2003

Quantitative Structure-Activity Relationship Models for Prediction of Estrogen Receptor Binding Affinity of Structurally Diverse Chemicals

Steven P. Bradbury
Patricia K. Schmeider
Gerald Ankley
Ovanes Mekenyan
John D. Walker

SETAC PRESS

Environmental Toxicology and Chemistry, Vol. 22, No. 8, pp. 1844–1854, 2003
Printed in the USA
0730-7268/03 $12.00 + .00

Annual Review

QUANTITATIVE STRUCTURE–ACTIVITY RELATIONSHIP MODELS FOR PREDICTION OF ESTROGEN RECEPTOR BINDING AFFINITY OF STRUCTURALLY DIVERSE CHEMICALS

PATRICIA K. SCHMIEDER,*† GERALD ANKLEY,† OVANES MEKENYAN,‡ JOHN D. WALKER,§ and STEVEN BRADBURY†

†U.S. Environmental Protection Agency, National Health and Environmental Effects Research Laboratory, Mid-Continent Ecology Division, 6201 Congdon Boulevard, Duluth, Minnesota 55804

‡Bourgas University (As. Zlatarov), Laboratory of Mathematical Chemistry, Department of Physical Chemistry, 118010 Bourgas, Bulgaria

§TSCA Interagency Testing Committee, U.S. Environmental Protection Agency, Office of Pollution, Prevention, and Toxic Substances, Washington, DC 20460

(*Received* 17 *July* 2001; *Accepted* 24 *June* 2002)

Abstract—The demonstrated ability of a variety of structurally diverse chemicals to bind to the estrogen receptor has raised the concern that chemicals in the environment may be causing adverse effects through interference with nuclear receptor pathways. Many structure–activity relationship models have been developed to predict chemical binding to the estrogen receptor as an indication of potential estrogenicity. Models based on either two-dimensional or three-dimensional molecular descriptions that have been used to predict potential for binding to the estrogen receptor are the subject of the current review. The utility of such approaches to predict binding potential of diverse chemical structures in large chemical inventories, with potential application in a tiered risk assessment scheme, is discussed.

Keywords—Estrogen receptor Relative binding affinity Structure–activity relationships Hazard identification

INTRODUCTION

The estrogen receptor (ER) is a member of the nuclear receptor superfamily, a group of ligand-inducible transcription factors (including androgen, thyroid hormone, retinoic acid receptors, etc.) that are represented throughout the animal kingdom in vertebrates, arthropods, and nematodes [1]. All nuclear receptor proteins have a similar architecture, including six domains referred to as A through F, with the ligand binding function found in the moderately conserved E domain. Ligand binding to the receptor is known to cause a conformational change in the receptor, which, dependent on the nature and conformation of the ligand, allows the recruitment of additional factors (coactivators or corepressors) that dictate the outcome of ligand-receptor-effector interactions with the DNA [2]. Much of the current understanding of ER molecular and cellular biology and pharmacology is a result of decades of study in the pharmaceutical industry. Research into the structure and function of the human ER, including studies of ER activation and antagonism, has been driven by the need to understand the role of the receptor in breast cancer as well as the search for drugs that can block undesirable effects while retaining the beneficial effects of estrogens. The need to predict the possibility of ER-mediated drug toxicities also requires understanding of ER-ligand interactions. The well-documented toxic effects of the synthetic estrogen diethylstilbestrol (DES), used as a pharmaceutical agent for the prevention of miscarriages in women [3], are manifested through its action as an ER agonist. In fact, DES is routinely used as a positive control in studies concerning ER activation.

In addition to being of interest from a pharmaceutical perspective, there is considerable concern regarding the potential for adverse effects in humans and wildlife from environmental exposures to chemicals that can potentially interact with and disrupt any of a number of endogenous hormone pathways (e.g., estrogen, androgen, thyroid). Collectively, chemicals with potential hormone-disruptive actions are referred to as endocrine-disrupting chemicals (EDCs). Chemicals of concern as environmental estrogens are known to include a seemingly bewildering array of structurally diverse environmental contaminants that have been shown to bind to ERs, which may in part reflect the seemingly promiscuous nature of this particular nuclear receptor [4]. Estrogenic EDCs have been postulated to be potentially responsible for a wide variety of adverse conditions in humans, including breast cancer and endometriosis in women and effects on reproductive health, such as decreased semen quality, in men [5,6]. Organochlorine chemicals, including some pesticides, polychlorinated biphenyls (PCBs), and/or their metabolites have been proposed to be potentially estrogenic in humans. Strong evidence also exists for adverse effects associated with specific EDCs in a variety of wildlife species, including invertebrates, fish, reptiles, and birds [7,8]. Perhaps the clearest examples of an effect attributable to environmental estrogens in fish are reported in studies of Sumpter and coworkers, who demonstrated that estrogenic xenobiotics were responsible for the abnormal production of vitellogenin (egg yolk precursor protein typically found in females) in male fish [9–11]. Chemicals identified as estrogenic in their studies included natural and synthetic steroidal chemicals emanating from human wastes and, in some instances, nonsteroidal degradation products of alkylphenol ethoxylate surfactants, again illustrating diverse chemical structures linked to a common adverse outcome.

The U.S. Environmental Protection Agency has been man-

* To whom correspondence may be addressed (schmieder.patricia@epa.gov).

dated to implement screening and testing programs to identify estrogenic chemicals associated with food-use pesticides and drinking water. A multi-stakeholder advisory group, the Endocrine Disruptor Screening and Testing Advisory Committee, recommended that the U.S. Environmental Protection Agency expand the screening to incorporate classes of industrial chemicals in the Toxic Substances Control Act inventory [12]. As a result, potentially estrogenic chemicals may need to be identified among thousands of compounds. The efficiency and effectiveness of a screening program, such as that being developed for EDCs, would be supported by the establishment of a rapid ranking and prioritization scheme to ensure that testing resources are focused on those compounds likely to cause adverse effects. One approach to chemical prioritization would be the utilization of quantitative structure–activity relationship (QSAR) models for prediction of ER binding affinity and ranking of untested chemicals.

This article reviews recent applications of QSAR models for the prediction of ER binding affinity, with consideration of model applications for ranking and prioritization of large chemical data sets. The prioritization concept is designed to support initial selection of chemicals of highest concern before initiating EDC tiered testing [12].

THE ESTROGEN RECEPTOR

Relationships between chemical structure and biological activity are highly dependent upon the quality and quantity of the biological data used to develop and evaluate QSAR models. In this context it is important to discuss data availability for prediction of estrogenic activity. There are multiple in vitro and in vivo measures of estrogen-dependent processes, including ER binding affinity, gene activation, cell proliferation, uterine growth, vitellogenin induction, etc. Although limited QSAR models have been developed for more biologically complex endpoints, such as ER-mediated gene activation [13] and cell proliferation [14,15], the majority of the modeling has involved prediction of the less biologically complex interaction of ligand binding to the ER. Most binding data available were measured using ER from uterine cytosol preparations from a number of species, including humans, rats, mice, cattle, and sheep [14]. Binding studies using oviparous species, such as fish, commonly utilize liver cytosol as their ER source. The discovery of a second human (h)ER receptor subtype [16], with subsequent work showing differences in hERα and hERβ ligand binding affinities [17–19], makes it necessary to closely evaluate the source of binding data when attempting to understand the biological implications of model predications. Although there is evidence of multiple ERs in wildlife species, including birds and fish [20], relatively little is known about differential binding affinities of these subtypes in species other than humans and rodents.

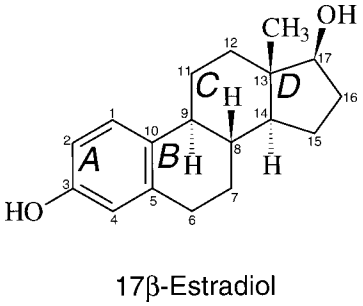
The discovery of multiple ER subtypes has been especially important in the pharmaceutical industry, where this information is leading to the development of selective ER modulators [1,18,19,21]. For instance, Kraichely et al. [22] have developed agonists and antagonists specific for hERα and hERβ subtypes. These compounds offer the potential to block adverse hERα-mediated events associated with breast cancer, while not interfering with the beneficial effects of hERβ-mediated processes. Additionally, the development of these novel ligands, often with the aid of QSARs, allows characterization of conformations induced in ER subtypes by ligands associated with different activities, which in turn enables a better understanding of how ligand-receptor conformations relate to estrogen agonist or antagonist behavior. This information also serves as the basis for development of QSAR models for ER antagonism [23].

