January, 1998

The Role of Ligand Flexibility in Predicting Biological Activity: Structure–Activity Relationships for Aryl Hydrocarbon, Estrogen, and Androgen Receptor Binding Affinity

Steven P. Bradbury
Ovanes G. Mekenyan
Gerald T. Ankley

SETAC PRESS

Environmental Toxicology and Chemistry, Vol. 17, No. 1, pp. 15–25, 1998
Printed in the USA
0730-7268/98 $6.00 + .00

Annual Review

THE ROLE OF LIGAND FLEXIBILITY IN PREDICTING BIOLOGICAL ACTIVITY: STRUCTURE–ACTIVITY RELATIONSHIPS FOR ARYL HYDROCARBON, ESTROGEN, AND ANDROGEN RECEPTOR BINDING AFFINITY

STEVEN P. BRADBURY,*† OVANES G. MEKENYAN,‡ and GERALD T. ANKLEY†
†U.S. Environmental Protection Agency, National Health and Environmental Effects Research Laboratory, Mid-Continent Ecology Division, 6201 Congdon Boulevard, Duluth, Minnesota 55804
‡Higher Institute of Chemical Technology, University of Bourgas, 8010 Bourgas, Bulgaria

(*Received* 13 *May* 1997; *Accepted* 29 *July* 1997)

Abstract—Recent studies indicate that the potency and agonist or antagonist activity of steroid hormone ligands are dependent, in part, on ligand–receptor binding affinity as well as the conformation of the ligand–receptor complex. The binding of ligands to hormone receptors is thought to involve interactions by which shapes of both the receptor and ligand are modified in the formation of the ligand–receptor complex. As a consequence, it is essential to explore the significance of ligand flexibility in the development of screening-level structure–activity relationships. In this review, examples are provided of techniques used to generate and screen ligand conformers in the development of quantitative structure–activity relationships and active analogue search algorithms. The biological endpoint modeled was binding affinity of natural ligands and xenobiotics to the aryl hydrocarbon, estrogen, and androgen receptors. These approaches may be useful in future studies to evaluate relationships between ligand structure, receptor binding affinity, and, ultimately, transactivational events associated with receptor interactions with DNA response elements and associated proteins. An improved understanding of ligand–receptor interactions in the context of well-defined effector systems will enhance the development of credible predictive models that can be used to screen large sets of chemicals for potential agonist or antagonistic activity.

Keywords—Structure–activity relationships Receptor binding affinity Aryl hydrocarbon receptor Androgen receptor Estrogen receptor

INTRODUCTION

Recent reports that a wide variety of natural and synthetic compounds are capable of acting as hormonal agonists and antagonists serve as timely examples of the need to advance mechanistically based screening techniques to support human health and ecological risk assessments [1,2]. Structure–activity relationships (SARs) could serve as screening tools to help prioritize untested compounds that act as hormone agonists or antagonists and thereby trigger more intensive and costly empirical evaluations [2]. The need for SAR development arises, in part, from the reality that there are not sufficient resources to perform in vitro or in vivo assays to screen all of the existing 70,000 industrial organic chemicals (not including pharmaceuticals and pesticides) and the 1,500 to 2,000 new industrial chemicals submitted each year in the United States for evaluation under the Toxic Substances Control Act [3]. In response to the need for screening approaches, a wide variety of SARs have been developed to predict hormone binding affinity (e.g., see Waller [4]), under the assumption that information concerning ligand binding is a useful endpoint in the problem formulation and hazard identification stages of ecological and human health risk assessments, respectively.

In this review, brief background information on the development and use of SARs in ecotoxicology and steroid hormone receptor pharmacology is provided. This background section is followed by a summary of recent research efforts to develop exploratory quantitative structure–activity relationships (QSARs) and active analogue search algorithms, in which the influence of three-dimensional (3D) flexibility of ligands was assessed when predicting receptor binding affinity.

