FORUM

The Practice of Structure Activity Relationships (SAR) in Toxicology

James D. McKinney,[*,1] Ann Richard,[*] Chris Waller,[†] Michael C. Newman,[‡] and Frank Gerberick[§]

*United States Environmental Protection Agency/NHEERL, 111 Alexander Drive, Research Triangle Park, North Carolina 27711;
†Sphinx Pharmaceuticals, Inc., Research Triangle Park, North Carolina 27709; ‡Virginia Institute of Marine Science,
Gloucester Point, Virginia 23062; and §The Procter & Gamble Company, Cincinnati, Ohio 45253

Both qualitative and quantitative modeling methods relating chemical structure to biological activity, called structure-activity relationship analyses or SAR, are applied to the prediction and characterization of chemical toxicity. This minireview will discuss some generic issues and modeling approaches that are tailored to problems in toxicology. Different approaches to, and some facets and limitations of the practice and science of, SAR as they pertain to current toxicology analyses, and the basic elements of SAR and SAR-model development and prediction systems are discussed. Other topics include application of 3-D SAR to understanding of the propensity of chemicals to cause endocrine disruption, and the use of models to analyze biological activity of metal ions in toxicology. An example of integration of knowledge pertaining to mechanisms into an expert system for prediction of skin sensitization to chemicals is also discussed. This minireview will consider the utility of modeling approaches as one component for better integration of physicochemical and biological properties into risk assessment, and also consider the potential for both environmental and human health effects of chemicals and their interactions.

Key Words: structure-activity relationships (SAR); SAR science; elements; models; prediction systems; issues in toxicology.

Why SAR?

Structure-activity relationships (SARs) are basic to toxicological investigations. Biological properties of new compounds are often inferred from properties of similar existing materials whose hazards are already known. However, toxicologists today are faced with the task of screening large numbers of diverse chemicals in different media, for an increasing array of toxicity endpoints, using limited resources and fewer animals. Animal and *in vitro* testing are still considered essential to the

support of risk assessment and regulatory action, but are often too costly and time consuming to be applied to the full range of chemicals for which some level of toxicological screening is necessary and desired. Computer-based modeling methods relating chemical structure to qualitative biological activity (SAR) and quantitative biological potency (QSAR) have been applied in many diverse problem settings. The resulting models are aimed toward the prediction and characterization of chemical toxicity (Golberg, 1983; Haque, 1980; Hansch and Leo, 1995; Hermens and Opperhuizen, 1991; Kaiser, 1987; Karche and De Villers, 1990; McKinney, 1985; Rand and Petrocelli, 1985). In addition, with accelerating trends toward improved understanding of the chemical mechanisms of toxicological endpoints and consolidation of toxicological data into databases, there are enhanced opportunities to incorporate such methods into existing toxicological investigations. Hence, it is important for both those who plan to use SAR models and those who plan to develop them to have a basic understanding of how an SAR model is constructed, as well as to learn the limits and potential of the technology.

This minireview is in large part a summary of material provided in a continuing education course at the Society of Toxicology 38th Annual Meeting in March 1999, entitled "The Practice of Structure Activity Relationships in Toxicology," and it is not intended to be an exhaustive review of the subject area. It will discuss some of the modeling approaches that are tailored to issues in toxicology and will stress QSAR as a valuable complement to experimental data and as a departure point for further inquiry into molecular mechanisms. Examples illustrate different approaches to, and some facets of, the practice of SAR as they pertain to current *in vitro* and *in vivo* toxicology analyses. Topics include the science of SAR in the context of toxicology and important elements for sound application, special application of 3D SAR methods and approaches, use of models to analyze biological activity of metal ions in toxicology, and the application of expert systems for screening and prediction of toxicologic outcome. There is

growing awareness (Conolly *et al.*, 1999; McKinney, 1996) of the importance of basic research on mechanisms of toxic action of chemicals as a means for enhancing understanding and providing a more rational basis for risk assessment. Structure-based modeling approaches are one component for better integration of physicochemical and biological properties into risk assessment.

Background

A structure activity relationship relates features of a chemical structure to a property, effect, or biological activity associated with that chemical. In so doing there can be both qualitative and quantitative considerations. The fundamental premise is that the structure of a chemical implicitly determines its physical and chemical properties and reactivities, which, in interaction with a biological system, determine its biological/toxicological properties. The process of developing a SAR is one of attempting to understand and reveal how properties relevant to activity are encoded within and determined by the chemical structure.

In the pharmaceutical and chemical industries, SARs have long been used to design chemicals with commercially desirable properties. This has been particularly the case in the area of drug design where chemicals with desired pharmacologic and therapeutic activities are sought. In the environmental health protection field, SAR is being used to predict ecological and human health effects, with applications varying widely. It is even being used to help industry design safer chemicals for commercial use as a part of their desirable properties.

Why should toxicologists be interested in SAR? Toxicologists generally operate in the domain of single-chemical investigations within a particular biological system. SAR offers a means for relating toxicological data across a spectrum of chemicals, and possibly biological endpoints, illuminating associations that transcend the particulars of single-chemical toxicological experiments, and conceivably revealing aspects of toxicological mechanisms that can be generalized across chemicals. When used in a predictive capacity, SARs have the potential to reduce the need for property measurements and animal testing, providing for more efficient screening of chemicals for a wide range of toxicity endpoints. This can ultimately lead to better environmental health protection through strategic application of limited resources aimed toward identifying the greatest chemical hazards.

The Science of SAR

SAR resides at the intersection of biology, chemistry, and statistics (Fig. 1). The focused linkage of these disciplines brought about through SAR activities has permitted the development of a research activity resembling the "science of SAR" (Hansch, 1969; Hansch *et al.*, 1989; Hermens, 1996; Topliss and Edwards, 1979). In relating structure to activity, the goal of SAR is to generalize across and outward from specific cases,
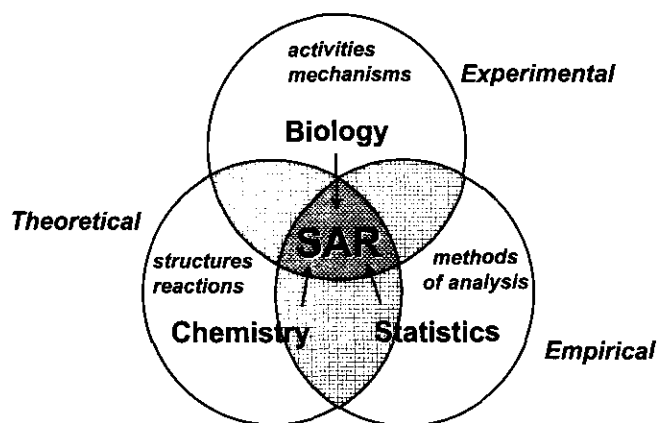

FIG. 1. SAR resides at the intersection of biology, chemistry, and statistics.

