12

SAR and Modeling

Ann M. Richard, Pauline Wagner, Richard Purdy, and Gilman Veith

ABSTRACT: SAR plays a prominent role in TSCA screening of new chemicals and existing chemicals in commerce. SAR models for bioavailability, ecotoxicology and health toxicology endpoints are being used to identlify chemicals with the greatest potential for ecological or health hazard, to set testing priorities, and to provide scientific support for a testing recommendation. SAR models vary considerably in accuracy and utility for screening application depending upon the quality of available data and level of current knowledge for a toxicity endpoint. The main deficiency of current SAR capabilities is inadequate data, and lack of knowledge of mechanisms of toxicity for many chemical classes and toxicity endpoints of potential regulatory concern. Knowledge or inference of a common mechanism of toxic action is crucial for selecting appropriate chemical analogues, guiding SAR model development, establishing model plausibility, and providing the necessary scientific rationale for model acceptance and use in prediction. This paper discusses the present role and capabilities of SAR in TSCA screening, general features and limitations of SAR, current and evolving SAR technologies, and advances most likely to lead to improvements in SAR models for use in TSCA screening. Although SAR has the clear potential to further reduce the need for testing or eliminate testing in some circumstances, the promise of SAR will not be fulfilled without proper application of these methods. This entails clear recognition of the limitations of SAR, and appreciation for the essential roles of research into mechanisms of toxicity, and strategic testing for further SAR model development and refinement.

A structure-activity relationship (SAR) relates features of chemical structure to a property, effect, or biological activity associated with that chemical. The fundamental premise is that the structure of a chemical determines its physical properties and activities. The term "structure-activity relationship" has taken on a wide range of meaning over the years, from heuristic chemical associations and human expert approaches that consider primarily structural features, to formal mathematical relationships that relate specific chemical attributes to a quantitative measure of the property or activity of interest, the latter being commonly referred to as "quantitative structure-activity relationships" (QSARs). In both the pharmaceutical and chemical industries, structure-activity considerations have long been used to design chemicals with commercially desirable properties. In the environmental protection field, SAR is being used to predict adverse ecological and health effects, with applications ranging from the prediction of relevant properties, such as chemical stability, bioavailability and bioaccumulation, to the prediction of various forms of chemical toxicity.

The focus of this workshop is testing and screening strategies for review of the Toxic Substances Control Act (TSCA) inventory of existing chemicals in commerce. This problem poses a significant and immediate challenge, not only in terms of the sheer numbers of chemicals that have undergone little testing or review (>10,000), but also in terms of the multiple exposure routes and ecological and health endpoints of potential concern. The foremost goal is to identify the chemicals that pose the greatest potential ecological and health risks, and to strategically allocate limited testing resources to best characterize these risks. SAR, coupled with exposure and use estimates, represents the top tier in a multiple tier screening approach for assessing chemical hazard, and provides the primary means for setting testing priorities. SAR currently plays a prominent

This manuscript has been reviewed by the U. S. Environmental Protection Agency and approved for publication. Approval does not signify that the contents necessarily reflect the views and policies of the Agency, nor does mention of trade names or commercial products constitute endorsement or recommendation for use.

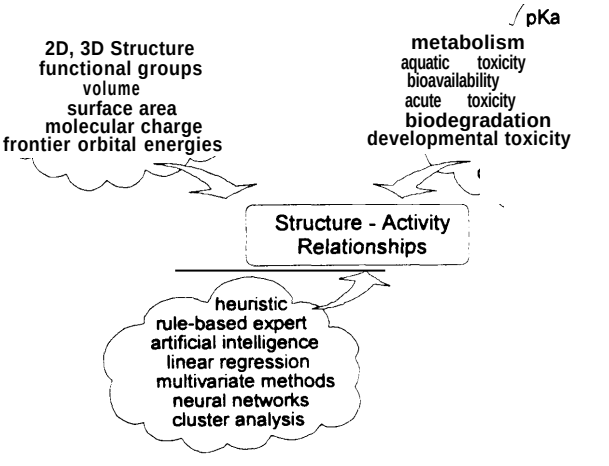
role in testing and screening strategies for TSCA review of new chemicals and existing chemicals in commerce. SAR screening is being used for hazard identification, to set testing priorities, to provide scientific support for a testing recommendation, and in a relatively new U.S. Environmental Protection Agency (EPA) initiative, to aid in the design of safer chemicals by suggesting modifications in structure predicted to minimize toxicity. SAR has the potential to further reduce the need for property measurements and animal testing, generate insight into mechanisms of action, and achieve better environmental protection by providing for more efficient screening of the TSCA inventory for a wide range of toxicity endpoints.

The following will consider some general characteristics of SAR, principles of application to toxicity screening, limitations and guidelines for use of current SAR technologies, current SAR capabilities being applied to the TSCA screening problem, and new technologies and advances that will lead to improvements in SAR capabilities for toxicity screening.

SAR FOR TOXICITY SCREENING: GENERAL CONSIDERATIONS

SAR approaches are extremely general with respect to the possible representations of chemical structure, the types of chemical or biological activity that can be modeled, and the methods for relating the two (figure 12-1). In contrast, an SAR model is highly specific to the particular set of chemicals, attributes, and experimental activities used in its derivation. An SAR model codifies and rationalizes existing data. It follows that the range of application, predictive accuracy, and ultimate relevance of the SAR model is wholly determined by the quality and quantity of existing data and knowledge upon which the SAR model is based. For example, an SAR model developed to predict mutagenicity based on qualitative (+/- activities) experimental data for a series of aromatic amines tested in the TA100 Salmonella reversion assay is not likely to be applicable to other chemical classes (e.g. small haloorganics), other strains of Salmonella (e.g. TA90), or the prediction of quantitative potencies. An obvious corollary is that an SAR model is only as relevant to the ultimate health effect of concern (e.g. carcinogenicity) as the toxicity endpoint that it purports to model (e.g. mutagenicity).

Figure 12-1: Structure-Activity Approaches



There are two complimentary goals of an SAR study: 1) to predict the activities of untested chemicals and prioritize chemicals according to relative activities; and 2) to provide a rational scientific basis for understanding and interpreting existing biological/chemical activity data in terms of chemical structure and mechanism of action. The former is the primary allure of SAR, while the latter provides the foundation and prescription for its success.