Much of our current understanding of ligand binding behavior, from the point of view of the receptor, has been gained by studies of hER using site-directed mutagenesis, as well as X-ray crystallographic analyses. These studies, which describe the characteristics of receptor ligand binding domains, also help in the interpretation of binding data and are critical to the generation of reasonable hypotheses as to the chemical structural descriptors most likely associated with ER binding. The ability to generate and test such hypotheses is important to the discovery of new, more selective, pharmaceuticals, but arguably, it is also important to the development of models to predict potential estrogenicity of pesticides and industrial chemicals.

The development of QSAR models for ER binding affinity for the extremely structurally diverse chemicals in commerce requires that different criteria be established for model predictability than those used for drug discovery. The development of drugs involves the search through a universe of often structurally similar chemicals to arrive at the best few candidates, with optimal activity for a specified endpoint (e.g., hERα binding). However, models for industrial chemical screening and prioritization for further testing of adverse effects often consider a larger universe of very diverse chemicals, and these models search for all candidates that might be capable of producing the activity. Risk management needs, therefore, may require that models be skewed to minimize the possibilities of false-negative predictions, at the expense of an elevated level of false-positive predictions (e.g., predictions of no binding activity when experimental data indicate otherwise) [12]. For drug design, however, one attempts to minimize the number of false-positives (e.g., prediction of binding affinity when there is none above a threshold of interest) to avoid expensive synthesis and evaluation of candidates unlikely to meet the necessary objective [24]. This appreciation of the different objectives in drug design versus risk assessment is needed when evaluating the types of QSAR models that have been developed for predicting ER binding. In summary, models must be evaluated in the context of the data on which they are based, the purpose for which they are designed, and the resulting constraints placed on the interpretation of model predictions. This review attempts to evaluate the types of QSAR modeling approaches that have been used to predict ER binding affinity in the context of these factors as they reflect the differing needs of the drug discovery versus risk assessment.

ER BINDING AFFINITY MODELS

Arguments can be made that it is necessary to map and model a receptor to fully understand ligand-receptor interactions. In fact, if receptor geometry is known, intermolecular docking studies can be performed [25]. Reports of the hERα crystal structure with bound agonists (17β-estradiol, E2, [26]; DES, [27]) and antagonists (raloxifene, [26]; 4-hydroxy tamoxifen, [27]) have advanced the understanding of critical features of the ER ligand binding domain and allowed for further exploration of the roles of activation factor regions of the protein in binding and coactivator interactions. Site-directed mutagenesis techniques have also allowed study of the contributions of single receptor residues or sequences critical



17β-Estradiol

Fig. 1. Structure of 17β-estradiol showing substitution and A, B, C, and D ring nomenclature.

to ER binding [28]. An attempt has also been made to extend this knowledge of the crystalline hER to additional species, such as the rainbow trout [29]. This information serves as a basis for docking experiments of ligands within the receptor and for gaining insights to ligand-receptor interactions. Receptor-site docking models, however, are computationally intensive and are therefore not practical as rapid screening tools for large numbers of chemicals. As such, docking models are not considered further in this review.

Whereas docking models may not be applicable for rapidly evaluating large chemical data sets, other mechanistically reasonable ligand-based QSAR modeling approaches have been developed with application to large numbers of diverse chemicals. Modeling approaches based on either calculations from two-dimensional (2-D), or three-dimensional (3-D) chemical structure, or combinations of both, have been used. These approaches are discussed separately below.

2-D QSAR ER binding models

A series of 27 QSAR regression equations were calculated by Gao et al. [14] for a variety of chemical classes (e.g., 16α-substituted estradiols, 11β-substituted estradiols, 17α-substituted estradiols, 17α-XCH=CH estradiols, multisubstituted estradiol derivatives, hexestrol derivatives, metahexestrol derivatives, 2-arylindene and indenone derivatives, 2,3-diarylindene derivatives, 2-phenylindole derivatives, etc.). The data sets used were obtained from the literature and largely comprised binding data from uterine cytosol preparations (predominently ERα) from human, rat, mouse, lamb, and calf, but they also included cellular proliferation and growth inhibition endpoints. Equations were based on data sets generated by chemical class, species, and incubation temperature, and they ranged from a minimum of six chemicals for the binding of 7α-undecylestradiol derivatives to calf uterine cytosol at 4°C to a maximum of 51 chemicals tested for their hER affinity. The authors correlated activity with physicochemical descriptors including commonly used substituent parameters [30]. Descriptors calculated from 2-D molecular information gave insights as to where and what modifications to the basic E2 structure (Fig. 1) are and are not tolerated through the study of a series of substituents at specific sites. Some of the general conclusions from the modeling of substituted estradiol derivatives included the following: A phenolic hydroxy group mimicking the 3-OH on the A-ring of E2 appears repeatedly as the most important factor for receptor-ligand interactions; substituents that increase the electron density on the phenolic ring also appear to increase binding affinity; there is apparently no consistent, positive hydrophobic interaction between ligand substituents and receptor (recognizing that the ABCD tetracyclic core structure of steroidal estrogens is already hydrophobic and contributes to the bulk of ligand binding); there is a hydrophobic interaction between ERα and hydrophobic constituents on the 11β position of estradiol; sometimes steric character, or in other cases polar character, of substituents results in reduced binding affinity; for nonsteroidal ligands, substituents that result in conformational restrictions resulting in deviations from planarity facilitate receptor binding; and, positive correlations with potency and hydrophobicity appear frequently in comparisons of ligand structure and biological endpoints measured in intact cells, probably reflecting the correlation between hydrophobicity and cell membrane penetration. Although this classical QSAR correlative analysis approach has great utility in better understanding the receptor and requirements for ligand binding within the structural series presented, the usefulness of this approach to predicting ER binding affinity for a large number of structurally unrelated chemicals in industrial chemical inventories is uncertain.

A recent application of a 2-D modeling approach to the prediction of ER binding affinity was that of Tong et al. [31], with their development of a holographic QSAR (HQSAR) approach. The technique is based on counts of substructural molecular fragments in which all linear, branched, and overlapping structural fragments, in the size range of four to seven atoms, are generated for each molecule. Generated fragments are placed into fixed-length arrays to produce molecular holograms. The model produced by HQSAR is dependent on length and information contained in the generated fragments. The information contained in the holograms is dependent on parameters that describe fragment size characteristics and fragment distinction parameters (atoms, bonds, connections, hydrogens, and chirality) [31]. The modeling approach was applied to chemicals evaluated for their binding to hERα, rER, or calf ER. The hER and rER data sets comprised 31 chemicals, including 19 steroids, four triphenylethylenes, three DES derivatives, two bis (4-hydroxylphenyl) alkanes, and three phytoestrogens. The calf ER data set had 37 congeners of 2-phenylindole, 10 congeners of 5,6-dihydroindolo[1,2-α]isoquinoline, and six estrogens and antiestrogens. The performance of this 2-D QSAR model was compared with two 3-D QSAR models applied to the same three data sets. The Comparative Molecular Field Analysis (CoMFA) and COmprehensive DEscriptors for Structural and Statistical Analysis (CODESSA) [31] models incorporate 3-D chemical structure assessments, and/or identification of a bioactive conformer, and molecular alignment (details of these modeling approaches are addressed in subsequent sections). Comparisons of model r^2 and q^2 (derived from leave-one-out [LOO] cross-validation procedure) fits, and number of principal components used in the partial least-squares analysis revealed the 2-D HQSAR model results were qualitatively comparable to the CoMFA models and outperformed the CODESSA model. In this initial assessment, the HQSAR models were seen as advantageous, in comparison to the 3-D approaches, because of the shorter computational times, and they were generally described to be more convenient than either CoMFA or CODESSA models [31], although exact logistic details were not provided. In addition to performing well in a comparison of model summary statistics, the HQSAR model showed good agreement with experimental data and CoMFA performance in the prediction of four compounds that had been excluded from the training set. As with many modeling approaches, the model fit and number as well as composition of principal components will

change depending on the characteristics of the training set used in model development.