BACKGROUND

Conceptual role of SARs and ecotoxicology

In the field of environmental toxicology, SARs and QSARs have developed as scientifically credible tools for predicting the ecological effect and fate of chemicals when little or no empirical data are available. Coupled with the use of these models is the use of analogue selection techniques in which data associated with structurally similar chemicals are used to estimate risk levels of compounds for which no data are available [5]. The proper application and continued acceptance of these predictive toxicology techniques require that scientifically credible methods or models be used to systematically assign chemicals to appropriate QSARs or analogues. This fundamental process in the use of predictive techniques addresses a major area of uncertainty in prospective ecological risk assessments for chemical stressors [6,7].

Traditionally, the selection of structural analogues or QSARs has been based on the implicit assumption that compounds from the same chemical class should behave in a toxicologically similar manner. Although this working hypothesis seems reasonable, the delineation of chemical classes is problematic, and research com-

* To whom correspondence may be addressed (bradbury.steven@epamail.epa.gov).

This manuscript has been reviewed in accordance with U.S. Environmental Protection Agency policy. However, mention of trade names, specific products or approaches does not constitute endorsement by the agency.

pleted over the past several years has challenged the notion that typical chemical classification schemes necessarily reflect any similarity in mode of toxic action [8,9]. As a consequence, QSAR development and application have been evolving from a chemical class perspective to one that is more consistent with assumptions regarding modes of toxic action. Incorporation of this philosophy in QSAR development requires a fundamental understanding of both toxic mechanisms and the critical structural characteristics and properties that govern a chemical's action via a specific mechanism.

To develop a toxicologically credible approach for predicting potency and biological character from chemical structure requires the establishment of a knowledge base that contains ''training sets'' of chemicals whose modes of toxic action and potency are well defined. In developing these predictive models, it is critical that the mode of action domain for a specific application be clearly defined. In particular, it is essential that the domain be defined for exposure regime, biological model and endpoint, and a range of chemical properties [8,9]. Failure to adequately specify a knowledge base across the various discrete and continuous variables associated within each of these three conceptual dimensions of the mode of action domain can lead to toxicologically meaningless information and statistically inadequate mode of action prediction schemes and QSARs. The use of this conceptual approach to develop mode of action classification schemes, and associated QSARs, for the prediction of 96-h median lethal concentration values (LC50s) in the fathead minnow (*Pimephales promelas*) has been described previously [8–10].

SARs and the prediction of endocrine disruption

The need to precisely define the toxicological domain for SAR development and application to predict the potential for a xenobiotic to disrupt the endocrinology of an organism is essential. As discussed in a number of workshops and symposia, endocrine disruption can result from a wide array of mechanisms. For example, Kavlock et al. [1] defined an endocrine disrupter as ''an exogenous agent that interferes with the production, release, transport, metabolism, binding, action, or elimination of natural hormones in the body responsible for the maintenance of homeostasis and regulation of developmental processes.'' This definition incorporates a bewildering number of possible mechanisms of endocrine disruption that must be experimentally resolved if there will be any hope of developing credible predictive toxicology models. Lists of putative endocrine-disrupting chemicals have been developed that, even for a given response (e.g., estrogenicity), represent an impressive range of structural characteristics that have led some to claim that there is no apparent chemical similarity underlying these effects. The data upon which some of these claims are based need to be carefully examined to ensure that the conclusions are not derived from attempts to discern chemical similarity among chemicals whose mechanisms of action are dissimilar (i.e., one would not expect similarity in chemical structure among xenobiotics with different modes of toxic action). By carefully defining mechanisms of endocrine disruption and potency in terms of specific endpoints and experimental models (see ''Overview of Steroid Hormone Receptor Pharmacology''), the basic assumption behind SAR development, which holds that structurally similar chemicals should have biologically similar activity, should be operative. A significant challenge, however, is to develop the toxicological knowledge bases from which biological similarity can be defined and chemical similarity quantified.