SAR methods are optimally applied to the prediction of intrinsic physical properties of a chemical where a single "mechanism" is operative and the property can be considered independent of external interactions. A number of high quality QSAR and computational chemistry models have been developed and are used routinely by industry and regulatory agencies to predict chemical properties such as vapor pressures, melting points, acid dissociation constants, spectral properties, chromatographic retention times, and octanol/water partition coefficients (commonly referred to as "logP"), to name but a few.

Such models have the advantages of significant cost savings over laboratory determinations, speed, ease of use, and no need for the availability or handling of the chemical of concern. Modeled properties also serve as key parameters for use in the development of SAR models for biological fate or effect. For example, an octanol/water partition coefficient, a property used extensively in QSAR studies of biological activity, approximates the ability of a chemical to transport through biological membranes and can be modeled easily and accurately by the computerized CLOGP method (13), yielding cost savings from $10,000 to $30,000 per chemical.

Toxicology provides a more severe SAR modeling challenge. In this case, the extrinsic chemical "property" being modeled is a biological endpoint, i.e. an activity determined by the complex interaction of a chemical within the biological system. Whereas an intrinsic chemical property relates unambiguously to a single physical process or mode-of-action, there are most often many possible mechanisms by which chemicals with different structural characteristics elicit a common biological activity or toxicity endpoint. This complexity coupled with lack of knowledge concerning mechanisms of toxicity introduces greater uncertainty and imposes greater restrictions on the application of SAR concepts to toxicology. The key to ameliorating these concerns is to restrict SAR models, whenever possible, to chemicals that elicit their effect by a common mode-of-action, and to incorporate whatever knowledge is available concerning the mechanism of toxicity into SAR model development. This does not necessarily require full, detailed knowledge of the molecular mechanism, but a common mode-of-action must be indirectly inferred or hypothesized to maximize validity and reliability, and minimize uncertainty in the SAR model. (The terms "mechanism" and "mode-of-action" are used interchangeably in the present text). This explicit linkage between SAR and mechanism of action is crucial to establishing the plausibility of an SAR model and providing the necessary scientific rationale for its acceptance and use (6).

It follows that SAR models will be most successful when applied to mechanistically well-defined toxicity endpoints. Such endpoints are more likely to consist of specific biochemical indicators (e.g. P-450 induction, inhibition of DNA repair), *in vitro* bacterial assays, tissue and organ-specific effects, and *in vivo* assays where a common unifying process, transformation, or event is central to the activity. Examples include the central role of: logP or bioavailability in narcosis mechanisms of acute aquatic toxicity; formation and stability of electrophilic nitrenium ion intermediates in mutagenicity of nitroaromatics; Ah receptor binding in toxic effects of dioxin and PCBs; and dermal penetration (logP) and acidity (pKa) in determining skin corrosivity. The most difficult types of toxicity endpoints to model with SAR are termed "apical" endpoints, i.e. typically whole animal *in vivo* assays of chronic disease or effect that consider as much of the integrated physical and biological process as possible in a single test (e.g. developmental toxicity, neurotoxicity behavioral effects, rodent carcinogenicity). While these assays are often considered most relevant and useful to human health or ecological risk assessment, they are also the most costly, most controversial in terms of animal usage, least likely to be available, and most difficult to interpret mechanistically. In these cases, restriction of the SAR to a narrowly defined chemical class is the best assurance that a common mode-of-action applies.

Another essential element of an SAR model is the data used in its development, i.e. the chemicals and activities or potencies. There are two fundamentally distinct types of SAR models for any toxicity endpoint, those that model the conditions for distinguishing between activity classes, e.g. "actives" and "inactive", and those that model the conditions for modulating potencies among a group of chemicals belonging to a common activity class, i.e. "actives" (5). The SAR requirements for being a member of the "active" class may be quite distinct from those that explain differences in potency among the actives. In addition, data requirements differ for the two objectives: sufficient test data on

negative, or inactive analogues are essential for establishing boundaries of an SAR, while test data on positives, or actives covering a wide range of potencies are required for QSAR development. Often, sufficient and appropriate negative test data for SAR model development are lacking. Testing recommendations are driven by concern for adverse effects and, hence, testing resources are applied to the chemicals considered to pose the greatest hazard, i.e. most likely to be active in a toxicity assay. In addition, negative test data are less likely to be published and available since they are perceived to be of less interest to the scientific community. A legitimate role of strategic testing, particularly in a research setting, should be to challenge and improve the quality of current SAR models and, in some cases, verification of a negative test prediction may be the best use of available resources. Often the most dramatic structure-activity differences among analogues, e.g. where a minor structural change eliminates or imparts an activity, are the most informative and useful in SAR analysis. For example, addition of a single methyl group in the bay region of the PAH, benz(a)anthracene, eliminates its carcinogenic activity due to steric crowding and blocking of metabolic activation to the ultimate carcinogen, i.e. the diol epoxide.

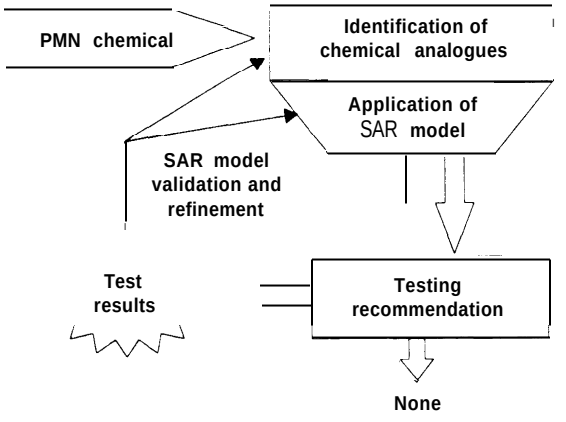
SAR has been most successfully applied to classes of organic chemicals where quantitative, reproducible activity data are available for pure chemicals with known structures. Since SAR requires knowledge of individual chemical structures, it cannot be applied to uncharacterized chemical mixtures. When SAR is applied to polymers, it generally deals with reactivity characteristics of the monomeric units. Also, very little SAR modeling has been done for inorganic chemicals, i.e. metals or metal complexes, due to sparsity of data on chemical analogues, and the greater challenges in characterizing and calculating the pertinent chemical characteristics of these species in biological systems. Subject to these constraints, conditions for optimal SAR model development and application include: restriction of the SAR model to a well-defined toxicity endpoint or single mode-of-action chemical class; availability of test data for a range of chemical structures, attributes, and potencies; use of mechanistically relevant molecular descriptors; a mechanism-based scientific rationale for SAR model and predictions; and prospective validation on test chemicals not used in SAR model derivation. Optimally, some knowledge of a possible or probable biological mechanism of action guides parameter development, provides the basis for determining chemical analogy, defines the region of chemical/activity space where the SAR model is likely to be applicable, i.e. places limits on model extrapolation, and provides scientific rationale for a model prediction (17). Even in the absence of explicit knowledge of mechanisms of toxicity, however, an SAR model developed mindful of the above constraints has the potential to generate insight into possible mechanisms of toxicity and guide further experimentation.