The 2-D HQSAR approach was compared with an additional 3-D QSAR modeling approach (i.e., COmmon REactivity PAttern [COREPA]) (see *3-D QSAR ER binding models: COREPA applications*). The authors of this comparison reported that, whereas the HQSAR and COREPA approaches for estimating relative binding affinity (RBA) would benefit from additional evaluation, the results of the exercise were sufficient to conclude that the uncertainties in the application of the 2-D HQSAR model were too great to consider the use of HQSAR to predict the ER binding affinity of the 23,460 additional chemicals for which ER binding predictions were being sought [32].

3-D QSAR ER binding models: Approach

The interaction of chemicals with steroid receptors is a 3-D process [2,18], with the receptor conforming to the shape of the ligand, and the ligand, if flexible, having its conformation altered by binding to the receptor [18]. Therefore, 3-D QSAR techniques that consider ligand conformational flexibility and permit quantification of chemical global and local steric and electronic characteristics might be expected to have greater predictive capability than correlative 2-D approaches. Hopfinger and Tokarski [25] differentiated 3-D QSAR modeling approaches into two types: receptor-dependent and receptor-independent analyses. Most of the modeling approaches applied to larger data sets utilize the receptor-independent modeling approach.

Perhaps the most widely available and often-employed receptor-independent 3-D QSAR approach utilizes CoMFA (for a complete description of this modeling approach, see Hopfinger and Tokarski [25] and Cramer et al. [33]). Briefly, the core of the CoMFA approach is a data table containing a column of the biological activity or measured property and multiple columns of structural parameters for every compound in the study. Each structural parameter column records the intensity of a particular type of field interaction, at a point in space, with a probe atom of specified charge and steric properties. A QSAR is sought that relates the intensity of the calculated field parameter properties to the biological activity being modeled. Because thousands of field property measurements are correlated to one measure of biological activity per compound, partial least-squares analysis is used to arrive at the optimum solution to the oversolved data set generated [25]. Using partial least-squares analysis, principal components are generated and accepted based on their improvement of the ability to predict the dependent variable (e.g., ER binding affinity). The predictive ability of a CoMFA model is evaluated by the cross-validated correlation coefficient q^2, while the general fit of the model to the data is assessed by r^2. An additional predictive correlation coefficient, q^2_{pred}, is assessed only after a QSAR has been used to predict the activity of chemicals not included in the training set. This q^2_{pred}, therefore, provides a bottom-line assessment of the analysis and should indicate deficiencies in a model that are due to the inability to detect all relationships in the training set [25].

Because a CoMFA model is a linear equation derived from thousands of terms that are difficult to interpret, contour maps of prevalent field properties are used to assist in visualizing model results (as further described by Hopfinger and Tokarski [25]). Typically, contour maps are displayed showing areas of the most positive and negative associations. These contour maps are helpful in suggesting new compounds likely to have high or low property values and for ranking potential compounds for further consideration (e.g., for synthesis in drug exploration applications).

A critical step in the CoMFA approach is the user's identification of the active conformation of each ligand in the training set; similarly, the user provides the relative-binding geometry (the alignment) of each ligand to the common receptor, or to a lead compound (e.g., E2). Typically, CoMFA has been described as performing best on rigid analogues with well-defined pharmacophores, which minimized uncertainty in aligning chemicals to the lead compound. However, as pointed out by Hopfinger and Tokarski [25], iterative application of CoMFA on a training set as a function of ligand conformation or relative alignment remains to be further explored, with no major limitation except cpu constraints. An exploration of this type was considered in the recent work of Xing et al. [34] when using CoMFA to compare estrogen receptor subtypes.

A different approach to 3-D QSAR modeling of receptor binding was first presented conceptually by Bradbury et al. [35,36] and Mekenyan et al. [37,38]. Recently, the approach has been more fully evaluated for prediction of ER binding affinity in two companion papers [39,40]. The COREPA approach differs conceptually from CoMFA in several key respects. One difference is the inclusion in COREPA of all energetically reasonable conformers for each chemical modeled, as opposed to a single low-energy conformer. Additional unique aspects of the COREPA modeling approach involve the process of selecting stereoelectronic descriptors as well as the statistical determination of parameters associated with the biological activity of interest. The approach also includes both active and inactive ligands in all modeling steps, and the contrast between active and inactive ligand parameters in identifying reactivity patterns is a critical feature. In this context, one seeks to gain insights into the receptor by understanding characteristics of not only ligands that bind but also characteristics of chemicals that do not bind. Descriptors that best distinguish active from inactive ligands are represented quantitatively and are statistically compared. Thus, the approach identifies distinct calculable chemical parameters associated with binding, which allows examination of the biological credibility of hypothesized chemical–biological interactions. In this sense, the COREPA approach resembles the 2-D QSAR linear regressions that also allow direct inspection of individual chemical parameters.

The COREPA approach was developed, in part, because of reports of the lowest energy gas phase conformer potentially being the least likely to interact with solvents or macromolecules [41]. Therefore, in complex systems such as biological tissues, it is quite possible that the minimal energy conformer does not interact with the receptor. An important observation from COREPA modeling studies to date is that there can be significant differences in the value of a given stereoelectronic descriptor among the energetically reasonable conformers of a single chemical. For example, in the analysis of the nonsteroidal ER ligand β-zearalanol, 49 reasonable conformers were generated, which covered a range of 0.449 eV for the lowest unoccupied molecular orbital (E_{LUMO}) (−0.4605 to −0.0112) and 0.189 eV for E_{HOMO} (−9.4623 to −9.2738). Similar variations were observed for other calculated parameters for other nonsteroidal and steroidal compounds [39]. Therefore, there can be as much variability in electronic structure between the conformers of a single chemical as might be observed between

chemicals. The observation that relatively small energy differences between conformers can result in significant variations in electronic structure highlights the necessity of including all energetically reasonable conformers when defining common reactivity patterns and the importance of this consideration in evaluating common features of ligands that bind the ER.

To provide a means to compare the COREPA approach to those previously described, a summary of the method is provided (see [13,37–39] for the conceptual basis, detailed mathematical formulations, and illustrations of the method). First, an algorithm is used to convert 2-D structures to 3-D structures, followed by an exhaustive conformer generation routine. Conformer generation in COREPA is based on a combinatorial procedure that initiates from molecular topology and generates all conformers consistent with steric constraints and expert rules [42]. Up to 500 of the sterically most distinct points from the conformational space for each chemical are selected (see [39] for conformer generation and optimization details). Geometric dissimilarity is assessed based on Euclidean distances between the sums of interatomic distances for the conformers. Each of the generated conformations is submitted to a strain-minimization technique, conformational degeneracy is assessed, and geometries are optimized, resulting in some of the conformers quenching into the same energy minima. Conformers are subsequently screened to eliminate those whose ΔH_f^0 is 20 kcal/mol or more greater than the conformer with the absolute energy minimum, to retain only energetically reasonable conformers. The rationale for using multiple conformers arises from experimental evidence that the free energy of binding for steroid hormones to receptors is in the range of -10 to -20 kcal/mol [43–45], which can provide the necessary energy to elevate conformers from the low(est) energy state during binding. Conformers selected within this range of $\Delta\Delta H_f^0$ are energetically reasonable from a thermodynamic and kinetic perspective [36–38,46]. Conformers of each chemical are considered to be a statistical ensemble, based on Boltzman's statistics. The selected electronic, steric, and physicochemical descriptors, chosen for hypothesized association with the biological endpoint, are calculated for each reasonable conformer of each chemical in the training set. All conformers of a given chemical are plotted across a molecular descriptor axis, thus forming a discrete distribution for the chemical relative to the selected descriptor.

After conformers for each chemical have been generated, a probabilistic approach is used to assess common reactivity patterns of biologically similar chemicals [38]. In step 1, training subsets of chemicals with similar biological response (active vs inactive) are selected. In the case of receptor binding predictions, those chemicals with RBAs above a user-defined threshold are considered active, and ligands with RBAs below a user-defined nonbinding threshold are considered inactive. In step 2, a set of descriptors associated with the biological activity of interest are established by a value of the parameter (e.g., E_{HOMO}) for each chemical conformer is calculated and distributed across a parameter axis as a continuous probability density function (Fig. 2). Probability distributions of the chemicals are normalized to unity, allowing assessment of similarity between ligands (with respect to a molecular descriptor) by using similarity indices based upon intersections of ligand distributions. The descriptors that provide the maximal measure of similarity among chemicals within a training subset and least overlap between active and inactive distributions are assumed to be related to biological activity. Well-defined, or distinct, patterns for descriptors are observed when the conformer distributions for the chemicals with the same biological activity are in phase. Finally, in step 3, common reactivity patterns for biologically similar molecules are obtained as products of the probabilistic distributions for specific stereoelectronic descriptors associated with active and inactive chemicals in the training sets. (See [38–40] for application of Euclidean distance to evaluate pattern overlap; to compare distributions derived from different chemical training sets; or to determine the extent to which conformer distributions are influenced by a specific chemical.) As with other QSAR model approaches, the stability of a reactivity pattern can be assessed by a LOO procedure (see original references for details).