developing an understanding of what constitutes a class of molecules that are active, what determines relative activity, and what distinguishes these from inactive classes. Included under the heading of SAR are activities ranging from the use of heuristics and expert judgment, to considerations of similarity/diversity of chemicals, to formal mathematical associations of properties and activity measures. The fundamental assumption in QSAR is that similar chemicals have sufficiently common mechanistic elements so as to share a common rate-determining step and similar energy requirements for activity. It is further assumed that differences in reaction rates will give rise to observed differences in activity or quantitative potency. The key is to identify aspects of structure pertaining to the rate-determining, molecular-triggering event in the mechanism of action for the chemical and biological actions of interest. Hence, the mechanism of action is a guiding concept in determining both the groupings of chemicals suitable for study and the molecular descriptors potentially most relevant to activity. Ultimately, it is the linkage of SAR to mechanism that enables a scientific rationale to be constructed to account for activity variations in existing chemicals. This, in turn, provides the most sound scientific basis for predicting the activity of new and untested chemicals. Having stated the ideal case, we are faced with the reality that many toxicity endpoints are complex, often poorly understood and characterized, and not resolvable to the level of a common mechanism of action. To the extent that we can resolve the toxicity problem, SARs may be capable of global discrimination among different mechanisms, e.g., categorizing by structural alerting fragments, and/or local discrimination within a more well-defined, mechanism-based class.

A good illustration of the former is provided by a recent report of enhanced MultiCASE models, developed in collaboration with the FDA, for predicting rodent carcinogenicity potential for pharmaceutical databases (Matthews and Contrera, 1988). The classifications are based primarily on MultiCASE-identified structural alerting features of molecules, and
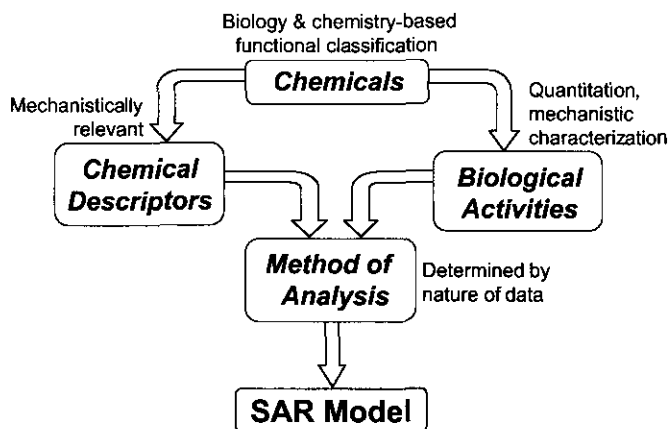
FIG. 2. Important elements in developing mechanistic SARs.

are potentially applicable to rough screening of a wide diversity of chemical structures and mechanisms of carcinogenicity. In contrast, a prominent example of a mechanism-based SAR application that has impacted risk assessment is the modeling of Ah receptor-binding capability of dioxin-like compounds, including the structurally related polychlorinated dioxins, dibenzofurans, and biphenyls, in particular. Extensive studies by Safe and coworkers (1990, 1994) and others, demonstrating rank order correlations between Ah-receptor binding and various measures of response, both *in vivo* and *in vitro*, were effective in establishing a common mechanism of action for these toxic responses. This, in turn, led to the development and use of toxic equivalency factors (TEFs) to arrive at concentrations of dioxin equivalents (TEQs) in human and ecological risk assessments involving exposures to complex mixtures of these compounds as they occur in real-world environments (Van den Berg *et al.*, 1998). In deriving TEF values, a variety of available data, including from *in vivo*, *in vitro*, and QSAR studies are usually weighed in using a tiered approach. This overall approach is basically a relative potency-ranking scheme in which the relative potency of each chemical is expressed as some fraction of the potency of 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD). A mechanism-based SAR analysis was particularly important in recognizing the close structural resemblance of the "coplanar PCBs" to TCDD and their associated, highly toxic properties (McKinney *et al.*, 1981).

Elements of SAR

There are several important elements to keep in mind in working toward the development of mechanistically based SARs (Fig. 2). As indicated above, it is desirable, when possible, to develop a mechanistic classification of the biological/toxicological activity of interest. This, in turn, determines the most relevant chemicals, associated properties, and descriptors to study, pertaining to the controlling/discriminating step(s) for the activity of interest. In addition, there are descriptors that are generically important for approximating the ability of a chem-

ical to reach the site of action; the most prominent example of such a descriptor is the octanol/water partition coefficient (Hansch and Dunn, III, 1972). Other ways of representing molecules may extend beyond those based on 2D structure, atoms and bonds, to those based on 3D structure, steric and electrostatic fields. The latter are most appropriate if a receptor-mediated mechanism is known or suspected. Finally, appropriate methods of analysis are needed for relating the activities and chemical structures of interest, which will depend on the nature of the activity measure (e.g., qualitative versus quantitative), and the extent to which the chemical mechanism of action is understood (e.g., receptor-mediated), etc. The goal is to strive at every step in the process to consider what is chemically and biologically plausible, to reasonably constrain the problem in these terms, and to derive models that have a strong scientific rationale and basis for interpretation.

SAR Models

An SAR model is defined and limited by the nature and quality of the data used in model development and is strictly applicable only in relation to the data set that was used to generate it, but that possibly has predictive capability within some reasonable boundary outside that data set. In evaluating an SAR model, it is important to define boundaries of application, by considering what sorts of molecules, and range of descriptor values, have activities that can be confidently predicted, and statistical measures of fit, significance, and robustness. Models can also lead to mechanistic hypotheses that guide future testing and validation. A process for model validation should test predictive capability, as well as explore the boundaries for model application and challenge the mechanistic hypotheses suggested by a well-constructed model.

SAR models are useful in research for purposes beyond prediction. They can offer rationalization of activity variations in existing data, argue for a common mechanism of activity (and additivity of effect) for a series of chemicals (Richard and Hunter III, 1996), identify outliers due to either experimental error or alternative mechanisms (Lipnick, 1991), narrow a dose range-finding experiment (by using a predicted dose as a first estimate), serve as a metric for comparison of different biological endpoints (Hansch *et al.*, 1995), and direct further research. The ideal SAR model should consider sufficient numbers of molecules for adequate statistical representation, have a broad range of quantitative activities (orders of magnitude) or adequate distribution of molecules in each activity class (active and inactive), and yield to mechanistic interpretation (Hermens, 1996). In toxicology modeling problems, this ideal is rarely encountered. For many toxicity endpoints of interest, diverse chemical structures, lack of knowledge of mechanisms, and large data gaps are more frequently the norm. These limitations on our ability to construct "classical" QSAR relationships, i.e., based on well-defined chemical classes, have led to various attempts to develop "global" SAR prediction

models for what are termed non-congeneric chemicals, i.e., large sets of structurally and mechanistically diverse chemicals (for some reviews, see, e.g., Benfenati and Gini, 1997; Benigni and Richard, 1996, 1998). Because SAR ultimately draws its validity from linkage to mechanism, however, any success achieved with these methods rests on the degree to which the global models are able to discern and adequately represent the mechanism-based SAR components of the larger data set (Lewis, 1992; Richard, 1995; Wagner *et al.*, 1995).