In attempting to apply SAR techniques to predict biological activity, it will also be critical to differentiate those processes and endpoints that can reasonably be expected to be related to chemical structure from those that cannot. For example, it is highly likely that binding affinity for specific receptors and associated transactivational endpoints can be related to chemical structure. Similarly, predictions concerning the binding of xenobiotics to transport proteins (e.g., Rickenbacher et al. [11]) and aspects of enzyme-mediated xeniobiotic metabolism (e.g., de Groot et al. [12]) are not unreasonable to anticipate. However, the prediction of endpoints associated with increasingly complex processes that are dependent on cascading events that incorporate tissue and organismal-level physiology are far less likely to be predictable from chemical structure alone. Predictions of in vivo dose–response relationships and interspecies responses will invariably require the use of physiologically based toxicokinetic and biologically based dose–response models in addition to mode of action and potency predictions derived from SARs.

Overview of steroid hormone receptor pharmacology

Katzenellenbogen et al. [13] have proposed that steroid potency and agonist versus antagonist activity within different in vivo and in vitro models can be viewed in terms of ligand-based, receptor-based, and effector-site-based mechanisms of selectivity. Ligand-based selectivity concerns the toxicokinetics and metabolism of hormones, which may contribute to differential responses at the tissue or cellular level. Receptor-based selectivity concerns the differential composition of receptors (concentration, subtypes, isoforms, and variants) in different tissues and cells, which can be responsible for tissue-selective or cell-selective action elicited by the same hormone. Effector-based selectivity is invoked to resolve those cases in which differential activity in tissues or cells is observed when there is no apparent difference in ligand toxicokinetics and a single receptor is involved. In these instances, even though the hormone, tissue, cell, and gene type are the same, differential activity can be elicited as a result of differences in the aggregate of coactivators, corepressors, and transcription factors (effector sites) with which the ligand–receptor complex interact. Consequently, Katzenellenbogen et al. [13] proposed that the pharmacology of nuclear hormone receptors should be viewed as a tripartite system involving the ligand, receptor, and effector. Within this model it is generalized that the potency of a ligand is determined through its interaction with the receptor (receptor binding), whereas agonistic or antagonistic activity is determined through the interaction of the receptor–ligand complex and the effector components. Varying activity is related to different conformational shapes of the receptor–ligand complex, which affects the nature of receptor–effector coupling. Thus, the activity, and to some degree the potency, of a ligand must be assigned within the context of a specific response–effector interaction [13].

Both the bipartite (i.e., a model that does not invoke the coupling of the receptor to the effector) and tripartite views of receptor pharmacology require that the ligand cause a conformational change in the receptor that reflects the shape and structure of the ligand [13]. In the tripartite model, ligands of different structure acting on the same receptor should give rise to conformationally different ligand–receptor complexes with the potential to differentially react with the suite of effector

systems, ultimately resulting in different patterns of gene regulation and expression [13]. The assumption that receptors undergo a progressive and substantial conformational change upon binding to ligands was recently confirmed [14–17] and suggests that receptor–ligand binding reflects an induced fit, where the receptor adapts to the shape of the ligand, whereas the ligand's conformation is altered to that of the ligand binding pocket [13,15].

SARs FOR RECEPTOR BINDING AFFINITY: EXPLORING THE ROLE OF LIGAND FLEXIBILITY

With conventional 3D SAR methods, the molecular structure of each chemical under investigation is represented by a single lowest energy conformer, as defined by quantum chemical or force-field methods. However, the most stable conformations may be less likely than other conformations to interact with a solvent or macromolecule [18]. To systematically evaluate conformer flexibility within the context of relevant biological environments and reactions (e.g., partitioning within a complex solution, substrate–receptor complex formation), a dynamic SAR approach [19] was incorporated in the studies summarized in this review. This approach assumes that in the complex reaction environments of biological systems, a molecule exists as a variety of conformers, with solvation and binding interactions capable of compensating for energy differences among the conformations. Biological activity is subsequently modeled as an electronic property for specifically selected conformers (or a set of conformers) rather than as a property derived from the single lowest-energy gas-phase conformer. Exploratory investigations using distributions of energetically reasonable ligand conformations in SARs for receptor binding affinity were undertaken because, as described above, it is likely that hormone receptor–ligand interactions reflect a continuum of conformational character, which in turn may be related to activity and potency.