To facilitate screening of large chemical data sets, common reactivity patterns have been coded into decision trees (see [13,39] for details). Decision trees consist of multiple hierarchically ordered rules that capture specific stereoelectronic descriptors that define the common reactivity patterns for the biological endpoint of interest. Each energetically reasonable conformer of a chemical is evaluated in the decision tree, with probabilities assigned for meeting parameter requirements [39]. If a chemical has to meet two successive requirements to equal or exceed an activity threshold, the total probability of meeting both requirements is obtained as a product of the probabilities of meeting the two requirements separately (see [39] for details). If the value of the descriptor calculated for a conformer falls outside of at least one of the parameter ranges, then the overall probability of having an RBA above that threshold is 0. The approach therefore offers flexibility in establishing hazard-ranking protocols for unknown compounds based on choices of RBA thresholds and confidence limits around pattern maxima. The probability outcomes from the decision tree are not viewed in absolute terms but instead permit a relative ranking of unknown chemicals in terms of the likelihood that the chemical possesses the biological activity of interest, above a user-defined threshold. This approach allows for the processing of numerous unknown chemical structures in a timely fashion, and has been applied, for example, to screen tens of thousands of chemicals in large chemical inventories.

The current formulation of COREPA uses a probabilistic characterization of conformer distributions, allowing comparisons between training sets as well as comparisons of individual chemical distributions to that of the training set. User-defined activity thresholds can also be applied to training sets, and the local electronic character of toxicophores can be evaluated. Yet, although the recently published version of COREPA allows the user to assess chemical similarity in terms of two or more descriptors simultaneously, the algorithm is limited in not providing evaluation of interdependencies of conformer distributions in a multivariate manner. Overall, because of the differences in the conceptual approaches and statistical measures used in COREPA and those used in CoMFA and 2-D approaches (linear regression and HQSAR), it is difficult to statistically compare models. For instance, the LOO analysis, as currently applied in COREPA, gives a qualitative indication of the stability of the model but does not provide a means of direct comparison to q^2_{LOO} statistics of CoMFA, CODESSA, or HQSAR approaches.

3-D QSAR ER binding models: CoMFA applications

One of the first uses of CoMFA to predict ER binding was described by Waller et al. [47], who performed an assessment

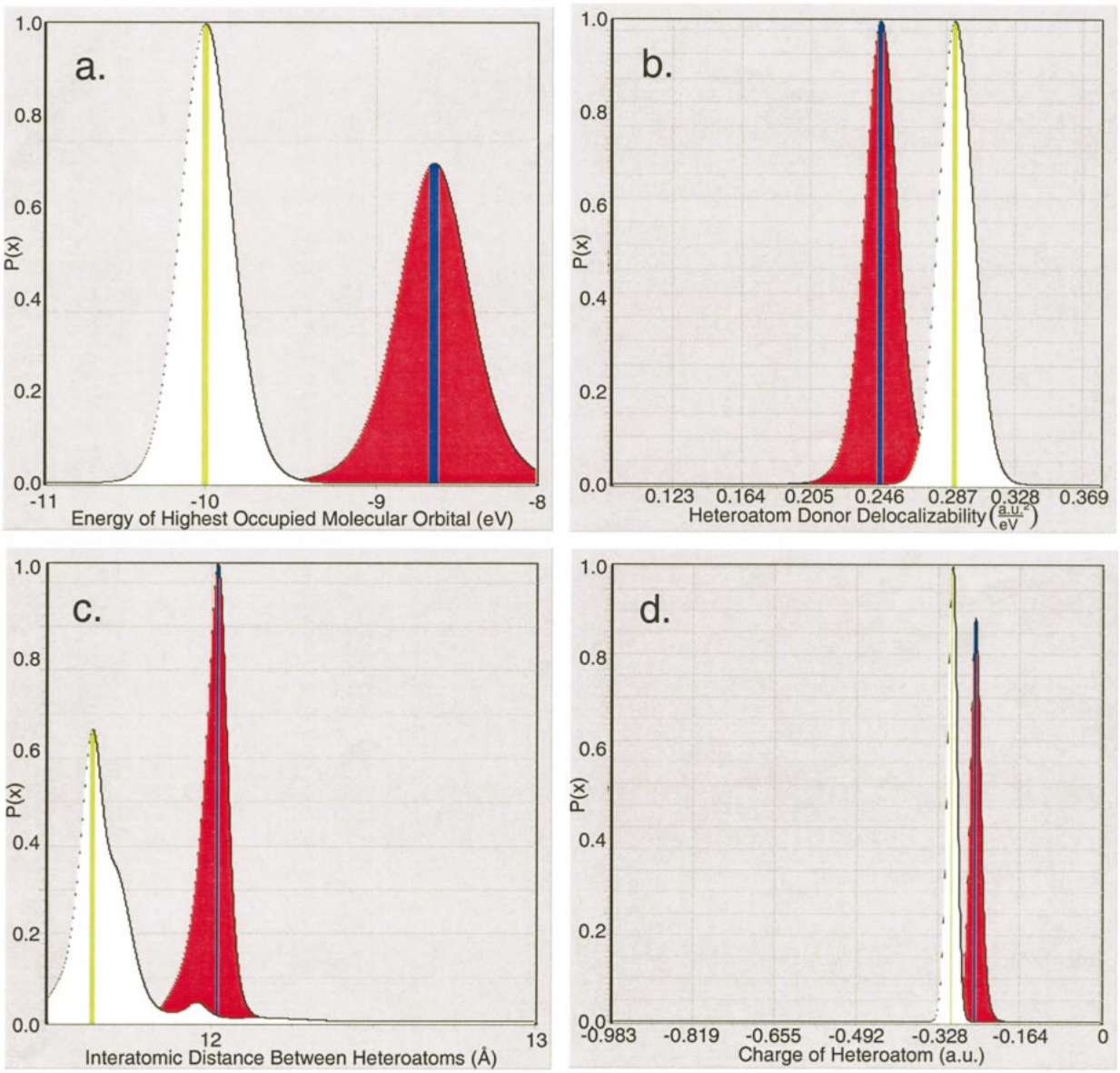


Fig. 2. The reactivity patterns for one estrogen receptor binding range representing the strongest affinity ligands (relative binding affinity >150%), based on the descriptors found to be associated with estrogen receptor binding: (**a**) global nucleophilicity (represented by energy of the highest occupied molecular orbital, [E_{HOMO}]); (**b**) electron donor delocalizabilities of heteroatoms [SE(R:O,N,Cl,F,S)]; (**c**) interatomic distances between electronegative heteroatoms [d(R_R)]; and (**d**) atomic charge on heteroatoms [Q(R:O,N,Cl,F,S)]. The reactivity pattern of active ligands is in red, whereas the pattern of inactive ligands is in white (modified from Bradbury et al. [39]).

of measured rat and mouse ER binding to a set of 58 chemicals. The training set spanned eight orders of magnitude in binding affinity. The low-energy structure of each chemical was calculated and aligned to E2. Steric (van der Waals) and electrostatic (Coulombic [charge–charge] interaction energies) were computed. Hydrophobic interactions were computed separately using Hydrophobic INTeractions (HINT) software [47]. Three models that were developed and evaluated against sets of natural and synthetic ER ligands included CoMFA, using steric and electrostatic parameters; HINT, for hydrophobic interactions; and CoMFA plus HINT.