Prediction Systems

Two main types of commercial toxicity prediction systems are currently available: the correlative or statistically based programs and the rule-based expert systems (see Benfenati and Gini, 1997; Chapter 6 in Hansch and Leo, 1995; and Richard, 1998a,b). Correlative systems, such as CASE/MultiCASE (Klopman, 1984) and TOPKAT (Enslein, 1993), typically process a large group of non-congeneric chemicals, without user bias or prior organization, and attempt to extract SAR associations from the data by statistical means. The biggest drawback of such systems is the ease with which a prediction is generated versus the need for careful scrutiny of the results. Typically, such methods are better at gross identification of "alerting" classes than at discerning finer activity variations within these classes. Rule-based systems, such as DEREK (Sanderson and Earnshaw, 1991) and ONCOLOGIC (Woo *et al.*, 1995), build associations and generalizations from small groups of chemicals, group similar-acting chemicals into classes based on organic chemistry definitions and limited mechanistic understanding, and use expert judgment and mechanism-based rationale within the classes. The rule-based systems typically are more limited in their application than the more correlative type approaches, but they may offer greater chemical and biological interpretableness for the chemicals they do predict.

3D-QSAR

Structure-activity methods that consider the 3D structure of modeled compounds in spatial relation to one another are collectively termed 3-dimensional QSAR (3D-QSAR) methods. These methods attempt to identify spatially-localized features across a series of molecules that correlate with activity, and represent requirements for ligand binding and complementarity to a postulated receptor binding site (Green and Marshall, 1995; Marshall and Cramer, III, 1988). These procedures extend the QSAR approach in 3 dimensions by choosing manually (Cramer III *et al.*, 1988) or automatically (Jain *et al.*, 1994), one particular geometry for each modeled compound and using the molecular scaffold (Cramer III, 1988), the pharmacophore (Van Drie *et al.*, 1989), and/or the molecular field (Kearsley and Smith, 1990) method for superimposition.

The underlying assumptions of 3D-QSAR methods are as follows:

- The modeled compound, and not its metabolites or other transformation products, is responsible for the biological effect.
- The proposed or modeled conformation is the bioactive one.
- All compounds are binding in the same way to the same site.
- The biological activity is largely explained by enthalpic processes (steric, electrostatic, hydrogen bonding, etc.).
- Entropic terms are similar for all compounds.
- The system is at equilibrium.
- Common solvent effects—diffusion, transport, etc.—apply to the studied molecules and thus are not considered.

Although enjoying much more extensive use in the area of drug design, the process of 3D-QSAR (specifically as applied in comparative molecular field analysis (CoMFA) will be described here in the context of its limited applications in the area of toxicology prediction. CoMFA is one of the earliest forerunners of current 3D-QSAR techniques, was developed from 1983–1987 (Cramer III and Bunce, 1987), continues to undergo refinement, and remains one of the most widely used 3D-QSAR methods today. In CoMFA, non-covalent ligand-receptor interactions are represented by steric (Lennard-Jones) and electrostatic (Coulombic) interactions with the ligand. The steric and electrostatic interactions of probe atoms with the ligand are calculated at uniform grid points, then tabulated for each molecule (row) in the series. The resulting matrix is analyzed with multi variate statistics (partial least squares or PLS), yielding an equation that relates the CoMFA field value to the activity. This process also highlights those features of the putative receptor that are being probed by the structure-activity data set.

In general, the objective of this and other related 3D-QSAR procedures is to place molecules with common alignments in a 3D grid (or region), calculate interaction values for each grid point, and place the values for each point in a QSAR table. Then create an equation, based on PLS regression, to describe the relationship between the values and the reported activities, verify the predictive ability of the QSAR by cross-validation (and determine the optimal number of components), visualize the final QSAR model by plotting coefficients in the corresponding regions of space, and use the final QSAR equation to estimate the biological activity for other new compounds not included in the model.

Requirements for successful development of a 3D-QSAR model include selecting appropriate compounds and biological data to serve as the training set and identifying a useful and meaningful alignment of the molecules for study. A general guideline is that at least 20 compounds are required to derive a QSAR, although useful QSARs have been obtained with as few as 7 compounds in the model. The quality and choice of biological data to be modeled is critical to successful development of a model. The range and distribution of biological data

are also important, with a normal distribution of data across as wide a range of activities as possible (minimum of 3 log units). The initial challenge is to choose structural conformers as close to the actual bioactive conformers as possible. In the absence of information on the bioactive conformer, default geometry optimization routines are typically employed, which determine a minimum energy conformation. The goal of the alignment procedure is then to superimpose conformers in such as way as to accurately reflect a common ligand-binding orientation to the receptor. Since actual bioactive conformers are seldom known, it has been useful to assume that ligands, regardless of chemical composition, bind in conformations and orientations that present similar steric and electrostatic potential patterns to the target receptor. This is the conceptual basis of a "pharmacophore" (Ariens, 1966), which is defined as the critical 3D arrangement of ligand-functional groups responsible for creating these patterns complementary to the target site(s). The alignment process orients a given molecular conformation in 3D-space relative to all the other molecules in the set. It is extremely important that this be done in a self-consistent manner since differences in field values must reflect structural variation. This may mean, in some cases, using conformations that are not necessarily of lowest energy. Alignment tools (see Klebe *et al.*, 1994 for a discussion) that have been used range from simple methods such as RMS fit, field fit, and Multifit methods to more sophisticated methods such as SEAL and Receptor/DISCO.

After determining the appropriate alignment of molecules for comparison, the 3D-QSAR fields are evaluated over a region usually defined as the "atoms" postulated to comprise a receptor site of known geometry. Steric and electrostatic fields are most often used for such purposes and are computed with a "probe atom" placed at the intersections of a 3D lattice. It is also possible to define a variety of other fields in 3D-QSAR that can reflect such things as partitioning and reactivity properties of molecules, such as HOMO/LUMO fields, polarizability grid fields, and hydrophobic fields.