Hence, there is a continuum of SAR modeling tools, biological endpoints, and considerations that impact on the relevance and utility of SAR models for use in toxicity screening. The next section will consider current SAR capabilities being applied to TSCA screening.

USE OF SAR IN TSCA REVIEW OF NEW AND EXISTING CHEMICALS

The bulk of the SAR expertise within EPA currently being brought to bear on the TSCA existing chemicals problem has evolved out of the Pre-Manufacture Notification (PMN) review process (2). Hence, the PMN process, and its strengths and weaknesses warrants some discussion. By law, TSCA requires companies wishing to manufacture a chemical not on the TSCA Inventory to submit a premanufacturing notice (PMN) to EPA. EPA then has 90 days in which to determine if the manufacture, processing or use of that chemical in commerce may present an unreasonable risk to human health or the environment. If this is determined, EPA has the legal authority to request further test data be submitted for the PMN chemical. The Structure-Activity Team (SAT) within the Office of Pollution Pre

Figure 12-2: EPA Screening Procedure



vention and Toxics at EPA was conceived in order to efficiently and systematically screen PMN chemicals for health and/or ecological hazard. The 20 or so members of the SAT represent a wide range of chemical, ecological, and toxicological disciplines. Some characteristics of the PMN process are as follows. Since no toxicity testing is required by law, test data accompany fewer than 5% of the PMN chemicals submitted. Hence, SAR frequently provides the sole means for evaluating these chemicals. In the area of ecotoxicology, a number of computerized, chemical class-based QSAR models have been developed for use in predicting physical/chemical properties, ability to degrade and bioconcentrate, and toxicity to fish, aquatic invertebrates, and algae (8, 10, 23). In the health toxicology area, models and SAR expertise vary considerably depending on the state of knowledge in the particular field of toxicology and, in contrast to ecotoxicology, most models are qualitative and heavily reliant on chemical analogy, rules and expert judgement.

The mandate of the SAT is primarily operational, i.e. to evaluate more than 2000 chemicals/yr within a 90-day deadline from the date of each PMN submission (20). The SAT operates under strict confidential business information (CBI) restrictions with respect to the chemicals it evaluates, which prohibits the sharing of chemical structures used in SAR model development with outside parties. Computers are used for data base searching, to aid the identification and retrieval of chemical analogues, for the calculation of chemical properties required for estimating bioavailability and fate, and for the application of QSAR models for eco-tox endpoints. Finally, there is an emphasis on mechanism-based approaches and interpretations, whenever possible, to reduce uncertainty in the SAR prediction, increase plausibility, and provide the necessary scientific rationale to support a testing recommendation.

The PMN screening process is summarized in figure 12-2. Upon receiving a PMN chemical submission, the SAT reviews the literature and in-house data bases of previously reviewed chemicals to identify possible analogues. Analogues consist of chemicals with similar structures or fictional groups to the PMN chemical, for which test data are available or a previous SAT assessment is on record. In some cases, when very little is known about the chemical or a toxicity endpoint, this is the extent of the SAR, i.e. available data for the analogue are assumed to apply to the PMN chemical. This information could flag the PMN chemical as a potential fish toxicant, developmental toxicant, carcinogen, etc. In other cases, additional SAR considerations or models apply to the chemical class, of which the PMN chemical is assumed to be a member, and are used in making a testing recommendation. This SAR hazard assessment is considered along with exposure data in making the “may present an unreasonable risk” determination, which may trigger a testing requirement under Section 5 of TSCA. The importance of the analogue selection step as the top-most tier in this overall process should be stressed. If suitable analogues are unavailable, or if inappropriate analogues are chosen with respect to toxic mode-of-action, inappropriate SAR considerations and incorrect judgement could be applied to the PMN chemical under review. Finally, an extremely important element of the PMN process occurs subsequent to the issuance of a testing requirement. Comparison of the toxicity test result to the SAT prediction provides the primary mechanism for the continual validation and refinement of the SAR models and assumptions used in the PMN process.

By virtue of their legal mandate, the EPA/SAT has evolved into a unique and valuable resource that plays a pivotal role in environmental and health protection. Nowhere else in the world has such a concentration and wide range of expertise been focused solely on the task of SAR model development and ecological fate and toxicity screening. The SAT is also unique in terms of its unparalleled access to unpublished, proprietary, and internally generated toxicity test data from a wide range of sources and for a wide variety of chemicals, data which are essential for the development and refinement of predictive SAR models. More than 24,000 chemicals have been screened through the PMN process since 1979, contributing greatly to the evolution and improvement of SAR expertise and models in current use. In addition, the SAT has engaged in a number of outside collaborations to further verify and improve upon existing models. An example is a recent collaboration with their European Community (EC) counterparts in which the SAT blindly evaluated 144 chemicals concurrently undergoing toxicity testing (19). This and other exercises have provided support for many of the SAR models and assumptions in current use by the EPA/SAT, while pointing to deficiencies in others. The overall performance results of the European Community exercise are available in the form of a joint EPA/EC summary report from either the European Union in Brussels, or the EPA as Document Number EPA 743-R-94-001. However, confidentiality restrictions required that all of the individual chemical identities associated with the exercise be destroyed, a loss that limits the potential benefit of the study to both the EPA and the outside SAR community.