The combined CoMFA/HINT QSAR model provided increased internal predictive ability and statistical robustness, in comparison with the standard CoMFA or HINT models alone. In the combined model, the contribution of steric, electrostatic, and hydrophobic fields were 27, 45, and 28%, respectively, providing a model constructed with three principal components, with an r^2 of 0.881, and a LOO cross-validation q^2_{LOO} of 0.590, a significant improvement over the standalone CoMFA models [47]. The predictability of the CoMFA/HINT model was assessed using an unbiased method; a LOO cross-validated analysis; and a fitted (i.e., non–cross-validated) analysis. The training set divided into eight subclasses (phenols, phthalates, phytoestrogens, DDT metabolites, PCBs, pesticides, DES-related compounds, and steroids) and predictability assessed for each group. The predictability was only within 1 to 2 log units using the unbiased assessment. Using the LOO cross-validated analysis, the best average error achieved was for PCBs, 0.65 log units, but this value was 0.9 to 1.5 log units for the other chemical classes. The fitted analysis predicted within 0.3 and 1.0 log unit average errors, with best results recorded for PCBs and phthalates.

Waller et al. [47] provided stereoviews of generalized steric, electrostatic, and hydrophobic field contributions from the combined CoMFA/HINT contour plots. These depictions provided generalizations of where steric bulk increases were, and were not, tolerated; where positive and negative potentials of ligands were favored; and where hydrophobic bulk was most desired in an ER ligand, in relation to the model ligand, E2. How these parameters apply to structurally divergent ligands (in comparison to E2) within the ER is not readily apparent. The authors suggest that the contour plots could aid in qualitative biological activity predictions and would be of most utility in computer-aided drug design, but the application to predictions of activity for very divergent chemical structures would not be as useful. The use of contour plots as negative images of the ligand binding domains was cautioned by Waller et al. [47], largely because of the limited information contained in the training sets and the hypothetical nature of the alignment rule. However, with that qualification, this model seemed to substantiate the authors' original pharmacophore hypothesis, suggesting that compounds with two electronegative centers separated by a rigid hydrophobic framework are preferred ER ligands [48].

Several limitations of the approach were pointed out by the authors, including lack of sufficient data for any one species (i.e., rat or mouse) over the set of diverse structures. The subjectivity of user-determined alignments to E2 was also acknowledged, with the application of the Steric and Electrostatic ALignment algorithm [49] presented as a preliminary approach to overcome some of the subjectivity for 3-D ligand-based identification.

Subsequent to the study by Waller et al. [47], Tong and coworkers [31] described the use of CoMFA and CODESSA to predict ER binding using data sets described previously for the HQSAR model (see *2-D QSAR ER binding models*). Steric and electrostatic 3-D descriptors were used in the CoMFA model, while CODESSA used 2-D and 3-D structures and quantum-chemical properties. For the CoMFA model, a single calculated lowest energy conformation was used for each chemical, and alignment was achieved by using a least-squares fitting of pharmacophore elements between E2 and each compound in the training set. The pharmacophore points of E2 used for alignment were the centroids of the A- and D-rings and the C7 atom of the B-ring (Fig. 1). For the 2-phenylindole derivatives in the training set, the investigators aligned the centroid of the phenyl ring of the indole moiety, the centroid of the 2-phenyl ring, and the indole nitrogen. The statistical models generated were used to predict biological activity of chemicals not utilized in the training set, in this case four chemicals. As mentioned earlier, comparison of statistical model parameters, and predictions for four chemicals excluded from the training set, revealed good agreement between the CoMFA and HQSAR models for the three data sets evaluated, with the CODESSA model not performing as well.

Recently, an approach was presented that involves the application of a variety of filters, substructural alerts, and classification trees, in combination with HQSAR and CoMFA QSAR modeling approaches, in a multiphase approach for screening large inventories of chemicals with regard to their potential to bind to the ER [24]. The authors use the series of prescreens to eliminate large numbers of chemicals from further consideration, so that only a restricted set of chemicals preliminarily identified as active ligands would subsequently be subjected to HQSAR and computationally intensive CoMFA QSAR modeling. However, the authors only evaluated the application of HQSAR and CoMFA in this study, leaving the preliminary filtering approach for subsequent assessment (see [24] for further details).

The training set used to develop the CoMFA and HQSAR models consisted of 130 diverse compounds with measured rER binding affinity, including steroids, DES derivatives, DDT and metabolites, bisphenol A derivatives, PCBs, alkylphenols, parabens, kepone, and α,α-dimethyl-β-ethylallenolic acid. Alignments to E2 were based on the phenolic A-ring, a D ring, 7α and 11β substituents, and a hydrophobic backbone. The alignment for chemicals not possessing elements similar to E2 was determined individually ''on the basis of an appropriate rationalization consistent with overall alignment strategy'' [24], thus introducing a degree of subjectivity that is difficult to reproduce or generalize for larger screening applications.

The CoMFA model generated from the training set of 130 chemicals had an r^2 of 0.908 and a q^2_{LOO} of 0.655. Predictions of ER binding affinity fell within a fivefold measure of the measured values for 85% of the chemicals. For comparison, a HQSAR model showed significantly poorer performance than the CoMFA model, with an r^2 of 0.756 and a q^2_{LOO} of 0.585. The q^2_{LNO} (leave-N-out) cross-validation statistic (generating by randomly selecting N groups to leave out one at a time) also indicated that the CoMFA model had greater stability than HQSAR for this larger and more structurally diverse set of chemicals [24], compared to the previous CoMFA to HQSAR evaluation, which used a smaller and more restricted data set [31]. Efforts to improve the CoMFA model, using region focusing (i.e., giving more weight to regions showing a greater influence on the standard partial least-squares model) and exploring all possible orientations and placements of aligned molecules in the CoMFA region [50], did not improve the model. Reducing lattice spacing or filtering energies also had no significant impact on model results. In addition, the model showed no improvement with the inclusion of the octanol–water partition coefficient, but it did show improvement with the inclusion of an indicator for presence or absence of a phenolic ring.

The CoMFA models, with and without a phenolic ring indicator, and the HQSAR model were tested for their ability to predict the binding of 48 chemicals not in the training set. This new data was extracted from larger data sets of Waller et al. [47] (mostly ERα from rodent uterine cytosol preparations) and Kuiper et al. [51] (hERα). Two chemicals that appeared in both data sets were treated separately, for a total of 50 considered in the analysis provided by Shi et al. [24]. The majority of chemicals comprising the two test sets were 21 biphenylols, 3 biphenyldiols, and 2 biphenyls, with the remainder including steroidal and nonsteroidal chemicals. Shi et al. [24] reported that while the two CoMFA models were similar in prediction of active chemicals, the model with the phenolic ring indicator did a better job of discriminating active versus inactive ligands in the two test sets. However, this conclusion was difficult to assess given that Kuiper et al. [51] used an RBA cutoff of log RBA ≥ -2.51, but Waller et al. [47] appeared to use an activity threshold of log RBA ≥ -5.00 (after conversion from pKi and normalization by Shi et al. [24]). This resulted in four chemicals with log RBA between -2.51 and -4.819 considered active in the Waller et al. [47] test set that would have been considered inactive by Kuiper et al. [51]. Nonetheless, both CoMFA models clearly were more predictive than the HQSAR model. Of 44 active ER

ligands, CoMFA and CoMFA with phenolic ring indicator predicted the activity of 31 and 33 of the chemicals, respectively, to within an order of magnitude of the measured values, resulting in a q^2_{pred} of between 0.62 and 0.71. The HQSAR model predicted 16 chemicals with more than an order of magnitude error, including six chemicals in error by more two orders of magnitude. The resulting HQSAR model q^2_{pred} of 0.15 and 0.22 for the two data sets indicate the poor predictability. While the two CoMFA models were reported to be similar in their prediction of active chemicals, the model with the phenolic ring indicator was described by Shi et al. [24] as doing a better job of discriminating active versus inactive ligands. It would seem difficult to support a conclusion of model performance drawn across all 50 chemicals when the parameter the models are said to be capable of discriminating is defined differently in the two data sets. Differences in activity thresholds are not uncommon between investigators as a result of differences in procedure and biological models, and often modelers are forced to use data from multiple sources because of limited data availability. Yet caution must be exercised when combining empirical data from multiple sources and when interpreting resulting QSAR model predictions.