The dynamic SAR approach: Conformer generation and selection

The dynamic SAR method combines an exhaustive conformer generation routine [20] with conformer screening algorithms that can be adapted to explore relationships between specific electronic properties and the biological activity under investigation [19]. The conformer generation technique is described by Ivanov et al. [20], which the reader is encouraged to consult for a detailed presentation of the approach. Essentially, the technique is a combinatorial procedure that initiates from molecular topology and generates all conformers in the context of steric constraints and expert rules. A unique aspect of the approach is that it incorporates the conformational flexibility of saturated cyclic molecular fragments, as opposed to other conformational analysis techniques, which explore conformational space formed by rotations around acyclic single bonds only [20].

Various screening paradigms can be used to remove conformers from subsequent analyses, including an initial screen to eliminate energetically unreasonable conformations. Conformers associated with the prevailing minimum and maximum ranges for the screening parameters are selected and organized in correlation samples. Within each sample, conformers derived from the same compound are considered as distinct structural representatives of the chemical (i.e., each conformer has different electronic parameters), but each conformer is associated with the toxicological value (dependent variable) for the ''parent'' two-dimensional (2D) compound. The identification of proposed ''active'' conformers is based on an evaluation of the regression statistics associated with QSARs derived using the different conformation samples. Finally, recent dynamic QSAR studies on the toxicity of unsaturated alcohols [19], semicarbazides [21], and α-terthienyls [22] indicated that models based on nonoptimized conformer geometries are not qualitatively different from those based on optimized structures when nonstrained chemicals are analyzed.

In developing and interpreting SAR models, there is always the challenge of assessing whether statistically significant correlations are based solely on chance. Obviously, the possibility of deriving statistically significant correlations by chance increases as the number of descriptors in the initial parameter pool (data matrix) used for deriving relationships increases. Although it is impossible to ever categorically declare that a SAR is not based on chance, regardless of the limited number of descriptors evaluated, the likelihood that a relationship is reasonable increases if the descriptors chosen for study can be based on a plausible hypothesis or set of hypotheses developed from a mechanistic understanding of the processes under consideration. For example, in the studies summarized here, the possibility of generating statistically robust but biologically meaningless correlations was minimized by restricting the choice of steric, physicochemical, and electronic descriptors to those hypothesized to be important for ligand binding to the aryl hydrocarbon receptor (AhR), the estrogen receptor (ER), and the androgen receptor (AR), based on previously published research. Another potential means of generating chance correlations concerns the number and nature of the conformer screens. As described in the subsequent sections, screens were primarily based on descriptors of steric bulk, interatomic distances, and electron-donating or -accepting capabilities because these attributes had previously been reported to be related to ligand binding to the AhR, ER, and AR. Thus, it is important to note that initial data sets based on all possible conformers and all possible descriptors were not used. Rather, specific conformer sets were evaluated individually with SARs derived from a restricted set of mechanistically plausible descriptors.

QSARs for AhR binding affinity

Our first evaluation of the dynamic QSAR method in terms of receptor–ligand interactions focused on binding of halogenated aromatic compounds to the AhR [23]. The relative affinity for the AhR of individual congeners of polychlorinated biphenyls (PCBs), polychlorinated dibenzofurans (PCDFs), and polychlorinated dibenzo-*p*-dioxins (PCDDs) is strongly correlated with the induction of hepatic cytochrome P4501A1-mediated activities and a variety of toxic responses, including thymic atrophy, body weight loss, immunotoxicity, and acute mortality [24–31]. Hence, accurate prediction of the affinity of chemicals for the AhR is key to assessing potential toxicity.