The current expertise of the EPA/SAT is being used in the development of a "Use Cluster Scoring System" for evaluation of the TSCA existing chemicals inventory. A tiered strategy has been implemented, the first step involving the identification of "use clusters", i.e. categories of common use chemicals, such as paints, rust inhibitors, plasticizers, etc., that are likely to have similar exposure scenarios (e.g. paints - occupational inhalation exposure). The second stage involves prioritization of these use clusters based on SAR, exposure, and available toxicity data. A more complete SAR toxicity evaluation using the models and expertise of the EPA/SAT, is then applied within the use cluster to establish testing priorities among the individual chemicals. All available test data, which may include possible occupational exposure health data for existing chemicals in commerce, are considered in the preliminary toxicity screening assessment.

The main deficiencies of the EPA/SAT approach for TSCA review of PMN chemicals or existing chemicals in commerce, shared by the SAR community at large, are inadequate data and lack of knowledge of mechanisms of toxicity for many of the chemical classes and toxicity endpoints of potential regulatory concern. In addition, the SAT has neither the mandate, nor the time or resources to evaluate new technologies or to carry out research to improve existing SAR models. Hence, a deliberate outreach effort must be made by the SAT to communicate and interact with industry and research groups with the potential to impact on the process. While the final SAR models and expertise developed by the EPA/SAT can and are being made available to the public (two examples being the ECOSAR program for eco-tox screening, and the ONCOLOGIC expert system for predicting chemical carcinogenicity), CBI confidentiality constraints prohibit the release of the primary data and chemical structures used in model development. CBI constraints are designed to safeguard the rights of industry, yet are in perpetuity under existing law, regardless of whether the chemical was ever produced or entered into commerce. These constraints limit the ability of outside parties to independently scrutinize, validate, and improve upon EPA/SAT models. Access to the data used in SAR development is valuable for defining the proper boundaries of application of the SAR model, for developing hypotheses concerning the structural basis for the toxic mode-of-action, and for refining or developing alternative SAR models.

QSAR/SAR models are also developed and used by industry for addressing TSCA require-

ments. However, these models tend to be tailored and restricted to the specialty chemicals produced and used within a particular industry (e.g. solvents, adhesives, etc.). These SAR models, and data on which they are based, are usually considered proprietary. There tends to be more limited SAR expertise within industry with respect to the wide range of toxicity endpoints of potential concern under TSCA, and more limited access to data than is available to the EPA/SAT. Particularly when in-house expertise is lacking, there is incentive for industry to take advantage of the PMN process for toxicity screening prior to large dollar investments in research and/or development. A PMN submission costs little and is performed within a short time frame. Even when expertise is available to industry, there is incentive to anticipate PMN toxicity estimates that would trigger a testing requirement, rather than to develop independent estimates. An EPA testing requirement for a PMN chemical often provides sufficient incentive for industry to redirect a line of research or abandon plans to manufacture a potentially toxic chemical, providing an effective means for serving the interests of environmental protection. For the review of existing chemicals in commerce, the economic incentive to avoid possible regulatory action is much greater since considerable investment in the chemical has already taken place. In this case, industry is more inclined to challenge testing requirements by independent SAR estimates.

CURRENT SAR APPROACHES AND EMERGING TECHNOLOGIES

The most widely used paradigm for QSAR study is the linear free-energy relationship (LFER) or Hammett equation approach, based on statistical linear regression fit of steric, electronic, and hydrophobicity terms to biological potency. This is a chemical class-based approach, designed to be applied to a range of structurally similar, or "congeneric" chemicals that are assumed to have a common mechanism of action. LFER equations are the basis of the ECOSAR compilation of QSARs for ecotoxicology used by the SAT, and the majority of published QSARs for biological endpoints. Within the QSAR paradigm, further incremental advances will come from better mechanism-based chemical classifications, generation of additional test data, and development of more mechanistically relevant molecular parameters and descriptors. For example, while it is now possible to predict with reasonable accuracy the acute toxicity of "unreactive" (i.e. narcosis mechanism) chemicals to many aquatic species, models are less reliable for chemicals acting by alternate mechanisms of action. In particular, there is movement in the QSAR field towards incorporation of more rigorous quantum mechanical properties related to potential reactivity and energy characteristics of molecules derived from their three-dimensional structure.

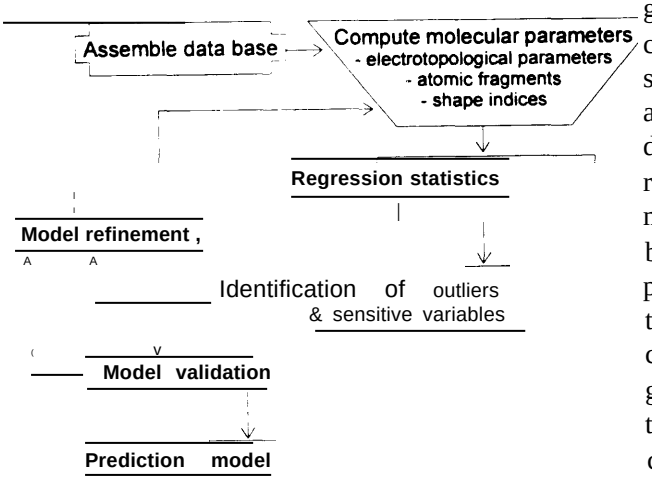
Each individual LFER QSAR equation is associated with a relatively narrow range of chemicals and a specific biological endpoint and, thus, has limited applicability to other toxicity prediction problems. An approach being advocated by Corwin Hansch, one of the pioneers of the QSAR field, is to process larger units of existing information than individual QSARs in order to generate insight into unifying features of biological processes (12). Over the past several years, Hansch and coworkers have compiled over 3000 existing QSARs from the literature into a computerized data base, CQSAR (9), for easy access, comparison, and study. The CQSAR data base is also being used as a validation tool, to judge individual QSARs in a larger biological context by lateral examination of related, or overlapping QSARs, i.e. QSARs for similar chemicals/different endpoints, QSARs for similar endpoints/different chemicals, or QSARs for different chemicals/different endpoints having a similar functional form. For example, a very general feature of the CQSAR data base is that >85% of the QSARs contain a major contribution from a hydrophobicity term (logPo/w), and the coefficient of this term is almost always in the range of 1-2. This argues that QSARs without a logP term, or with a logP coefficient significantly deviating from 1-2 should be considered either novel or suspect.