The use of predictive models for ranking and prioritization of large chemical inventories requires that the models be evaluated with regard to anticipated rates of false-negative and false-positive predictions, as mentioned previously. The minimization of false-negative predictions is especially desirable for risk assessments in which ranking of chemicals for further tiered testing risk is the objective. However, for drug design it is often desirable to minimize false-positives in order to avoid expensive unproductive leads. Regardless of the intended usage, the rate of false-negative or -positive predictions is assessed against a user-established criteria. In the preceding example, the authors reported no false-negative predictions for the 42 distinct chemicals in the test set, even for chemicals with activities of one million–fold below that of E2 [24]. If, in fact, a cut-off of log RBA < -4.0 is used for both data sets, the statement is accurate. However, if the Kuiper et al. [51] cut-off of log RBA < -2.5 is used as the criteria, at least one false-negative prediction would be made for the hER data set and five or six for the mER data set, depending on the CoMFA model used. For instance, application of the CoMFA model with a phenolic ring indicator would predict the three substituted biphenyls to be negative using the log RBA < -2.5 cut-off. One must also note that application of this threshold to the mER data set would mean that seven chemicals having experimental log RBA values of <-2.5 would now be considered inactive. As mentioned above, this assessment is complicated by the application of different thresholds to the experimental data, but it is done to illustrate the importance of considering limitations of the experimental data when establishing user-defined criteria as well as the intended use of model predictions. It is also uncertain what the results of the present analysis would have been if phase I and phase II (i.e., rejection filters, 2-D structural alerts, pharmacophore searches, and classification and regression trees) of the multiphase approach had been applied to prescreen and eliminate chemicals from consideration by the more cpu-intensive CoMFA QSAR models, as proposed by Shi et al. [24]. It would seem that consideration of chemicals determined to be inactive by previously applied 2-D screens and filters would be required for determination of an overall false-negative rate. Regardless of these uncertainties concerning the proposed multiphase approach of Shi et al. [24], it is apparent from the information presented that the use of the 3-D CoMFA QSAR models was preferable to the 2-D HQSAR model for prediction of ER binding for large diverse chemical inventories.

3-D QSAR ER binding models: COREPA applications

A COREPA model was developed to predict hERα binding affinity ranges using a training set of 45 chemicals (26 steroids and 19 nonsteroidal ligands) [39]. Reactivity patterns were established for identifying and ranking compounds within a series of binding affinity ranges relative to estradiol binding of 100%, including RBA values of >150%, 100 to 10%, 10 to 1%, and 1 to 0.1%. Local, global, and steric descriptors used in the model were restricted to those hypothesized to be associated with ER binding, based on previous studies with a variety of model receptors [35–38,45,47,48,52–56] (see [39] for computational details). Through the COREPA analysis the authors determined that ER RBA could be predicted based on three parameters: global nucleophilicity (represented by energy of the highest occupied molecular orbital, E_{HOMO}), interatomic distances between nucleophilic heteroatoms, and electron donor capability of these heteroatoms.

The stereoelectronic requirements of the reactivity pattern associated with each RBA activity range were organized in a hierarchical decision tree whose output was an estimated probability that a conformer would bind to the hERα within a given RBA range. The first part of the tree consists of absolute screens (i.e., the necessary structural requirements for eliciting minimal ER binding affinity, set at RBA $\geq$ 0.1%, for this demonstration model). For example, enantiomers of steroids were required to have trans–trans (B/C trans and C/D trans) ring fusion as an absolute stereochemistry screen. Global nucleophilicity was an absolute electronic requirement, with an E_{HOMO} of -9.95 eV selected as the necessary nucleophilicity threshold. The presence of negatively charged (i.e., electron donor) atomic sites was also a basic requirement, specified as any heteroatomic site [Q(R) = O, N, Cl, F, S] with a donor-delocalizability (i.e., atomic nucleophilicity) in the range of 0.239 to 0.279 (arbitrary units)^2/eV. Conformers that had E_{HOMO} values of less than -9.95 eV, electronegative sites not meeting the specified donor delocalizability, or steroids not conforming to stereochemical requirements of the natural enantiomer were assigned a 0% probability to bind to hERα with an RBA >0.1%. Conformers that passed these absolute requirements were then compared to the E_{HOMO} (-8.99 eV $< E_{HOMO}$), interatomic distance [11.77 $<$ d(R_R) $<$ 12.22], and charge [$-0.272 <$ Q(R) < -0.233] screens associated with the activity pattern of chemicals having an RBA >150%. The identification of a ligand with a binding affinity within an RBA range requires that at least one conformer meets all three specified parameter ranges. If a compound was not identified as having an RBA >150%, it was then screened to determine if it had an RBA between 10 and 100% [E_{HOMO} (-9.44 eV $< E_{HOMO}$)], interatomic distance [10.62 $<$ d(R_R) $<$ 10.95], and charge [$-0.273 <$ Q(R) < -0.236], and so on until all activity range screens are applied. In general, most predictions were within an order of magnitude of observed RBA values. Consistent with the conservative bias in selection of reactivity patterns, the majority of predictions that were not within an order of magnitude of the observed RBA values overpredicted binding potential.

The COREPA for hERα binding affinity [39] was further assessed for its ability to predict ER binding affinity using

experimental data from three species (rER, mouse ER, and hER) [40] combined from nine different investigators and for chemical structures not well-represented in the training set [39]. The model, developed to demonstrate an approach for chemical ranking and prioritization for regulatory application, was evaluated by placing chemicals in predicted RBA ranges. False-negatives and -positives were defined as all chemicals falling outside of the order of magnitude range in which the observed value was found. Since the cut-off for activity used by the authors was <0.1 RBA (or log RBA < -1), it is not possible to assess the rate of false-negatives that would have been achieved had a much lower cut-off (e.g., log RBA < 2.5 [51], or log RBA < -4 [24]) been used, except to say that the false-negative rate would be even lower.

The modeling approach has also been applied to a more homologous series of chemicals (alkylphenols) with relatively weak binding affinity in order to determine more precise predictors of activity within a confined range and, in this instance, for the more biologically complex endpoint of ER gene activation [13]. The COREPA approach has also been recently extended to the discrimination of ER antagonists from agonists [23] and has also been applied to prediction of ligand affinity to other steroid hormone receptors [36,37,57].

Although the scientific understanding of the dynamic nature of ligand-receptor-effector interactions and the importance of conformational change in these interactions has gained wide acceptance [2], the availability of computational methods for consideration of chemical conformational flexibility is still somewhat exploratory. This is particularly true when attempting to apply such methods to large numbers of chemicals in a timely manner. Several exploratory methods for the generation and timely evaluation of multiple conformers of hundreds to thousands of chemicals have been considered in the COREPA modeling approach (discussed previously) and its many modifications [58]. One method for reducing computational time for initial screening of large chemical inventories has been the consideration of only a single conformer per chemical, but not necessarily the single lowest energy conformer. This assessment has been accomplished through use of a directed tweak algorithm for a rapid (and approximated) conformational search [59] in conjunction with broadening the descriptor windows in decision trees. This approach was taken in an effort to further minimize false-negatives, although at the expense of increasing the possible number of false-positive predictions. The 3-D screening of large inventories by accurate analysis of chemical conformational space has also been achieved using a genetic algorithm [38]. New procedures are under development to allow adjustment of the number of conformers encompassing conformational space of chemicals in an attempt to reduce effects due to the nondeterministic character of genetic algorithms. Further assessments of the sensitivity of model predictions to the methods of conformer generation and evaluation are ongoing. This is an area that is certain to receive more attention in the near future as more computational approaches become available and as awareness of the importance of considering the conformational space of flexible chemicals increases.

SUMMARY

Although the models reviewed here have all been applied to the prediction or description of ER binding, the purposes for model development and intended application were not identical in all instances. Even with similar use in mind, the conceptual as well as practical differences in model development preclude direct comparison.

Models developed to describe relationships of biological activity to quantifiable aspects of chemical structure are highly dependent upon the quality, quantity, and specificity of the biological endpoint data used. Predictive models will, of course, be most reliable when chemicals of unknown activity fall within the structure-space of the training sets used for model development. Extrapolations beyond the area of the training set, with respect to model parameters determined to be associated with activity, will be more prone to prediction error. In this context it is important to discuss data availability for QSAR models to predict estrogenic activity. Theoretically, one could use in vitro or in vivo measures of estrogenicity, including ER binding affinity, gene activation, cell proliferation, uterine growth, etc. Practically, however, the more closely the biological endpoint accurately reflects the direct chemical–biological interaction, the more accurate a QSAR model will be. Precise modeling of any in vivo endpoint is only possible when the relationship between the applied dose and chemical concentration at the site of action is known, or predictable. Whole-organism responses for which the ultimate site of action is unknown, the chemical form and concentration acting at that site is unknown, and for which the toxicodynamic response is downstream from the site of chemical action, will continue to elude the effective and scientifically credible application of QSAR modeling techniques.