Numerous attempts to develop both SARs and QSARs for predicting AhR binding affinity have been made. The most simple of these consists of a model that presumes the structural characteristics necessary for binding to the receptor are planarity and the ability to occupy a 3- × 10-Å rectangle [24]. More sophisticated modeling efforts have focused on quantitative analysis of the effects of ligand substitutents on receptor binding (e.g., in charge–transfer interactions) and dispersion interactions between the receptor and aromatic xenobiotics [32–35]. However, virtually all modeling of binding of halogenated aromatic chemicals to the AhR has focused, implicitly or explicitly, on minimum energy conformers. Although this

Table 1. Observed aryl hydrocarbon receptor binding affinities for 14 polychlorinated biphenyls and key electronic descriptors associated with the most planar conformers of each chemical

No.	Compound	Parameters: Observed log(1/EC50)[a]	E_{LUMO}[b] (eV) Eqn. 1	$E_{(HOMO-LUMO)}$[b] (eV) Eqn. 2
1	3,3′,4,4′-Tetrachlorobiphenyl	6.15	0.906	8.225
2	2,3,4,4′-Tetrachlorobiphenyl	4.55	0.816	8.328
3	3,3′,4,4′,5-Pentachlorobiphenyl	6.89	1.020	8.181
4	2′,3,3′,4′,5-Pentachlorobiphenyl	4.85	0.935	8.286
5	2,3,3′,4,4′-Pentachlorobiphenyl	5.37	0.925	8.286
6	2,3′,4,4′,5-Pentachlorobiphenyl	5.04	0.951	8.240
7	2,3,4,4′,5-Pentachlorobiphenyl	5.39	0.946	8.250
8	2,3,3′,4,4′5-Hexachlorobiphenyl	5.15	1.045	8.203
9	2,3′,4,4′,5,5′-Hexachlorobiphenyl	4.80	1.052	8.208
10	2,3,3′,4,4′,5-Hexachlorobiphenyl	5.33	1.029	8.243
11	2,2′,4,4′-Tetrachlorobiphenyl	3.89	0.812	8.319
12	2,2′,4,4′5,5′-Hexachlorobiphenyl	4.10	0.807	8.164
13	2,3,4,5-Tetrachlorobiphenyl	3.85	0.807	8.394
14	2,3′,4,4′,5′,6-Hexachlorobiphenyl	4.00	1.040	8.199

[a] Data from Safe et al. [27] and Bandiera et al. [25,26].
[b] E_{LUMO} and $E_{HOMO-LUMO}$ values for the most planar, optimized conformers of each chemical.

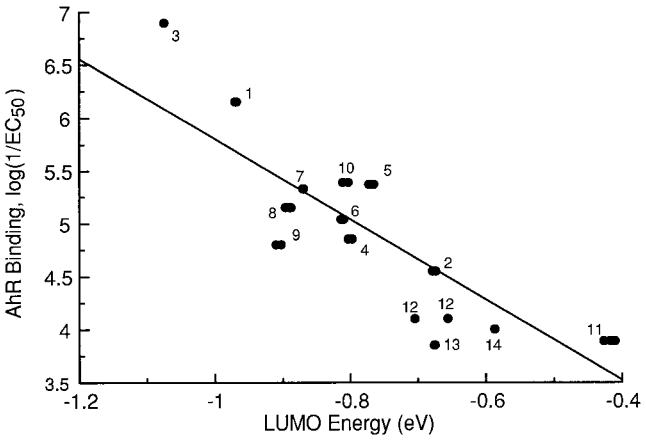

Fig. 1. Variation of observed arylhydrocarbon receptor (AhR) binding affinity versus energy of the lowest unoccupied molecular orbital (LUMO) (E_{LUMO}) for the most planar, optimized conformers of 14 polychlorinated biphenyl congeners (see Table 1 for compound names; modified from Mekenyan et al. [23]). EC50 = concentration of test chemical that reduces specific binding of 2,3,7,8-[^{3}H] tetrachlorodibenzo-*p*-dioxin to 50%.

likely is not of great concern for structurally rigid molecules such as PCDFs and PCDDs, in the case of the relatively flexible PCBs, we hypothesized that consideration of specific conformers as the active conformers might result in QSARs superior to those that previously had been developed [23].