Table 12-1: TOPKAT Available Models

- Rat acute oral LD50 (19 submodels, 400 chemicals)
- Rat chronic oral LOAEL
- Mouse inhalation LC50
- Developmental toxicity potential (3 submodels, w/ and w/o maternal toxicity)
- Carcinogenicity
- Mutagenicity
- Fathead minnow acute LD50
- Daphnia EC50
- Biodegradability
- Skin/eye irritancy (Draize test)

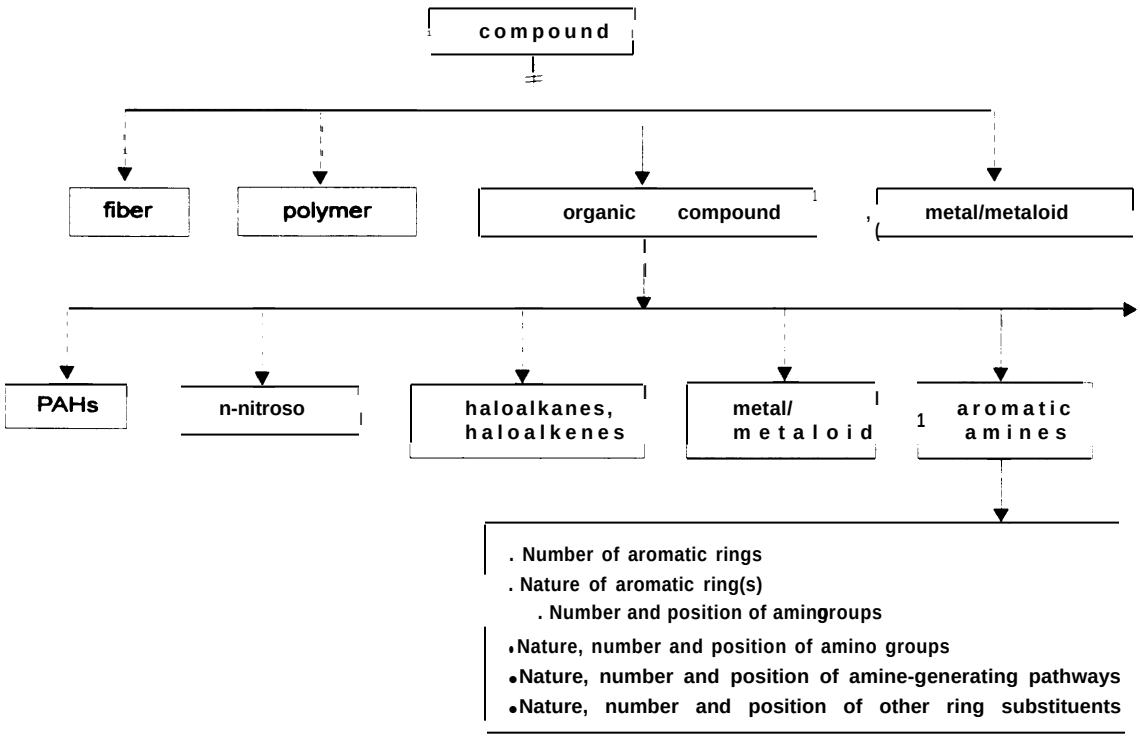
An example of a commercially available SAR program, in use by some industry and government groups, is the TOPKAT computer-based toxicity prediction program (18). TOPKAT is based on LFER concepts, but has typically been applied to SAR modeling of large data sets of "non-congeneric" chemicals, i.e. chemicals representing many chemical classes and mechanisms of action. TOPKAT applies traditional statistical approaches, such as multiple-linear regression and discriminant analysis, to identify SAR associations between structure-derived chemical properties and activity. Indicator variables, i.e. variables that take on a value of O or 1 depending on the presence or absence of a molecular feature, provide an approximate means for incorporating multiple chemical classes into a common SAR model, e.g. a parameter could be "turned on" or "turned off' if a molecular feature such as a nitro group were present. TOPKAT models have been developed and are available for the endpoints listed in table 12-1. The model development procedure is summarized in figure 12-3 and culminates in an SAR model for toxicity prediction. The main limitations of the TOPKAT approach, shared by other statistically-based computerized SAR programs such as CASE and ADAPT, are: limitations in chemical descriptors; dissociation of SAR model development from biological mechanism considerations; and the abandonment of the chemical class restrictions of traditional QSAR (16). Models have been developed for large, chemically diverse data sets associated with complex toxicity endpoints known to represent many possible modes-of-action and, since little effort has been made to incorporate mechanism considerations, models tend to be difficult to interpret and scientifically rationalize. In addition, there has been a tendency towards over-reliance on statistical indicators of model predictive capabilities and underestimation of the inherent uncertainty of these models due to their biological component. For these reasons, TOPKAT, and other statistically-based toxicity prediction programs are not currently used by the SAT for TSCA screening.

Figure 12-3: TOPKAT Method



TOPKAT does, however, have some useful features and legitimate uses. One of TOPKAT's greatest strengths is its high quality data bases compiled from private sources and an exhaustive search of the literature, where each experiment and activity call is carefully evaluated prior to data base incorporation. TOPKAT provides ready access to this existing data and an automated means for identifying chemical analogues based on structural features. TOPKAT also employs conservative statistical analysis and validation procedures. A TOPKAT analysis of a non-congeneric data set is potentially useful for generating mechanism hypotheses when very little prior knowledge is available for classifying chemicals according to mechanism. TOPKAT can also serve as a potentially valuable supple

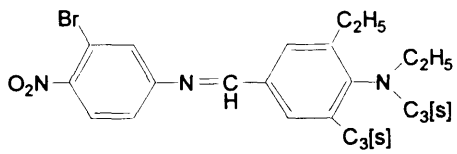
Figure 12-4: OncoLogic Cancer Expert System



ment to expert rule-based approaches when the latter are relatively undeveloped (e.g. in the developmental toxicity field), or as an approximate preliminary screening capability when little expertise is available. Recognizing current limitations of their approach, TOPKAT developers are moving towards the goal of more "mechanistic" models in the sense of restricting model development to smaller, more well-defined activity endpoints and chemical classes (1 1). (See examples in table 12-1.) These initiatives offer hope for improving the utility of such models for toxicity prediction in TSCA screening.