Several statistical tools (see, for example, Cramer III, *et al.*, 1988) are used to analyze 3D-QSAR parameters to arrive at the final QSAR model and to examine the stability of the derived equation. These tools include cross-validation to examine the internal predictability of the model, cross-validated r^2 (i.e., q^2) to estimate the variance predicted by the model, and bootstrapping to test the stability of QSAR numerical values. An important aspect of the modeling process that aids in evaluation and interpretation is the graphical representation of the 3D-QSAR results. Since each coefficient in a 3D-QSAR equation corresponds to a field type and a 3D coordinate in the region, the 3D-QSAR coefficients can be graphically displayed as scatter or contour plots. The fields may also be color coded according to their level of contribution to the model (e.g., positive or negative), to aid in interpretation of the model and to communicate the nature and role of specific structural prop-
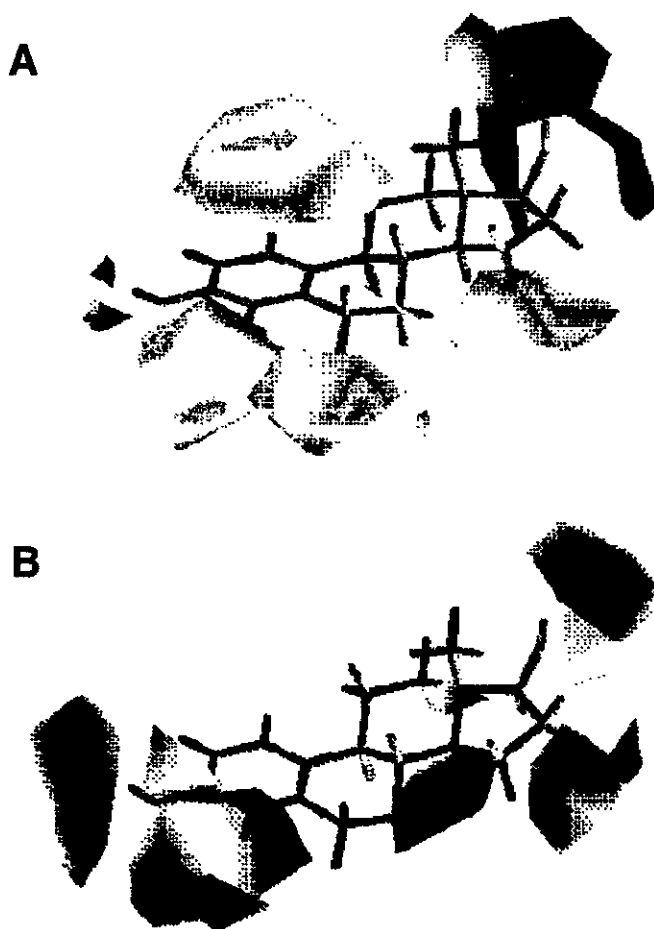

FIG. 3. The template structure is estradiol and the target property for the CoMFA model being displayed is estrogen-receptor binding affinity. (A) Steric fields: negative values represent regions (yellow contours appearing above and below the plane of the phenolic A-ring portion of the molecule) of space where steric bulk should be removed relative to the template structure and positive values (green contours in the vicinity of the D-ring) suggest that steric bulk should be kept or enhanced. (B) Electrostatic fields; red contours appearing on either end of the molecule (off A- and D-rings) designate areas where negative charge appears to be beneficial to binding, whereas blue contours suggest that partial positive charge is desired.

erties in the models. The goal is to ultimately use the final QSAR equation to make predictions, noting the field points that are outside of the model's highlighted graphical regions requiring extrapolation (i.e., novel structural space). Successful 3D-QSAR models in the area of toxicity prediction have primarily centered on endpoints known to be receptor-mediated. Examples include models for estrogen, androgen, and dioxin receptors (Waller *et al.*, 1996b,c; Waller and McKinney, 1995), associated enzyme induction (Waller and McKinney, 1992), and specific P-450 bioactivation activities (Waller *et al.*, 1996a). Graphical representation of the estrogen receptor binding CoMFA model is shown in Figure 3 (A and B), with steric and electrostatic field contour plots (with estradiol used as the template structure for alignment)

models for what are termed non-congeneric chemicals, i.e., large sets of structurally and mechanistically diverse chemicals (for some reviews, see, e.g., Benfenati and Gini, 1997; Benigni and Richard, 1996, 1998). Because SAR ultimately draws its validity from linkage to mechanism, however, any success achieved with these methods rests on the degree to which the global models are able to discern and adequately represent the mechanism-based SAR components of the larger data set (Lewis, 1992; Richard, 1995; Wagner et al., 1995).

Prediction Systems

Two main types of commercial toxicity prediction systems are currently available: the correlative or statistically based programs and the rule-based expert systems (see Benfenati and Gini, 1997; Chapter 6 in Hansch and Leo, 1995; and Richard, 1998a,b). Correlative systems, such as CASE/MultiCASE (Klopman, 1984) and TOPKAT (Enslein, 1993), typically process a large group of non-congeneric chemicals, without user bias or prior organization, and attempt to extract SAR associations from the data by statistical means. The biggest drawback of such systems is the ease with which a prediction is generated versus the need for careful scrutiny of the results. Typically, such methods are better at gross identification of "alerting" classes than at discerning finer activity variations within these classes. Rule-based systems, such as DEREK (Sanderson and Earnshaw, 1991) and ONCOLOGIC (Woo et al., 1995), build associations and generalizations from small groups of chemicals, group similar-acting chemicals into classes based on organic chemistry definitions and limited mechanistic understanding, and use expert judgment and mechanism-based rationale within the classes. The rule-based systems typically are more limited in their application than the more correlative type approaches, but they may offer greater chemical and biological interpretableness for the chemicals they do predict.

3D-QSAR

Structure-activity methods that consider the 3D structure of modeled compounds in spatial relation to one another are collectively termed 3-dimensional QSAR (3D-QSAR) methods. These methods attempt to identify spatially-localized features across a series of molecules that correlate with activity, and represent requirements for ligand binding and complementarity to a postulated receptor binding site (Green and Marshall, 1995; Marshall and Cramer, III, 1988). These procedures extend the QSAR approach in 3 dimensions by choosing manually (Cramer III et al., 1988) or automatically (Jain et al., 1994), one particular geometry for each modeled compound and using the molecular scaffold (Cramer III, 1988), the pharmacophore (Van Drie et al., 1989), and/or the molecular field (Kearsley and Smith, 1990) method for superimposition.

The underlying assumptions of 3D-QSAR methods are as follows:

- The modeled compound, and not its metabolites or other transformation products, is responsible for the biological effect.
- The proposed or modeled conformation is the bioactive one.
- All compounds are binding in the same way to the same site.
- The biological activity is largely explained by enthalpic processes (steric, electrostatic, hydrogen bonding, etc.).
- Entropic terms are similar for all compounds.
- The system is at equilibrium.
- Common solvent effects—diffusion, transport, etc.—apply to the studied molecules and thus are not considered.

Although enjoying much more extensive use in the area of drug design, the process of 3D-QSAR (specifically as applied in comparative molecular field analysis (CoMFA) will be described here in the context of its limited applications in the area of toxicology prediction. CoMFA is one of the earliest forerunners of current 3D-QSAR techniques, was developed from 1983–1987 (Cramer III and Bunce, 1987), continues to undergo refinement, and remains one of the most widely used 3D-QSAR methods today. In CoMFA, non-covalent ligand-receptor interactions are represented by steric (Lennard-Jones) and electrostatic (Coulombic) interactions with the ligand. The steric and electrostatic interactions of probe atoms with the ligand are calculated at uniform grid points, then tabulated for each molecule (row) in the series. The resulting matrix is analyzed with multi variate statistics (partial least squares or PLS), yielding an equation that relates the CoMFA field value to the activity. This process also highlights those features of the putative receptor that are being probed by the structure-activity data set.