Our study focused on two groups of structurally dissimilar PCBs and sets of PCDFs and PCDDs, first as discreet groups, then as a combined set of all the chemicals [23]. For illustrative purposes, here we describe the modeling results obtained with only one of the sets of PCBs. The AhR binding data for 14 PCB congeners with varying chlorination patterns on the two phenyl rings were obtained from Safe et al. [27] and Bandiera et al. [25,26]. The test chemicals and AhR binding data, expressed as the log(1/EC50), are shown in Table 1. Here, the EC50 is defined as the concentration of the test chemical necessary to reduce specific binding of 2,3,7,8-[^{3}H]tetrachlorodibenzo-*p*-dioxin to 50% of the maximal value obtained in the absence of the competitor. Affinity of the PCBs for the AhR spans nearly three orders of magnitude and includes very potent agonists (e.g., 3,3′,4,4′,5-pentachlorobiphenyl) as well as relatively nontoxic congeners (e.g., 2,2′,4,4′-tetrachlorobiphenyl). Conformers for the PCBs were generated as described previously; for the 14 congeners we found 102 conformations using a torsional resolution of 30° and a nonbonded cutoff of 1.8 Å. Screening and computation of molecular descriptors of interest were performed after geometry optimization or by direct single-point (1SCF) calculations of the nonoptimized confirmations.

Because of the demonstrated importance of planarity in some binding interactions with the AhR, we used a steric parameter for planarity (P1) to screen both the optimized and nonoptimized PCB conformers before modeling (for the definition of P1, see Mekenyan et al. [23]). This planarity parameter was not intended to be used as an independent variable in developing QSARs but as a basis for screening conformers with respect to their possibility of assuming a planar or nearly planar conformation. Following identification of the most planar conformations, structures were further screened to retain only those that were within 20 kcal/mol of the lowest-energy conformation for a given congener. For the studies described in this review [23,36,37], ΔH_f^0 values for the energy minima for both the unoptimized and optimized conformers were typically within 5 to 15 kcal/mol, which is presumed to be within the free energy of binding of ligands to nuclear steroid receptors (e.g., −12.1 kcal/mol for the binding of E_2 to the ER; see Wiese and Brooks [38]).

In the complete analysis, described by Mekenyan et al. [23], we generated several robust mono- and biparametric models for this set of PCBs using different electronic and steric descriptors representative of different aspects of binding of aromatic xenobiotics to the AhR [23]. For the sake of brevity we present only the monoparametric models, which highlight the necessity of considering conformational flexibility of ligands in QSAR derivation. In selecting regressors for generating these models, we considered that it has been shown that chemicals that have greater ability to accept electron density through charge–transfer interaction should bind to the AhR with greater affinity than those with lower electron acceptor properties [33–35,39]. We hypothesized, therefore, that stronger electron acceptors would have lower energy of unoccupied and occupied frontier orbitals (E_{LUMO} and E_{HOMO}, respectively) and a lower energy difference in these frontier orbitals ($E_{HOMO-LUMO}$) that can be related to molecular reactivity [40,41]. Thus, consistent with mechanistic assumptions in previous studies, two of the key electronic descriptors calculated for the QSAR modeling effort were E_{LUMO} and $E_{HOMO-LUMO}$ (Table 1).

In a reference analysis, we evaluated models with the optimized, lowest-energy conformers of the 14 PCBs (i.e., those geometries that typically would be used for conventional QSAR modeling); coefficients of determination (r^2) for these models with E_{LUMO} and $E_{HOMO-LUMO}$ were quite low (0.38 and 0.41, respectively). As hypothesized, marked improvement in the correlation between AhR binding and E_{LUMO} (Eqn. 1; Fig. 1) and $E_{HOMO-LUMO}$ (Eqn. 2) was noted when P1 was used to select the most planar, optimized conformers:

$$\log(1/EC50) = 2.01\ (\pm 0.32) - 3.79\ (\pm 0.41)\ E_{LUMO}$$

$$n = 36\ (14), \quad r^2 = 0.715, \quad s^2 = 0.152, \quad F = 85.11 \qquad (1)$$

Table 2. Observed estrogen receptor binding affinities for 12 hydroxylated polychlorinated biphenyls and key electronic descriptors associated with the most polarizable conformers of each chemical