Traditional QSAR studies attempt to discover new, previously unknown mathematical relationships for predicting activity from chemical structure. In contrast, an expert system aims to reproduce the human expert decision process for evaluating chemical toxicity by codifying current knowledge. An example of the latter is the OncoLogic cancer prediction expert system (15), being developed as a collaborative effort between the cancer experts within the EPA/SAT and outside expert systems programming consultants. OncoLogic is an artificial intelligence, rule-based expert system that can be applied to a wide range of non-congeneric chemicals, but that relies on a chemical class, mechanism-based approach to cancer prediction (22). It incorporates literally thousands of discreet rules for characterizing each of a variety of chemical classes based on the cancer expertise of the SAT. Due to the enormous size of this undertaking, the program is currently operational only for metals, polymers, fibers, and a few classes of organic chemicals, with capabilities for other chemical classes still under development. OncoLogic has a hierarchical structure as represented in figure 12-4. In the top-most levels, structural considerations are used to exclude molecules from concern on the basis of factors such as molecular weight, volubility, and bioavailability. A chemical is then classified according to properties and structure features until sufficient characterization allows application of

Figure 12-5: Sample OncoLogic Carcinogenesis Evaluation Justification Report



Summary: The level of caranogenicity concern from this aromatic amine is LOW - MODERATE.

Justification: In general, the level of carcinogenicity concern of an aromatic amine is determined by considering the number of rings, the presence or absence of heteroatoms in the rings; the number and position of amino groups; the nature, number and position of other nitrogen-containing "amine generating groups", and the type, number and position of additional sustituents.

...

The evaluation of this compound proceeds as if the di-akyl substituted amino group, NR1R2 [where R1= ethyl; R2=sec-propyl] were a free amino group. The influence of the N-akyl groups on the bioactivation of the compound is considered at the end of the evaluation.

...

The reduced electron conducting properties of the intercyclic linkage are expected to lower the overall level of concern. Therefore the level of carcinogenicity is reduced to MODERATE.

the SAR rules for a specific chemical class, such as shown for aromatic amines. A sample output is also shown in figure 12-5, illustrating a portion of the lengthy, mechanism-based rationale provided to support the prediction of LOW-MODERATE concern for this sample chemical. The rules consider issues such as metabolism and activation to reactive electrophiles, and accurately reproduce and communicate the mechanism-based rationale of the EPA/SAT cancer experts. Since these rules are distinct program units, they can be modified as knowledge advances. By communicating the detailed rationale for EPA/SAT cancer predictions, OncoLogic allows industry and others to identify and challenge prevailing SAR assumptions by directed research. OncoLogic also makes the current cancer prediction expertise of the SAT more accessible and widely available within and outside EPA.

An expert system is only as good as the rules and knowledge upon which it is based. Although a few other expert systems are currently available for toxicity prediction, most are based on more limited knowledge and expertise than OncoLogic, and have more limited appeal for toxicity screening. One possible exception is in the area of metabolism prediction. Many chemicals require metabolic activation as a precondition to toxicity. In many of these cases, modeling the conditions for metabolic activation, or modeling the metabolizes instead of the parent compounds, is the key to developing a successful SAR for toxicity prediction. OncoLogic, for example, incorporates numerous metabolism rules for organic chemicals. Metabolism expert systems, such as MetabolExpert (14), provide industry and the larger SAR community valuable access to expertise concerning likely metabolic pathways and products for many chemicals of concern.

As has been stressed, mechanism-based chemical classification is one of the primary requirements for successful SAR model development, providing the scientific basis for the chemical analogue selection step in figure 12-2. The criteria for analogue selection is a key area of uncertainty in many SAR models since it is usually based on organic chemistry principles derived independent of the biology, and may not reflect similarity in terms of biological mechanisms of toxicity. The problem of choosing appropriate analogues is illustrated by the example of peroxisome proliferators. These chemicals are structurally diverse, yet have highly similar pleiotropic, toxicological responses that strongly suggest a common receptor-mediated mode-of-action. Hence, it is the biological

response, not the apparent chemistry, that directs one to group these chemicals into a common class for the purpose of SAR model development.

A few research groups are considering biological means for classification of chemicals for use in SAR studies. In recent years, Bradbury and coworkers at EPA's Environmental Ecology Laboratory have moved away from traditional chemical class-based QSARs for predicting aquatic toxicity and towards the generation of biological mechanism-based QSARs (7). They first established a mode-of-action knowledge base covering a broad range of chemicals, exposure regimes and endpoints. Empirical assessment of toxicity mechanism was then determined by consideration of joint toxic action studies, physiologically-based toxic response syndromes, and single chemical dose-response curves, yielding a variety of toxic mode-of-action classifications (e.g. baseline narcosis, oxidative phosphorylation uncouplers and respiratory inhibitors). Only after such biological classifications were determined were efforts centered on QSAR analysis and understanding the chemical mechanisms and structural criteria for underlying activity. Hence, in applications, biologically-based chemical classifications would define criteria for choosing appropriate chemical analogues and identify the relevant QSAR for use in a toxicity screening application.

A second example is provided by the Rules Induction Method for Predicting Chemical Carcinogenesis developed by Bahler and Bristol in a collaborative effort between NIH/NIEHS and academia (3). This is an automated, decision-tree approach where rules for use in prediction are mathematically induced from available data, rather than obtained by human experts. In contrast to TOPKAT, the rules are derived from both chemical and biological "attributes". These include: Salmonella mutagenicity (SAL); electrophilic structural alerts; route of administration; MTD; subchronic organ pathology (up to 59 organ types, up to 40 morphological lesions); and miscellaneous *in vitro* short term test results. As a predictive screening tool, this approach has the limitation that it requires subchronic pathology information from the rodent bioassay, and some *in vitro* assay information. However, significant cost savings would be realized in generating these data as opposed to carrying out a fill 2-year rodent carcinogenicity bioassay. This biological information was also utilized by "human experts" in a recent NTP-44 prospective carcinogenicity prediction exercise where human expert predictive performance was judged superior to the performance of "pure" SAR methods, i.e. methods based on chemical structure alone (l). Perhaps because the Rules Induction Method and human experts used much the same information in their assessments, the Rules Induction Method performed nearly as well as the human experts in the prediction exercise. An alternative use of this information, germane to the present discussion, is as a means for defining biological mechanism-based chemical classifications for subsequent SAR analysis, i.e. using each rule branch to define a possibly distinct mode-of-action chemical class. Two sample rules are shown in table 12-2. All chemicals satisfying Rule#l comprise a subclass of active carcinogens likely to be mechanistically distinct from other active carcinogens in terms of biological attributes. Hence, the structural features common to chemicals in this rule class, and distinct from chemicals belonging to the remaining actives or inactives could constitute an SAR model for prediction of carcinogenicity.