Quantitative structure–activity relationship models have been developed for several ER-related in vitro endpoints, such as ER binding affinity [14,24,31,35,36,39,40], gene activation [13], and cell proliferation [14,15]. The ER ligand binding affinity serves as an example of a well-defined endpoint for which chemical–biological interaction can be successfully predicted. The QSAR models developed to predict ER binding have demonstrated sufficient predictive capability to warrant evaluation as a tool for prioritizing and ranking large numbers of diverse chemicals for subsequent tiered in vitro or in vivo screening for endocrine disruption. However, a multistep iterative approach should be adopted when extending existing models to screen large, structurally diverse chemical inventories. First, preliminary (i.e., existing) models need to be examined to determine the most important structural parameters that are mechanistically descriptive of the biological activity being modeled (e.g., ER binding). Next, the preliminary model is used to select nonsimilar chemicals from the inventory to be screened. Carefully selected compounds representative of unique structures are empirically evaluated. The new information is used to assess whether chemical structural descriptors selected from preliminary models continue to quantify meaningful aspects of dissimilar chemical structures that are predictive of the biological activity of concern. This then leads to development of a next generation QSAR model based on the more structurally robust training set of chemicals that better represents the structural diversity of the chemical inventory. The chemical inventory is again screened and prioritized with regard to the biological activity of interest. Additional modeling iterations, including chemical selection, endpoint measurement, development of next-generation predictive models, and subsequent screening of chemical inventories, continues only until acceptability criteria are met. Model acceptance is evaluated in comparison to user-defined acceptability criteria (e.g., false-negative, false-positive rates), which is ultimately a function of the risk management application.

The application of this type of approach is recommended for credible advancement of QSAR models for prioritization and ranking of large numbers of chemicals when chemical screening for potential to cause adverse environmental effects is a risk assessment goal.

Acknowledgement—The authors thank A. Richard and C. Russom for their thoughtful comments on the draft manuscript. D. Spehar and R. LePage assisted in manuscript preparation. Reference herein to any specific commercial product, process, or service by trademark, manufacturer, or otherwise, does not necessarily constitute its endorsement, recommendation, or favoring by the Toxic Substances Control Act (TSCA) Interagency Testing Committee (ITC) or any of the 16 U.S. Government organizations represented on the ITC, including the U.S. Environmental Protection Agency (U.S. EPA). Views expressed in this article do not necessarily reflect policies of the ITC or any of the U.S. Government organizations represented on the ITC, including the U.S. EPA. This article has been subjected to review by the National Health and Environmental Effects Research Laboratory and approved for publication. Approval does not signify that the contents reflect the views of the Agency, nor does mention of trade names or commercial products constitute endorsement or recommendation for use.

REFERENCES

1. Egea PF, Klaholz BP, Moras D. 2000. Ligand-protein interactions in nuclear receptors of hormones. *FEBS Lett* 476:62–67.
2. Katzenellenbogen JA, O'Malley BW, Katzenellenbogen BS. 1996. Tripartite steroid hormone receptor pharmacology: Interaction with multiple effector sites as a basis for the cell- and promoter-specific action of these hormones. *Mol Endocrinol* 10: 119–131.
3. Herbst A, Bern HE. 1981. *Developmental Effects of Diethylstilbestrol (DES) in Pregnancy*. Thieme-Stratton, New York, NY, USA.
4. Elsby R, Ashby J, Sumpter J, Brooks A, Pennie W, Maggs J, Lefevre P, Odum J, Beresford N, Paton D, Park BK. 2000. Obstacles to the prediction of estrogenicity from chemical structure: Assay-mediated metabolic transformation and the apparent promiscuous nature of the estrogen receptor. *Biochem Pharmacol* 60:1519–1530.
5. Toppari J, Larsen JC, Christiansen P, Giwercman A, Grandjean P, Guillette LJ Jr, Jégou B, Jensen TK, Jouannet P, Keiding N, Leffers H, McLachlan JA, Meyer O, Müller J, Rajpert-De Meyts E, Scheike T, Sharpe R, Sumpter J, Skakkebæk NE. 1996. Male reproductive health and environmental xenoestrogens. *Environ Health Perspect* 104:741–803.
6. Kavlock RJ, Daston GP, DeRosa C, Fenner-Crisp P, Earl Gray L, Kaattari S, Lucier G, Luster M, Mac MJ, Maczka C, Miller R, Moore J, Rolland R, Scott G, Sheehan DM, Sinks T, Tilson HS. 1996. Research needs for the risk assessment of health and environmental effects of endocrine disruptors: A report of the U.S. EPA–sponsored workshop. *Environ Health Perspect* 104:715–740.
7. Ankley GT, Giesy JP. 1998. Endocrine disruptors in wildlife: A weight of evidence perspective. In Kendall R, Dickerson R, Suk W, Giesy JP, eds, *Principles and Processes for Assessing Endocrine Disruption in Wildlife*. SETAC, Pensacola, FL, USA, pp 349–368.
8. Tyler CR, Jobling S, Sumpter JP. 1999. Endocrine disruption in wildlife: A critical review of the evidence. *CRC Crit Rev Toxicol* 28:319–361.
9. Purdom CE, Hardiman PA, Bye VJ, Eno NC, Tyler CR, Sumpter JP. 1994. Estrogenic effects of effluent from sewage treatment works. *Chem Ecol* 8:275–285.
10. Desbrow G, Routledge EJ, Brighty GC, Sumpter JP, Waldock M. 1998. Identification of estrogenic chemicals in STW effluent. 1. Chemical fractionation and in vitro biological screening. *Environ Sci Technol* 32:1549–1558.
11. Jobling S, Sumpter JP. 1993. Detergent components in sewage effluent are weakly oestrogenic to fish: An in vitro study using rainbow trout (*Oncorhynchus mykiss*) hepatocytes. *Aquat Toxicol* 27:361–372.
12. U.S. Environmental Protection Agency. 1998. Endocrine disruptor screening and testing advisory committee (EDSTAC). Report 98. Office of Prevention, Pesticides and Toxic Substances, Washington, DC.
13. Schmieder P, Aptula A, Routledge E, Sumpter J, Mekenyan O. 2000. Estrogenicity of alkylphenolic compounds: A 3-D structure-activity evaluation of gene activation. *Environ Toxicol Chem* 19: 1727–1740.
14. Gao H, Katzenellenbogen J, Garg R, Hansch C. 1999. Comparative QSAR analysis of estrogen receptor ligands. *Chem Rev* 99: 723–744.
15. Wiese TE, Polin L, Palomino B, Brooks SC. 1997. Induction of the estrogen specific mitogenic response of MCF-7 cells by selected analogues of 17β-estradiol: A 3-D QSAR study. *J Med Chem* 40:3659–3669.
16. Mosselman S, Polman J, Dijkema R. 1996. ERβ: Identification and characterization of a novel human estrogen receptor. *FEBS Lett* 392:49–53.
17. Kuiper GG, Carlsson BO, Grandien K, Enmark E, Haggblad J, Nilsson SGJA. 1997. Comparison of the ligand binding specificity and transcript tissue distribution of estrogen receptors alpha and β. *Endocrinology* 138:863–870.
18. McDonnell DP. 1999. The molecular pharmacology of SERMs. *Trends Endocrinol Metab* 10:301–311.
19. Kim S-H, Katzenellenbogen JA. 2000. Triarylethylene bisphenols with a novel cycle are ligands for the estrogen receptor. *Bioorg Med Chem Lett* 8:785–793.
20. Hawkins M, Thornton J, Skipper J, Dotte A, Thomas P. 2000. Identification of a third distinct estrogen receptor and reclassification of estrogen receptors in teleosts. *Proc Natl Acad Sci USA* 97:10751–10756.
21. Levenson AS, Jordan VC. 1999. Selective oestrogen receptor modulation: Molecular pharmacology for the millennium. *Eur J Cancer* 35:1974–1985.
22. Kraichely DM, Sun J, Katzenellenbogen JA, Katzenellenbogen BS. 2000. Conformational changes and coactivator recruitment by novel ligands for estrogen receptor- and estrogen receptor-: Correlations with biological character and distinct differences among SRC coactivator family members. *Endocrinology* 141: 3534–3545.
23. Schmieder PK, Koleva Y, Mekenyan OG. 2002. A reactivity pattern for discrimination of ER agonism and antagonism based on 3-D molecular attributes. *SAR/QSAR Environ Res* 13:353–364.
24. Shi L, Fang H, Tong W, Wu J, Perkins R, Blair R, Branham W, Dial S, Moland C, Sheehan DM. 2001. QSAR models using a large diverse set of estrogens. *J Chem Inf Comput Sci* 41:186–195.
25. Hopfinger A, Tokarski J. 1997. Three-dimensional quantitative structure-activity relationship analysis. In Charifson PS, ed, *Practical Application of Computer-Aided Drug Design*. Marcell-Dekker, New York, NY, USA.
26. Brzozowski AM, Pike ACW, Dauter Z, Hubbard RE, Bonn T, Engstrom O, Ohman LGGL, Gustafsson JA, Carlquist M. 1997. Molecular basis of agonism and antagonism in the oestrogen receptor. *Nature* 389:753–758.
27. Shiau AK, Barstad D, Loria PM, Cheng L, Kushner PJ, Agard DA, Greene GL. 1998. The structural basis of estrogen receptor/coactivator recognition and the antagonism of this interaction by tamoxifen. *Cell* 95:927–937.
28. Ekena K, Weis KE, Katzenellenbogen JA, Katzenellenbogen BS. 1997. Different residues of the human estrogen receptor are involved in the recognition of structurally diverse estrogens and antiestrogens. *J Biol Chem* 272:5069–5075.
29. Mori T, Sumiya S, Yokota H. 2000. Electrostatic interactions of androgens and progesterone derivatives with rainbow trout estrogen receptor. *J Steroid Biochem Mol Biol* 75:129–137.
30. Hansch C, Leo A, Hoekman D. 1995. *Exploring QSAR. Hydrophobic, Electronic, and Steric Constants*. American Chemical Society, Washington, DC.
31. Tong W, Lowis DR, Perkins R, Chen Y, Welsh WJ, Goddette DW, Heritage TW, Sheehan DM. 1998. Evaluation of quantitative structure-activity relationship methods for large-scale prediction of chemicals binding to estrogen receptor. *J Chem Inf Comput Sci* 8. 5A-I.
32. Walker JD, Fang H, Perkins R, Tong W. 2003. QSARs for EDPSD 2: The integrated 4-phase model. *QSAR Combinatorial Science* 22:89–105.
33. Cramer RI, Patterson D, Bunce J. 1988. Comparative molecular