Compound	Observed pEC50 (Meq)[a]	Descriptors: LogP Eqn. 5	AF[b] (f^{HOMO}) (a.u.) Eqn. 5	D^2 Eqn. 6	$E_{HOMO\text{-}LUMO}$[b] (eV) Eqn. 6
2,4,6-Trichloro-4′-biphenylol	−1.63	5.22	0.300	4.007	8.866
2,3,4,5-Tetrachloro-4′-biphenylol	−1.98	5.81	0.249	4.141	8.889
2-Chloro-4,4′-biphenyldiol	−1.95	3.37	0.300	3.995	8.652
2,6-Dichloro-4′-biphenylol	−2.59	4.63	0.300	3.795	8.995
2,5-Dichloro-4′-biphenylol	−2.70	4.63	0.300	0.899	8.910
3,4,5-Trichloro-4-biphenyldiol	−3.00	5.22	0.140	4.146	8.394
3,3′,5,5′-Tetrachloro-4,4′-biphenydiol	−3.13	5.14	0.138	4.339	8.298
2-Chloro-4-biphenydiol	−3.40	4.04	0.158	3.697	8.844
4′-Chloro-4-biphenylol	−3.59	4.04	0.167	3.979	8.446
2,3,5,6-Tetrachloro-4,4′-biphenyldiol	−3.70	3.33	0.257	4.136	8.504
4,4′-Biphenydiol	−4.00	2.66	0.156	3.979	8.429
4-Biphenylol	−4.00	3.33	0.190	3.683	8.610

[a] Negative log of the EC50 for displacing [^{3}H] estradiol from mouse cytosolic uterine estrogen receptor (observed data from Korach et al. [52]).
[b] Mean values of AF(f^{HOMO}) and $E_{HOMO\text{-}LUMO}$ are provided for the most polarizable conformer sets. Note that AF(f^{HOMO}) values are derived from unoptimized conformers (Eqn. 5).

$$\log(1/EC50) = 25.8\ (\pm 2.23) - 2.48\ (\pm 0.26)\ E_{HOMO\text{-}LUMO}$$
$$n = 36\ (14), \quad r^2 = 0.721, \quad s^2 = 0.149, \quad F = 87.85 \tag{2}$$

where n is the cardinality of the correlation sample (i.e., the number of conformers associated with the PCBs), r^2 is the variance, s^2 is the standard error of the estimate, and F is the Fisher criterion.

We also modeled receptor binding of the PCBs using nonoptimized structures to generate acceptable models based on both E_{LUMO} and $E_{HOMO\text{-}LUMO}$:

$$\log(1/EC50) = 0.39\ (\pm 0.65) - 6.14\ (\pm 0.84)\ E_{LUMO}$$
$$n = 17\ (14), \quad r^2 = 0.781, \quad s^2 = 0.167, \quad F = 53.39 \tag{3}$$

$$\log(1/EC50) = 38.4\ (\pm 4.81) - 3.89\ (\pm 0.56)\ E_{HOMO\text{-}LUMO}$$
$$n = 17\ (14), \quad r^2 = 0.762, \quad s^2 = 0.182, \quad F = 48.04 \tag{4}$$

This particular result is noteworthy in that the correlations derived using the nonoptimized, but energetically reasonable, planar PCB conformations yielded models comparable to those obtained using optimized conformers, which requires a more computationally intensive approach.

QSARs for predicting ER binding affinity

The ER has demonstrated considerable structural tolerance for ligands [42–46]. Besides the natural hormone (estradiol, E_2) and related steroidal analogues [38,42,47], many different structural classes of nonsteroidal compounds have high ER binding affinity, such as bibenzyls, stilbenes, triarylethylenes, phenylindoles, phenylindenes, coumarins, and lactones [46]. The development of mechanistically based SARs, however, is hampered by a lack of knowledge concerning the 3D structure of the ER. In the absence of a well-defined 3D structure for the ER, a determination of the steric and electronic requirements for binding of E_2 derivatives is starting to provide the means to assess the receptor's characteristics [42]. However, the influence of an aqueous environment on ligand and ER geometry and the potential conformational flexibility of ER ligands [38] create challenges to mapping the estrogen binding domain. Ligands can also induce alterations in the tertiary structure of the ER binding complex, which in turn may alter binding to estrogen hormone responsive elements and subsequent transactivational events [13,48–51].