In general, the objective of this and other related 3D-QSAR procedures is to place molecules with common alignments in a 3D grid (or region), calculate interaction values for each grid point, and place the values for each point in a QSAR table. Then create an equation, based on PLS regression, to describe the relationship between the values and the reported activities, verify the predictive ability of the QSAR by cross-validation (and determine the optimal number of components), visualize the final QSAR model by plotting coefficients in the corresponding regions of space, and use the final QSAR equation to estimate the biological activity for other new compounds not included in the model.

Requirements for successful development of a 3D-QSAR model include selecting appropriate compounds and biological data to serve as the training set and identifying a useful and meaningful alignment of the molecules for study. A general guideline is that at least 20 compounds are required to derive a QSAR, although useful QSARs have been obtained with as few as 7 compounds in the model. The quality and choice of biological data to be modeled is critical to successful development of a model. The range and distribution of biological data

indicating areas of positive and negative contributions of steric bulk and areas of positive or negative charge.

3D-QSAR has been shown to be useful in the identification of potential toxicants, particularly for activities known to be receptor-mediated or involving specific binding proteins. 3D-QSAR holds additional promise as a means to infer and better understand the specific molecular requirements for the receptor interaction. Limited data and understanding relative to receptor-mediated toxicities and the necessity for rigorous conformational analysis and superposition analyses (alignment problem) are presently hampering large-scale application of 3D-QSAR models as predictive tools. As toxicological understanding at the molecular level progresses and the computational tools are further developed and validated, rapid and accurate predictions of certain toxicological activity based on 3D molecular structure will fulfill more of its current promise.

Application to Metals

The study of the biological activities of organic compounds, which encompass a large proportion of drugs and environmental chemicals, have often applied QSAR methods and approaches. In addition, much of our present day knowledge and understanding of mechanisms by which foreign chemicals affect biological systems is derived from studies with organic compounds. This work has benefited from SAR-based approaches and has led to some fundamental principles that help us to understand and sometimes predict the biological effects of a given organic chemical. Two basic approaches have been particularly helpful in guiding our ability to predict biological activities of organic compounds. These include recognizing structural similarities to compounds known to be important in intermediary metabolism and related life-giving processes (i.e., concepts of lethal synthesis and antimetabolites [Peters, 1963]), discerning specific actions at discrete pharmacological receptors (i.e., the concept of pharmacophores and toxicophores discussed earlier), and anticipating nonspecific effects based on physicochemical properties and reactivities of molecules. A special case of the last is the covalent binding hypothesis in which chemically reactive substances are assumed to react nonenzymatically with cellular macromolecules such as proteins and nucleic acids.

In attempting to extend the above considerations to metal compounds (Hanzlik, 1981), one is faced with a much more limited knowledge about the normal physiological functioning of metals in biological systems and the considerably greater range of chemical properties and reactivities offered by metal compounds of various types. In addition, there has been some success in drawing parallels between the biochemical toxicology of organic and inorganic chemicals based on key chemical properties or processes that may be common to both groups. These include the relationship between bonding and binding, the ability of metals to function as electrophilic species with "alkylating-like" properties, the relative importance of metal

ion size versus charge, and the role of metals as "antimetabolites" in isomorphous interchange processes. The potential of 3D-QSAR methods to study isomorphous interchange processes involving metal ions is of interest. In addition, metal compounds can act as initiators or catalysts *in vivo*, and can be involved in complexing and redox processes in absorption, storage, metabolism, and excretion.

Recent studies (Newman *et al.*, 1998) using metal-ligand binding characteristics to predict metal toxicity and the development of quantitative ion character-activity relationships (QICARs) are showing promise as a screening approach and in situations analogous to those in which QSARs are being applied. Since the major focus in pharmacology and to a large extent in human toxicology has been on organic drugs and poisons, QICARs have not been well developed. In addition, chemical speciation complicates prediction because several metal species usually are present simultaneously and the bioavailability of each is ambiguous. However, some of this ambiguity can be removed by judicious application of the free ion-activity model (FIAM). This model is an extension of the free ion-hypothesis in which the bioactivity of a dissolved metal is correlated with its free ion concentration or activity.

In recent work (Newman *et al.*, 1998), inter-metal trends in toxicity were successfully modeled with ion characteristics reflecting metal binding to ligands associated with a wide range of effects. In general, models for metals with the same valence (i.e., divalent metals) were better than those combining mono-, di-, and trivalent metals. Ion characteristics that were most useful in QICAR model construction included the softness parameter and absolute value of the log of the first hydrolysis constant. The softness index quantifies the ability of the metal ion to accept an electron during interaction with a ligand. It reflects the importance of covalent interactions relative to electrostatic interactions in determining inter-metal trends in bioactivity. Interestingly, softness or molecular polarizability is often an important factor in molecular recognition and binding processes for organic compounds. The hydrolysis constant reflects the tendency for a metal ion to form a stable complex with intermediate ligands such as O donor atoms in biomolecules. There is not a clear counterpart for this on the organic chemical side, and it appears to be a distinctive feature that can be important in determining the relative bioactivity of metals. The first stable reduced state also contributed substantially to several of the 2-variable models. Most models were useful, for predictive purposes, based on an F-ratio criterion and cross-validation, but anomalous predictions did occur if speciation was ignored. The importance of speciation may have confounded attempts to model simple mixtures in complex media. In these cases, quantitative attempts to predict metal interactions in binary mixtures, based on metal-ligand complex stability, were not successful.

There are several resolvable issues that need further attention before the QICAR approach has the same general usefulness as the QSAR approach. These issues include development

and testing of more explanatory variables, careful evaluation of ionic qualities used to calculate explanatory variables, better understanding of models capable of predicting effects for widely differing metals (e.g., metals of different valence states), effective inclusion of chemical speciation, examination of more effects, and assessment of the applicability of QICARs to complex phases such as sediments, soils, and food.

Application of Expert Systems

Allergic contact dermatitis is a cell-mediated immunological response to chemicals that contact and penetrate the skin. It is the most common occupational skin disease and represents a major non-occupational, environmentally related problem. Allergic contact dermatitis is a prominent pathological condition in which understanding of the chemistry has been shown to be the key to understanding the various elements of the toxicity (Ashby *et al.*, 1995; Kimber, 1996; Lepoittevin and Berl, 1996). Chemical reactions and interactions are involved throughout the process, beginning with the crossing of the cutaneous barrier (mainly controlled by the physicochemical properties of the allergen), through the formation of the hapten-protein complex (in which chemical bonds are involved), or during the recognition process between the antigen and the receptors on T lymphocytes (involving the rapidly developing area of supra molecular chemistry).

To cause sensitization, a chemical has to penetrate the skin, where it may be metabolized, and subsequently react with Langerhans cell surface proteins to form new chemical structures that are recognized as foreign. Thus, it might be anticipated that SAR approaches and considerations could be particularly useful in understanding and predicting the relationship between such contact allergic properties of chemicals and their molecular structure. Important chemical factors in contact sensitization include molecular properties affecting bioavailability (appropriate molecular size, polarity, and hydrogen bonding to bring about skin penetration, slow transit, and initiation of binding), chemical stability (sufficient to reach viable tissues of the skin in a reactive form), and protein reactivity (to form stable bonds with proteins either directly or via metabolic activation to, usually, electrophilic species). Reactive chemical species shown to be important include acylating/alkylating/ arylating agents, Michael electrophiles, aldehydes and related carbonyl reagents, free-radical generators, and thiol exchange agents. In view of the previous discussion on the ability of metals to function as electrophilic species with "alkylating-like" properties, it should not be surprising to find that certain metals or metal salts can lead to contact hypersensitivity or dermatitis. This supports the view that metal coordination complexes can be sufficiently stable, and the protein modification sufficiently important, to lead to allergy.