field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins. *J Am Chem Soc* 110:5959–5967.
34. Xing L, Welsh WJ, Tong W, Perkins R, Sheehan DM. 1999. Comparison of estrogen receptor alpha and beta subtypes based on comparative molecular field analysis (CoMFA). *SAR/QSAR Environ Sci* 10:215–237.
35. Bradbury SP, Mekenyan OG, Ankley GT. 1996. Quantitative structure-activity relationships for polychlorinated hydroxybiphenyl estrogen receptor binding affinity: An assessment of conformer flexibility. *Environ Toxicol Chem* 15:1945–1954.
36. Bradbury SP, Mekenyan OG, Ankley GT. 1998. The role of ligand flexibility in predicting biological activity: Structure-activity relationships for aryl hydrocarbon, estrogen and androgen receptor binding affinity. *Environ Toxicol Chem* 17:15–25.
37. Mekenyan O, Ivanov J, Karabunarliev S, Bradbury SP, Ankley GT, Karcher W. 1997. A computationally-based hazard identification algorithm that incorporates ligand flexibility. 1. Identification of potential androgen receptor ligands. *Environ Sci Technol* 31:3702–3711.
38. Mekenyan O, Nikolova N, Karabunarliev S, Bradbury S, Ankley G, Hansen B. 1999. New developments in a hazard identification algorithm for hormone receptor ligands: COREPA-C. *Quant Struct-Act Relat* 18:139–153.
39. Bradbury S, Kamenska V, Schmieder P, Ankley G, Mekenyan O. 2000. A computationally-based identification algorithm for estrogen receptor ligands. Part I. Predicting hERα binding affinity. *Toxicol Sci* 58:253–269.
40. Mekenyan O, Kamenska V, Schmieder P, Ankley G, Bradbury S. 2000. A computationally-based identification algorithm for estrogen receptor ligands. Part II. Evaluation of a hERα binding affinity model. *Toxicol Sci* 58:270–281.
41. Eliel E. 1993. *Chemistry in Three Dimensions*. Springer, Berlin, Germany.
42. Ivanov JM, Karabunarliev SH, Mekenyan OG. 1994. 3DGEN: A system for an exhaustive 3D molecular design. *J Chem Inf Comput Sci* 34:234–243.
43. Anstead GM, Wilson SR, Katzenellenbogen JA. 1989. 2-Arylindenes and 2-arylindenones: Molecular structures and considerations in the binding orientation of unsymmetrical nonsteroidal ligands to the estrogen receptor. *J Med Chem* 32:2163–2171.
44. Anstead GM, Carlson KE, Katzenellenbogen JA. 1997. The estradiol pharmacophore: Ligand structure-estrogen receptor binding affinity relationships and a model for the receptor binding site. *Steroids* 62:268–303.
45. Wiese TE, Brooks SC. 1994. Molecular modeling of steroidal estrogens: Novel conformations and their role in biological activity. *J Steroid Biochem Mol Biol* 50:61–73.
46. Ivanov JM, Mekenyan OG, Bradbury SP, Schüürmann G. 1998. A kinetic analysis of the conformational flexibility of steroids. *Quant Struct-Act Relat* 17:437–449.
47. Waller CL, Oprea TI, Chae K, Park H-K, Korach KS, Laws SC, Wiese TE, Kelce WR, Gray LE Jr. 1996. Ligand-based identification of environmental estrogens. *Chem Res Toxicol* 9:1240–1248.
48. Waller CL, Minor DL, McKinney JD. 1995. Using three-dimensional quantitative structure-activity relationships to examine estrogen receptor binding affinities of polychlorinated hydroxybiphenyls. *Environ Health Perspect* 103:702–707.
49. Kearsley SK, Smith GM. 1990. An alternative method for the alignment of molecular structure: Maximized electrostatic and steric overlap. *Tetrahedron Computed Methodology* 3:615–633.
50. Wang R, Gao Y, Lin L, Lai L. 1998. All orientation search and all-placement search in comparative molecular field analysis. *J Mol Modeling* 4:276–283.
51. Kuiper GG, Lemmon J, Carlsson B, Corton J, Safe S, van der Saag P, van der Burg B, Gustafsson JA. 1998. Interaction of estrogenic chemicals and phytoestrogens with estrogen receptor β. *Endocrinology* 139:4252–4263.
52. Goldstein RA, Katzenellenbogen JA, Luthey-Schulten ZA, Seielstad DA, Wolynes PG. 1993. Three-dimensional model for the hormone binding domains of steroid receptors. *Proc Natl Acad Sci USA* 90:9949–9953.
53. VanderKuur JA, Wiese TE, Brooks SC. 1993. Influence of estrogen structure on nuclear binding and progesterone receptor induction by the receptor complex. *Biochemistry* 32:7002–7008.
54. Lewis DFV, Parker MG, King RJB. 1995. Molecular modelling of the human estrogen receptor and ligand interactions based on site-directed mutagenesis and amino acid sequence homology. *J Steroid Biochem Mol Biol* 52:55–65.
55. Waller CL, Juma BW, Gray LE Jr, Kelce WR. 1996. Three-dimensional quantitative relationships for androgen receptor ligands. *Toxicol Appl Pharmacol* 137:219–227.
56. Wurtz JM, Bourguet W, Renaud JP, Vivat V, Chambon P, Moras D, Gronemeyer H. 1996. A canonical structure for the ligand-binding domain of nuclear receptors. *Nat Struct Biol* 3:87–94.
57. Ankley G, Mekenyan O, Kamenska V, Schmieder P, Bradbury S. 2002. Reactivity profiles of ligands of mammalian retinoic acid receptors: A preliminary COREPA analysis. *SAR/QSAR Environ Sci* 13:365–377.
58. Mekenyan O, Nikolova N, Schmieder P. 2003. Dynamic 3D QSAR techniques: Applications in toxicology. *J Mol Struct* 622: 147–165.
59. Hurst T. 1994. Flexible 3-D searching: The directed tweak technique. *J Chem Inf Comp Sci* 34:90–196.