Table 12-2: Sample Induction Rules for Chemical Carcinogenisis Predictor

Rule #1:

IF chemical mutates Salmonella
AND adjusted rat MTD=<750 mg/kg/day,
THEN class is positive.
(Rule true for 90% of 147 chemicals in training set)

Rule #5:

IF chemical does not mutate Salmonella
AND there is no subchronic pathology in male rat pituitary, spleen, or urinary/bladder,
AND there is no subchronic pathology in female rat kidney
THEN class is negative

A discussion of promising technologies for toxicity screening also should include mention of major advances in computational chemistry and 3D modeling and visualization that are yielding greater understanding of the detailed molecular-level interactions and changes ultimately responsible for the toxicity. Such methods are advancing in tandem with increasing computational capabilities, and increasing knowledge of chemical reaction mechanisms, the structure and function of biological receptors, metabolic enzyme activity (cytochrome P-450s, and glutathione), and DNA interactions implicated in various forms of toxicity. Computational chemistry studies have had a significant impact, for example, on understanding of the structural and electronic requirements for DNA-adduct formation and carcinogenicity of polycyclic aromatic hydrocarbons (incorporated as rules in OncoLogic).

Another example illustrating how 3D modeling tools, more commonly employed in drug design, can be applied to toxicity problems once a receptor-based mechanism for a toxicity endpoint has been proposed or established is provided by a recently published 3D-QSAR model for endocrine disrupters developed by Wailer and coworkers at the National Health and Environmental Effects Research Lab of EPA (21). The toxicity of a class of endocrine disrupters was postulated to be due to interaction with the steroid hormone receptor. Since the structure of this receptor was unknown, ligand requirements of the receptor were inferred from a comparison of the three-dimensional structures of known steroid receptor ligands (such as estradiol). The final computerized 3D QSAR model provides a means for predicting the potential receptor binding affinity of any chemical relative to endogenous steroids given the 3D structure of the untested chemical. After sufficient validation, such a model could serve as a rapid screen for potential endocrine disrupters and be used for setting testing priorities, i.e. by identifying chemicals most likely to compete with endogenous steroid ligands.

Major advances are being realized in information and computational fields that could eventually lead to improved SAR models for toxicity prediction. Advances in neural networks, artificial intelligence, molecular visualization and modeling all have the potential to generate previously undiscovered models from existing data. However, these models will be subject to the same biology-imposed constraints as previously discussed, and share many of the same limitations as current methods. For example, the major disadvantage of current neural network-based SARs for toxicity prediction is that the model cannot be easily interpreted in terms of the original molecular parameters and, hence, the scientific basis for the NN model is practically undecipherable, making it difficult to scientifically rationalize a model prediction or define the bounds of application of the model.

CONCLUSIONS

Improvement in current SAR models used in TSCA chemical screening will be achieved most effectively by close interaction and feed-back between SAR application and toxicity prediction, laboratory testing, validation, and research into chemical mechanisms of toxicity. SAR is an extremely multidisciplinary field, applicable to a wide range of problems and endpoints. Since SAR modelers often lack expertise in toxicology, and toxicologists tend to be unfamiliar with the tools and assumptions of SAR modelers, there is a need for increased interaction, collaboration, and education between these two groups. The SAR modeler can guide the toxicologist in choosing experimental measures of toxicity, appropriate chemicals for SAR model design, and, in cases where a preliminary SAR model exists, approximate dose ranges to test for an effect. The toxicologist can provide the SAR modeler with insight into possible modes-of-action, practical and experimental design constraints (i.e. a reality check), and sources of uncertainty and error in the data.

The EPA/SAT is the regulatory arm that bears primary responsibility for the development and application of SAR to TSCA chemical screening, and the SAR expertise, models, and data used by

the SAT represent an extremely valuable resource for serving the interests of health and environmental protection. However, the SAT operates in relative isolation from the larger SAR community, and each could benefit from increased communication and collaboration. Although some outreach efforts have been made by the SAT, through development and dissemination of computerized SAR programs such as ECOSAR and OncoLogic, a major obstacle to increased collaboration is the confidential nature of the data used in SAR model development. Similarly, industry and other government regulatory agencies, such as FDA, often have large stores of toxicity data that are considered proprietary. A recent ECVAM workshop on "Integrated Use of Alternative Approaches for Predicting Toxic Hazard" produced the following recommendations (4):

- "Companies should be encouraged to make non-confidential data available to external groups, perhaps via an independent organisation such as ECVAM. For confidential data, they should be encouraged to review the need to maintain that confidentiality on a regular (continual) basis."
- "Regulatory agencies should be encouraged formally to establish (Q)SARs utilising submission data. . . . Companies should also be encouraged to develop (Q)SARs using their confidential data. Such (Q)SAR models should then be placed in the public domain, along with supporting non-confidential data."

While legitimate and defensible concerns of industry regarding the need for confidentiality of chemical structures and processes should not be minimized, there also should be greater acknowledgment of the value of available toxicity data, and recognition that more universal access to high quality toxicity data for SAR model development serves the best interests of the entire SAR community.

A number of issues have been identified that impact on the accuracy and utility of current SAR models for TSCA screening. To reiterate, the major trends likely to lead to the greatest improvement in SAR models for toxicity prediction are:

- greater understanding of mechanisms of toxicity for endpoints of potential concern;
- greater use of biologically-based chemical classifications in SAR development;
- better ways to represent molecules and their detailed biological interactions in SAR models;
- continued use of prospective validation and testing to validate and refine SAR models;
- greater understanding of role of metabolism in various forms of chemical toxicity;
- increased knowledge of role of biological receptors in toxicity and elucidation of the structure, function, and ligand requirements of relevant receptors;
- testing to fill crucial data gaps for chemical classes and toxicity endpoints of potential concern;
- greater effort to declassify some CBI and proprietary toxicity data that have little commercial value for use in SAR development;
- improved interaction between SAR users in industrial, academic and government research, and regulatory agencies to improve SAR models .