In addition to the nature and reactivity of certain chemical groupings in initiating activity, the compatibility of spatial geometry can also be an important factor contributing to structure-activity relationships, especially in studies of cross-allergy among structurally related families of chemicals. Receptor molecules are typically highly selective with respect to molecule size and shape, and molecules must have similar 3D characteristics to be recognized by true protein bioreceptors. This suggests a possible role for 3D-QSAR approaches in studying the cross-allergic properties of structurally related allergens.

Modeling of contact hypersensitivity is an area where both rule-based and correlative SAR methods have been applied with some success (Ashby *et al.*, 1995; Barratt *et al.*, 1994a,b; Graham *et al.*, 1996; Payne and Walsh, 1994). Skin sensitization databases are available that are searchable by chemical structure, permitting quick identification of structural analogs and easy access to their associated skin sensitization data. This in turn permits one to assess the skin sensitization (predictive testing) potential of chemicals (whole or as substructures) and provides a basis for building QSAR models and using SAR approaches in risk assessment. This can be particularly important since currently no validated, regulatory accepted, *in vitro* methods are available for assessing the skin sensitizing potential of chemicals, although methods that can be useful in fundamental research have been described (Hauser and Katz, 1988). In the absence of *in vitro* methods, rule-based systems like DEREK can serve as a first step in a strategic approach for screening contact allergens and for prioritization of further testing. In addition to classifying chemicals as potential sensitizers or not, more work is needed to derive QSAR models that also have the ability to assess the relative potency of chemical allergens.

DEREK

It has been known for some time that chemical contact allergens are capable of reacting with skin proteins either directly or after appropriate biochemical transformation. The correlation of protein reactivity of chemicals with their skin sensitization potential is well established (Dupuis *et al*, 1982; Lepoittevin *et al*, 1998). At present, it is not possible to predict relative sensitization potency on the basis of physicochemical properties alone. However, one expert rulebase system is available that correlates the structural alerts for protein reactivity of chemicals with their skin sensitization potential. DEREK (an acronym for "deductive estimation of risk from existing knowledge") is a program that embodies both a controlling program and a chemical rulebase (Barratt *et al.*, 1994a,b). In the ideal case, structural alerts used to identify potential sensitizing chemicals need to include those structural features that determine skin penetration and metabolism (both activation and deactivation), chemical reactivity, and immune recognition. However, DEREK, as presently constituted, places heavy emphasis on the chemical reactivity component.

Prior to conducting any preclinical testing on a new ingredient, the chemical can be evaluated for skin sensitization

alerts using DEREK; this expert system makes it is possible to evaluate a large number of chemicals without preclinical testing. Thus, the identification of skin sensitization structural alerts can be extremely helpful in guiding the product development process. It is important to note, however, that some molecules may contain a structural alert, but may not be skin sensitizers, perhaps because their skin permeability is too low or they do not form an immunoreactive moiety within the epidermis. In addition, the fact that the chemical does not trigger a skin sensitization alert in DEREK doesn't guarantee that the chemical is not a sensitizer, since its chemistry may be new to DEREK. In spite of its current limitations, the use of DEREK provides a powerful first step in a strategic approach to the identification of contact allergens.

Structure Database

In addition to using DEREK, a skin sensitization database has been developed that is searchable by chemical structure. The system is designed so that structural analogs and their associated skin sensitization test data can be located in minutes. The skin sensitization data has been gathered from multiple sources. Guinea pig and local lymph-node data on known skin sensitizers have been obtained from the published literature (for example Andersen and Maibach, 1985; Ashby et al., 1993; Cronin and Basketter, 1994). Skin sensitization data have, in addition, been obtained from public databases such as TSCATS and IUCLID. Currently, this skin sensitization database contains approximately 3500 chemicals that are associated with skin sensitization test data. A relational database is used to store the skin sensitization data.

For new ingredients, the structure or structural fragments are used to search the skin sensitization database for structural analogs. Depending on the similarity between the unknown compound and strength of the skin sensitization data associated with the analogs identified in the database, valuable information can be provided to the risk assessment process. Chemical structure searching provides an unambiguous method for the identification of novel compounds as well as structural analogs that, when associated with skin sensitization test data, can be used to predict the skin sensitization potential of the chemical. The use of such structure activity relationships has significantly reduced development times, test costs and animal usage.

CONCLUSION

Given the huge range and variability of possible interactions of chemicals in biological systems, it is highly unlikely that SAR models will ever achieve absolute certainty in predicting a toxicity outcome, particularly in a whole-animal system. However, in different degrees, this caveat applies to any experimental or computational model requiring extrapolation among levels of biological organization (e.g., biochemical to *in vitro* to *in vivo*) or among species. Much more can be done to

improve the scope and utility of SAR approaches by improving the linkages among the various scientific elements of the SAR problem: chemical, biological, and statistical. Certainly, new technologies to refine biofunctional understanding (e.g., DNA arrays to classify chemicals according to gene expression pathways) and better understanding of the mechanistic elements pertinent to an expression of toxicity in whole systems will be useful for refining SAR analyses. In addition, more effective ways are needed to make toxicity databases widely accessible, and bring all relevant information to bear, derived from both expert judgment and quantitative analysis, on the prediction problem.

SAR is an extremely multi-disciplinary field, potentially applicable to a wide range of problems and endpoints. In the environmental and human health area alone, there have been a number of applications for pollution prevention, toxicity screening, and risk assessment (for a review, see Walker, 2000). SAR work has also been useful in guiding mechanistic studies and predicting endocrine-disrupting activities, the environmental fate and ecological effects of chemicals, and environmental-human health interactions (Walker, 2000). Although such broad application potential is desirable and useful, it has also increased the opportunity for the misuse of such methods and approaches. Toxicology is entering a new era of mechanistic emphasis (Stevens and Marnett, 1999). Because SAR ultimately draws its validity from linkage to mechanisms, a mechanism-based approach has been emphasized that slowly builds up a database and scientific understanding of the interaction of chemicals with various forms of life and life-giving processes at the molecular level. In this regard, SAR, in conjunction with the techniques of physical organic chemistry and biochemistry, will further advance our scientific understanding of life-giving processes as well as produce practical benefits to society in terms of improved health outcomes.

REFERENCES

Andersen, K. E., and Maibach, H. I. (1985). Contact allergy predictive tests in guinea pigs: *Current Problems in Dermatology*. Karger, New York.

Ariens, E. J. (1966). Molecular pharmacology, a basis for drug design. *Fortschr Arzneimittelforsch* 10, 429.

