAR modelling [8]. The main change is reflected in the growth of a parallel and quite different conceptual approach to the modelling of the relationships among a chemical’s structure and its activity/properties.

In the Hansch approach (so called mechanistic), still applied widely and followed by many QSAR modelers, molecular structure is represented by only a few molecular descriptors (typically log Kow, Hammett constants, homo/lumo, some steric parameters) selected personally by the modeler and inserted in the QSAR equation to model a studied end-point. Alternatively, in a different approach (so called statistical) the chemical structure is represented, in the first preliminary step, by a large number of theoretical molecular descriptors which are then, in a second step, selected by different chemometric methods (such as for instance evolutionary techniques, Genetic Algorithms, etc) as the best correlated with response and included in the QSAR model (the algorithm). Thus, descriptor selection is performed with the final and crucial aim to maximize, as an optimization parameter, the predic-

tive power of the QSAR model, as the real utility of any model is considered its predictivity.

In fact, the first aim of any modeler should be validation for the predictive application of the QSAR model, for both the mechanistic approach and the statistical one. The famous "Kubinyi Paradox" [15] , emphasized also by Tropsha et al. in their famous papers: *Beware of* Q^2 [16] and *The Importance of being Earnest* [17] is that: The "best fit" models are not the best ones for prediction! In fact, a QSAR model must, first of all, be a real model, robust and predictive, to be considered a reliable model [18] ; only a stable and predictive model can be usefully interpreted for its mechanistic meaning, even so this is not always easy or feasible.

Validation and reliability of QSARs

Obviously, to meet the requirements of the REACH legislation it is essential to use (Q)SAR models that produce reliable estimates, i.e., validated (Q)SAR models. QSAR model validation has been recognized by specific OECD expert groups as a crucial and urgent point in recent years, and this has led to the development, for regulatory purposes, of the "OECD principles for the validation of (Q)SAR models" [2] .Thus, reliable QSAR model must 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.

The need for descriptor interpretability depends on the application, as a validated mathematical model relating a target property to chemical features may be all that is necessary, particularly when predicted data are needed for screening of large libraries of chemicals, though it is obviously desirable to attempt some explanation of the 'mechanism' in chemical terms [19, 20].

It needs to be recognised that the predictions, made by estimation techniques, cannot be more reliable than the experimental data they are trying to predict and that the error of predicted data are generally of the experimental error order. Nevertheless, a (regulatory) decision based on the use of a reliable model (developed by using high quality data and verified by statistical validation) will be more useful than a decision based on poor quality experimental data (animal tests).

Secondly, the answer depends on the endpoint. Estimation methods are very reliable for most of the physicochemical endpoints. They are reliable for many of the environmental endpoints for an important part of the chemicals universe. They are less reliable for complex toxicological endpoints related to human health, but highly useful for priority setting purposes.

Thirdly, these tools can only be applied by experts, knowing the applicability domains of these tools and their limitations. In other words for these tools it is important to know what they are doing, and more importantly, what you should not do. The message is: leave their use to experts!.

A common criticism of (Q)SARs, when considered in isolation, is that

they are only useful for making reliable estimations for limited classes of chemicals (the so-called "applicability domain"). For this reason, an important issue in the validation of a (Q)SAR model is the establishment of its applicability domain. The important question is not only whether a given (Q)SAR model is valid, but for which classes of chemicals it can be used to make reliable predictions. In vitro tests are also associated with applicability domains which define limitations in their applicability to classes of chemicals.

It will not be possible to completely replace animal tests in REACH, but if different types of alternative approaches are combined in an intelligent manner, the number of animal tests required can be reduced significantly. (Q)SARs and in vitro tests will be useful as partial replace-

ments of animal tests. In particular QSAR methods can be applied for screening chemicals and prioritizing the more dangerous for experimental tests; additionally, can be applied in the design of safer alternative chemicals. No method other than QSAR is applicable to chemical design and to detect *a priori*, only from the drawn structures, the property/activity of any compound.

A lot of commercial software are available: in those packages QSAR models have been already developed by QSAR modellers and the user can just press a button for obtaining a predicted value of the requested endpoint for a requested chemicals. However, the great danger is that the quality of the obtained prediction is not always clear and the reliability of the data are not guarantee. Also the application of commercial software should be limited to QSAR experts.