Screening the TSCA existing chemical inventory for all manner of potentially harmful effects in a timely manner is a huge challenge that cannot be met by testing alone. While testing deals with the generation of new data, SAR is above all the study of existing data and how to make best use of these data to predict the biological activity and properties of chemicals for which data are unavailable. SAR provides the only real alternative to expensive and time consuming laboratory testing. Hence, reliance on SAR methods will no doubt increase in response to increased budgetary and societal pressures to reduce costs and limit the use of animals in toxicity screening. While these methods hold great promise, the danger is that in response to such pressures SAR models will be invoked prema-

turely for some toxicity endpoints, will be extended beyond where they are likely to be valid or reliable, and will be used without sufficient oversight and testing verification. SAR offers a means for achieving better health and environmental protection by enabling a strategic and intelligent application of limited testing resources, by identifying the highest priority risk chemicals, and by attacking a much larger portion of the problem than is currently being addressed by testing alone. Better SAR models also have the clear potential to further reduce the need for testing or eliminate testing in some circumstances. SAR models will improve in tandem with increased understanding and availability of data upon which to base and refine such models. However, fulfilling the promise of SAR requires proper application of these methods, clear recognition of the limitations of SAR, and appreciation for the essential roles of research and strategic testing in SAR model development and refinement.

REFERENCES

1. Ashby, J., and Tennant, R. W., "Prediction of Rodent Carcinogenicity for 44 Chemicals: Results," *Mutagenesis* 9:7-15, 1994.
2. Auer, C. M., et al., "SAR – The U.S. Regulatory Perspective," *SAR and QSAR in Environmental Research* 2:29-38,1994.
3. Bahler, D., and Bristol, D. W., "The Induction of Rules for Predicting Chemical Carcinogenesis in Rodents," *Intelligent Systems for Molecular Biology,* L. Hunter, et al. (eds.) (Menlo Park, CA: AAAI/MIT Press, 1993).
4. Barret, M. D., et al., "The Integrated Use of Alternative Approaches for Predicting Toxic Hazard. Report and Recommendations of ECVAM (European Centre for the Validation of Alternative Methods) Workshop," *Alternatives to Laboratory Animals,* in press.
5. Benigni, R., Andreoli, C., and Giuliani, A., "QSAR Models for Both Mutagenic Potency and Activity: Application to Nitroarenes and Aromatic Amines," *Environmental and Molecular Mutagenesis* 24:208-219, 1994.
6. Benigni, R., and Giuliani, A., "Quantitative Structure-Activity Relationship (QSAR) Studies in Genetic Toxicology: Mathematical Models and the 'Biological Activity' Term of the Relationship," *Mutation Research* 306:181-186, 1994.
7. Bradbury, S., "Quantitative Structure Activity Relationships and Ecological Risk Assessment: An Overview of Predictive Aquatic Toxicology Research," *Proceedings of Workshop on Screening Technologies (1995),* in press.
8. Clements, R. G., et al., "The Use of Quantitative Structure-Activity Relationships (QSARs) as Screening Tools in Environmental Assessment," *Environmental Toxicology and Risk Assessment - 2nd Volume,* ASTM STP 1216, J.W. Gorsuch et al. (eds.) (Philadelphia, PA: American Society for Testing and Materials, 1993.
9. CQSAR: Comprehensive Structure Activity Relationships Using Multivariate Information, Biobyte Corp., Claremont, CA.
10. ECOSAR: A Computer Program for Estimating the Ecotoxicity of Industrial Chemicals Based on Structure Activity Relationships, National Center for Environmental Publications and Information, U.S. Environmental Protection Agency, Cincinnati, OH.
11. Enslein, K., et al., "Computation of Developmental Toxicity Potential by QSTR Models in the TOPKAT Program, " *Animal Test Alternatives,* in press.
12. Hansch, C., "Quantitative Structure-Activity Relationships and the Unnamed Science," *Accounts of Chemical Research 26:147-153, 1993.*
13. Leo, A., and Weininger, D., "CLOGP: Medchem Software Release 3.5," *Medicinal Chemistry Project* (Claremont, CA: Pomona College, 1992).
14. MetabolExpert: Expert System for Predicting Metabolizes, CompuDrug, USA, Austin, TX.
15. OncoLogic: An Expert System that Evaluates the Carcinogenic Potential of Chemicals, LogiChem, Inc., Boyertown, PA.
16. Richard, A. M., "Application of SAR Methods

to Non-Congeneric Data Bases Associated with Carcinogenicity and Mutagenicity: Issues and Approaches," *Mutation Research* 305:73-97, *1994.*

17. Richard, A. M., "Role of Computational Chemistry in Hazard ID: Mechanism-based SARs," *Proceedings of ATSDR sponsored Workshop on Screening Technologies (1995),* in press.
18. TOPKAT: Toxicity Prediction Software, Health Designs Incorporated, Rochester, NY, 14604
19. U.S. Environmental Protection Agency/ European Community, *Joint Project on the Evaluation of (Quantitative) Structure-Activity Relationships: Final Report,* July, 1993.
20. Wagner, P.J., Nabholz, V. and Kent, R.J., "The New Chemicals Process at EPA: Structure Activity Relationships for Hazard Identification and Risk Assessment," *Proceedings of ATSDR sponsored Workshop on Screening Technologies (1995),* in press.
21. Wailer, C.L., Minor, D. L., and McKinney, J. D., "Examination of the Estrogen Receptor Binding Affinities of Phenyl-substituted Hydrocarbons of Environmental Concern Using Three-dimensional Quantitative Structure-activity Relationships," *Environmental Health Perspectives,* in press.
22. Woo, Y.-T., et al., "Development of Structure Activity Relationship Rules for Predicting Carcinogenic Potential of Chemicals," *Proceedings of ATSDR sponsored Workshop on Screening Technologies (1995),* in press.
23. Zeeman, M., Nabholz, J. V., and Clements, R. G., "The Development of SAR/QSAR for Use Under EPA's Toxic Substances Control Act (TSCA): An introduction" *Environmental Toxicology and Risk Assessment - 2nd Volume,* ASTM STP 1216, J.W. Gorsuch et al. (eds.) (Philadelphia, PA: American Society for Testing and Materials, 1993).