Ashby, J., Basketter, D. A., Paton, D., and Kimber, I. (1995). Structure-activity relationships in skin sensitization using the murine lymph node assay. *Toxicology* 103, 177–194.

Ashby, J., Hilton, J., Dearman, R. J., Callander, R. D., and Kimber, I. (1993). Mechanistic relationship among mutagenicity, skin sensitization, and skin carcinogenicity. *Environ. Health Perspect.* 101, 62–67.

Barratt, M. D., Basketter, D. A., Chamberlain, M., Admans, G. D., and Langowski, J. J. (1994a). An expert system rulebase for identifying contact allergens. *Toxic. in Vitro* 8, 1053–1060.

Barratt, M. D., Basketter, D. A., Chamberlain, M., Payne, M. P., Admans, G. D., and Langowski, J. J. (1994b). Development of an expert system rulebase for identifying contact allergens. *Toxic. in Vitro* 8, 837–839.

Benfenati, E., and Gini, G. (1997). Computational predictive programs (expert systems) in toxicology. *Toxicology* 119, 213–225.

Benigni, R., and Richard, A. M. (1996). QSARs of mutagens and carcinogens:

Two case studies illustrating problems in the construction of models for noncongeneric chemicals. *Mutat. Res.* **371**, 29–46.

Benigni, R., and Richard, A. M. (1998). Quantitative structure-based modeling applied to the characterization and prediction of chemical toxicity. *Methods* **14**, 264–276.

Conolly, R. B., Beck, B. D., and Goodman, J. I. (1999). Stimulating research to improve the scientific basis of risk assessment. *Toxicol. Sci.* **49**, 1–4.

Cramer R. D. III, and Bunce, J. D. (1987). The DYLOMMS method: Initial results from a comparative study of approaches to 3D-QSAR. In *QSAR in Drug Design and Toxicology* (D. Hadzi and B. Jerman-Blazic, Eds.), pp. 3–12. Elsevier, Amsterdam.

Cramer R. D., III, Patterson, D. E., and Bunce, J. D. (1988). Comparative Molecular Field Analysis (CoMFA): 1. Effect of shape on binding of steroids to carrier proteins. *J. Am. Chem. Soc.* **110**, 5959.

Cronin, M. T., and Basketter, D. A. (1994). Multivariate QSAR analysis of a skin sensitization database. *SAR QSAR Environ. Res.* **2**, 159–179.

Dupuis, G., and Benezra, C. (1982). *Allergic contact dermatitis to simple chemicals. A molecular approach.* Marcel Dekker, New York, NY.

Enslein, K. (1993). The future of toxicity prediction with QSAR. *In Vitro Toxicology* **6**, 163–169.

Golberg, L. (Ed.) (1983). *Structure-Activity Correlation as a Predictive Tool in Toxicology.* Hemisphere Publishing, New York.

Graham, C., Gealy, R., Macina, O. T., Karol, M. H., and Rosenkranz, H. S. (1996). QSAR for allergic contact dermatitis. *Quant. Struct. Act. Relat.* **15**, 224–229.

Green, S., and Marshall, G. R. (1995). 3D-QSAR: A current perspective. *Trends Pharmacol. Sci.* **16**, 285–291.

Hansch, C. (1969). A quantitative approach to biological structure-activity relationships. *Acct. Chem. Res.* **2**, 232.

Hansch, C., and Dunn, W. J., III (1972). Linear relationships between lipophilic character and biological activity of drugs. *J. Pharm. Sci.* **61**, 1–19.

Hansch, C., Hoekman, D., Leo, A., Zhang, L., and Li, P. (1995). The expanding role of quantitative structure-activity relationships (QSAR) in toxicology. *Toxicol. Lett.*79, 45–53.

Hansch, C., Kim, D., Leo, A., Novellino, E., Silipo, C., and Vittoria, A. (1989). Toward a quantitative comparative toxicology of organic compounds. *Crit. Rev. Toxicol.* **19**, 185–226.

Hansch, C., and Leo, A. (1995). *Exploring QSAR.. Fundamentals and Applications in Chemistry and Biology.* ACS Professional Reference Book, American Chemical Society, Washington, DC.

Hanzlik, R. P. (1981). Toxcity and metabolism of metal compounds: Some structure-activity relationships. In *Environmental Health Chemistry* (J. D. McKinney, Ed.), pp. 467–496. Ann Arbor Science Publishing, Ann Arbor, MI.

Haque, R. (Ed.) (1980). *Dynamics, Exposure, and Hazard Assessment of Toxic Chemicals.* Ann Arbor Science, Ann Arbor, MI.

Hauser, C., and Katz, S. I. (1988). Activation and expansion of hapten- and protein-specific T helper cells from non-sensitized mice. *Proc. Natl. Acad. Sci. U.S.A.* **85**, 5625–5628.

Hermens, J. L. M. (1996). Structure-activity relationships. In *Toxicology: Principles and Applications* (R. J. M. Niesink, J. deVries, and M. A. Hollinger, Eds.), pp. 239–268. CRC Press, New York.

Hermens, J. L. M., and Opperhuizen, A. (Eds.) (1991). *QSAR in Environmental Toxicology IV.* Elsevier, Amsterdam.

Jain, A. N., Koile, K., and Chapman, D. (1994). Compass: Predicting biological activities from molecular surface properties. Performance comparisons on a steroid benchmark. *J. Med. Chem.* **37**, 2315–2327.

Kaiser, K. L. E., Ed. (1987). *QSAR in Environmental Toxicology II,* Reidel, Dordrecht, Holland.

Karcher, W., and DeVillers, J., Eds. (1990). *Practical Applications of Quan-*

titative Structure-Activity Relationships (QSAR) in Environmental Chemistry and Toxicology. Kluwer, Dordrecht and Boston.

Kearsley, S., and Smith, G. (1990). An alternative method for the alignment of molecular structures: Maximizing electrostatic and steric overlap. *Tetrahedron Comput. Methodol.* **3**, 615.

Kimber, I. (1996). Altenative methods for contact sensitization testing. In *Toxicology of Contact Dermatitis* (K. Kimber and T. Mauer, Eds.), pp. 140–151. Taylor and Francis, London.

Klebe, G., Mietzner, T., and Weber, F. (1994). Different approaches toward an automatic structural alignment of drug molecules: Applications to sterol mimics, thrombin and thermolysin inhibitors. *J. Comput. Aided Mol. Des.* **8**, 751–758.

Klopman, G. (1984). Artificial intelligence approach to structure-activity studies. Computer automated structure evaluation of biological activity of organic molecules. *J. Amer. Chem. Soc.* **106**, 7315–7320.

Lepoittevin, J.-P., Basketter, D. A., Goossens, A., and Karlber, A-T. (1998). *Allergic Contact Dermatitis: The Molecular Basis.* Springer, Berlin.

Lepoittevin, J.-P., and Berl, V. (1996). Molecular basis of allergic contact dermatitis. In *Dermatotoxicology,* 5th ed. (Francis N. Marzulli and Howard I. Maibach, Eds.), pp. 147–160. Taylor and Francis, Washington, D.C.