QSAR experts for REACH in Italy

In Italy, few groups of QSAR experts are involved in scientific works and projects for the development of QSAR models, validated according to the OECD principles [2], with the aim to be applicable in REACH. In particular: the Benfenati group in Mario Negri Institute (Milan) has been published many papers on this topic [as an example 21], has been involved in different EU-Projects, more recently in CAESAR [22]; Benigni in Istituto Superiore di Sanità (Rome) has a very long activity on the QSAR modelling of chronic effects (carcinogenicity, mutagenicity [23]) and is currently involved in a FP7-EU project for REACH OpenTox [24]; Gramatica in Insubria University has published a lot of papers on QSAR models developed according to the OECD principles of validation [18] (mainly on POPs [25] and Endocrine Disruptors (EDs) [26]), she is now leader of the QSAR workpackage in the FP7-EU project for REACH CADASTER [27]; Todeschini in Milano Bicocca is mainly interested in molecular descriptors [14] and modelling software [11]. He is now coordinator of a PRIN07 project, with Gramatica, on the development of QSAR models for REACH focused on PBTs and EDs. The identification and assessment of PBTs and EDs, on which topics Insubria group has a long activity, is an important task under REACH, as they are included in the Authorization procedure.

References

[1] http://europa.eu.int/comm/environment/chemicals/Reach.htm)
[2] www.oecd.org/document/23/0,2340,en_2649_201185_33957015_1_1_1_1,00.html
[3] H. Meyer, *Arch. Exp. Pathol. Pharmakol.,* 1899, **42,** 109.
[4] C.E. Overton, *Studien Uber die Narkose*, Fischer, Jena, Germany, 1901.
[5] L.P. Hammett, *Chem. Rev.,* 1935, **17,** 125.
[6] L.P. Hammett, Physical Organic Chemistry, 2nd Ed., McGraw-Hill, New York, 1970.
[7] R.W. Taft, *J. Am. Chem. Soc*., 1952, **74,** 3120.
[8] a) C. Hansch *et al., Nature*, 1962, **194,** 178; b) C. Hansch, T. Fujita, *J. Am. Chem. Soc.,* 1964, **86,** 1616.
[9] A.R. Katritzky, V.S. Lobanov, CODESSA, Version 5.3, University of Florida, Gainesville, 1994.
[10] MolConnZ, Ver. 4.05, 2003, Hall Ass. Consult., Quincy, MA.
[11] R. Todeschini *et al.,* DRAGON-Software for the calculation of molecular descriptors. Ver. 5.4 for Windows, 2006, Talete srl, Milan, Italy.
[12] J. Devillers, A.T. Balaban, (Eds.) Topological Indices and Relat ed Descriptors in QSAR and QSPR, Amsterdam, Gordon BReach Sci. Pub., 1999.
[13] M. Karelson, Molecular Descriptors in QSAR/QSPR, Wiley-InterScience, New York, 2000.
[14] R. Todeschini, V. Consonni, Handbook of Molecular Descriptors, Wiley-VCH, Weinheim (Germany), 2000.
[15] a) J.H. van Drie, *Curr. Pharm. Des.,* 2003, **9,** 1649; b) J.H. van Drie, in Computational Medicinal Chemistry for Drug Discovery, P. Bultinck *et al.* (Eds.), Marcel Dekker, 2004.
[16] A. Golbraikh, A. Tropsha, *J. Mol. Graph Mod.,* 2002, **20,** 269.
[17] A. Tropsha *et al., QSAR Comb. Sci.,* 2003, **22,** 69.
[18] P. Gramatica, *QSAR Comb. Sci.,* 2007, **26**, 694.
[19] D.J. Livingstone, *J. Chem. Inf. Comput. Sci.*, 2000, **40,** 195.
[20] N.S. Zefirov, V.A. Palyulin, *J. Chem. Inf. Comput. Sci.*, 2001, **41,** 1022.
[21] C.Y. Zhao *et al., Chemosphere,* 2008, **73,** 1701.
[22] www.caesar-project.eu/index.php
[23] R. Benigni, C. Bossa, *Toxicology Mechanism and Methods*, 2008, **18,** 137.
[24] www.opentox.org
[25] P. Gramatica, E. Papa, *Environ. Sci. Technol*., 2007, **41,** 2833.
[26] H. Liu *et al., Chem. Res. Toxicol.,* 2006, **19,** 1540.
[27] www.cadaster.eu

RIASSUNTO

La necessità di metodi alternativi per il REACH

I metodi in vitro ed in silico sono previsti nella legislazione europea REACH per prioritizzare i composti più pericolosi, per focalizzare esperimenti costosi, riducendo i test su animali, e per colmare la mancanza di dati. Viene qui presentata una breve introduzione storica sulla modellistica QSAR e l'importanza della validazione di tali modelli secondo i principi OECD, per ottenere dati predetti attendibili.