Lewis, D. F. V. (1992). Computer-assisted methods in the evaluation of chemical toxicity. In *Reviews in Computational Chemistry* (K. B. Lipkowitz and D. B. Boyd, Eds.), pp. 173–221. VCH Publishers, New York.

Lipnick, R. L. (1991). Outliers: their origin and use in the classification of molecular mechanisms of toxicity. *Sci. Total Environ.* **109/110**, 131–153.

Marshall, G. R., and Cramer, R. D. III (1988). Three dimensional structure-activity relationships. *Trends Pharmacol. Sci.* **9**, 285–289.

Matthews, E. J., and Contrera, J. F. (1998). A new highly specific method for predicting the carcinogenic potential of pharmaceuticals in rodents using enhanced MCASE QSAR-ES software. *Regul. Toxicol. Pharmacol.* **28**, 242–264.

McKinney, J. D. (1985). Monograph on structure-activity correlations in mechanism studies and predictive toxicology. *Environ. Health Perspect.* **61**, 349.

McKinney, J. D. (1996). Reactivity parameters in structure-activity, relationship-based risk assessment of chemicals. *Environ. Health Perspect.* **104**, 810–816.

McKinney, J. D. and Singh, P. (1981). Structure-activity relationships in halogenated biphenyls: Unifying hypothesis for structural specificity. *Chem. Biol. Interact.* **33**, 271–283.

Newman, M. C., McCloskey, J. T., and Tatara, C.P. (1998). Using metal-ligand binding characteristics to predict metal toxicity: Quantitative ion character-activity relationships (QICARs). *Environ. Health Perspect.* **106**(Suppl. 6), 1419–1425.

Payne, M. P., and Walsh, P. T. (1994). Structure-activity relationships for skin sensitization potential: Development of structural alerts for use in knowledge-based toxicity prediction systems. *J. Chem. Inf. Comput. Sci.* **34**, 154–161.

Peters, R. A. (1963). *Biochemical Lesions and Lethal Synthesis.* Pergamon, Oxford.

Rand, G. M., and Petrocelli, S. R., Eds. (1985). *Fundamentals of Aquatic Toxicology.* Hemisphere, New York.

Richard, A. M. (1995). Role of computational chemistry in support of hazard identification (ID): mechanism-based SARs. *Toxicol. Lett.* **79**, 115–122.

Richard, A. M. (1998a). Structure-based methods for predicting mutagenicity and carcinogenicity: are we there yet? *Mutat. Res.* **400**, 493–507.

Richard, A. M. (1998b). Commercial toxicology prediction systems: A regulatory perspective. *Toxicol. Lett.* **102–103**, 611–616.

Richard, A. M., and Hunter, E.S., III (1996). Quantitative structure-activity

relationships for the developmental toxicity of haloacetic acids in mammalian whole embryo culture. *Teratology* **53**, 352–360.

Safe, S. (1990). Polychlorinated biphenyls (PCBs), dibenzo-*p*-dioxins (PCDDs), dibenzofurans (PCDFs), and related compounds: Environmental and mechanistic considerations which support the development of toxic equivalency factors (TEFs). *Crit. Rev. Toxicol.* **21**, 51–88.

Safe, S. (1994). Polychlorinated biphenyls (PCBs): Environmental impact, biochemical and toxic responses, and implications for risk assessment. *Crit. Rev. Toxicol.* **24**, 87–149.

Sanderson, D. M., and Earnshaw, C. G. (1991). Computer predicition of possible toxic action from chemical structure; the DEREK system. *Hum. Exp. Toxicol.* **10**, 261–273.

Stevens, J. L., and Marnett, L. J. (1999). Defining molecular toxicology: A perspective (editorial). *Chem. Res. Toxicol.* **12**, 747–748.

Topliss, J. G., and Edwards, R. P. (1979). Chance factors in studies of quantitative structure-activity relationships. *J. Med. Chem.* **22**, 1238–1244.

Van den Berg, M., Birnbaum, L., Bosveld, A. T. C., Brunstrom, B., Cook, P., Feeley, M., Giesy, J., Hanberg, A., Hasegawa, R., Kennedy, S. W., Kubiak, T., Larsen, J. C., van Leeuwen, F. X., Liem, A. K., Nolt, C., Peterson, R. E., Poellinger, L., Safe, S., Schrenk, D., Tillitt, D., Tysklind, M., Younes, M., Waern, F., and Zacharewski, T. (1998) Toxic equivalency factors (TEFs) for PCBs, PCDDs, PCDFs for humans and wildlife. *Environ. Health Perspect.* **106**, 775–792.

Van Drie, J. H., Weininger, D., and Martin, Y. C. (1989). ALADDIN: An integrated tool for computer-assisted molecular design and pharmacophore recognition from geometric, steric, and substructure searching of 3-dimensional molecular structures. *J. Comput. Aided Mol. Des.* **3**, 225–251.

Wagner, P. M., Nabholz, J. V., and Kent, R. J. (1995). The new chemicals process at the Environmental Protection Agency (EPA): Structure-activity

relationships for hazard identification and risk assessment. *Toxicol. Lett.* **79**, 67–73.

Walker, J. D., Ed. (2000). Handbooks on QSARs for (a) *Pollution Prevention, Toxicity Screening, Risk Assessment, and WWW Application;* (b) *Predicting Endocrine Disruption Potential of Chemicals;* (c) *Predicting Ecological Effects of Chemical;* (d) *Environmental Fate of Chemicals; and* (e) *Predicting Effects of Chemicals on Environmental-Human Health Interactions.* SETAC Press, Pensacola, FL.

Waller, C. L., Evans, M. V., and McKinney, J. D. (1996a). Modeling the cytochrome p450-mediated metabolism of chlorinated volatile organic compounds. *Drug Metab. Disp.* **24**, 203–210.

Waller, C. L., Juma, B. W., Gray, L. E., Jr., and Kelce, W. (1996b). Three-dimensional quantitative structure-activity relationships for androgen receptor ligands. *Toxicol. Appl. Pharmacol.* **137**, 219–227.

Waller, C. L., and McKinney, J. D. (1992). Comparative molecular field analysis of polyhalogenated dibenzo-*p*-dioxins, dibenzofurans, and biphenyls. *J. Med. Chem.* **35**, 3660–3666.

Waller, C. L., and McKinney, J. D. (1995). Three-dimensional quantitative structure-activity relationships of dioxins and dioxin-like compounds: Model validation and Ah receptor characterization. *Chem. Res. Toxicol.* **8**, 847–858.

Waller, C. L., Oprea, T. I., Chae, K., Park, H. K., Korach, K. S., Laws, S. C., Wiese, T. E., Kelce, W. R., and Gray, L. E., Jr. (1996c). Ligand-based identification of environmental estrogens. *Chem. Res. Toxicol.* **9**, 1240–1248.

Woo, Y.-T., Lai, D., Argus, M., and Arcos, J. (1995). Development of structure-activity relationship rules for predicting carcinogenic potential of chemicals. *Toxicol. Lett.* **79**, 219–228.