As discussed previously, PCB conformations were found to be critical in evaluating SARs for their binding to the AhR. To further study the role of PCB conformer flexibility on biological activity, the multiple conformer SAR method was used to model ER binding affinity of a series of hydroxylated PCBs (PCHBs) [36], previously investigated by Korach et al. [52] and Waller et al. [45]. Based on experimental and theoretical findings for the involvement of hydrogen bonding [42] and a negative isopotential surrounding the E_2 A ring in ER binding [38,47,50,51] and the assumption that there are cationic sites (protonated amino acid residues) associated with ER ligand binding domain [42,53,54], we hypothesized that PCHBs with greater binding affinity should have higher electron-donating abilities associated with their corresponding phenolic ring. Such chemicals should have higher donor and lower acceptor superdelocalizabilities and frontier charges on the HOMO and LUMO orbitals, respectively, as well as larger negative charges on the phenolic oxygen and the aromatic carbons in the A ring.

The 12 PCHBs that were studied are listed in Table 2. Cytosolic uterine ER binding affinities from ovariectomized mice, as reported by Korach et al. [52], were expressed as the concentration (in molar equivalents) of competitor required to displace 50% of ER-bound [^{3}H] estradiol (EC50). The observed binding affinity values, expressed as the negative log of the EC50s (pEC50s), are presented in Table 2. Seventy-two nonoptimized conformations were derived for the 12 PCHBs. During optimization, some of the PCHB conformations quenched into the same energy minima, which reduced the number of conformers to 29. The ΔH_f^0 values for the energy minima for these unoptimized and optimized conformers were within the free energy of binding for E_2 to the ER (−12.1 kcal/mol; see Wiese and Brooks [38]).

Global electronic parameters used in screening conformers or assessed as potential regressors in the QSAR analyses in-

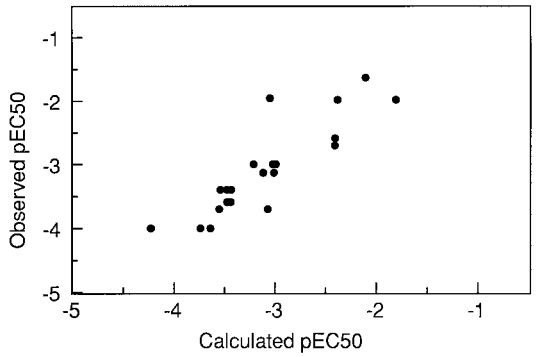

Fig. 2. Plot of observed versus calculated estrogen receptor binding affinities for the most polarizable hydroxylated polychlorinated biphenyl nonoptimized congeners (see Table 2 and Eq. 5 in the text; modified from Bradbury et al. [36]). pEC50 = negative log of chemical concentration required to displace 50% of E_2 from the ER.

cluded E_{HOMO}, E_{LUMO}, $E_{HOMO\text{-}LUMO}$, and volume polarizability (VolP). Volume polarizability (VolP < 0) is defined as a sum of atomic self-polarizabilities and, thus, the averaged ability of a compound to change electron density at its atoms during chemical interactions [55,56]. Lower absolute values of VolP reflect higher-charge localizations and more polarizable (less lipophilic) molecules [55]. Atomic charges (q_i), bond orders (p_{ij}), frontier atomic charges (f_i^{HOMO} and f_i^{LUMO}), and acceptor and donor superdelocalizability indices (S_i^E and S_i^N), and frontier donor and acceptor superdelocalizability indices, as well as atomic self-polarizabilities (π_i), were calculated as local electronic indices (I denotes a specific atom in a molecule; e.g., with the local electronic descriptors of S_{O3}^E and S_3^E, O3 and 3 correspond to the phenolic oxygen and its neighboring aromatic carbon atom in the E_2 A ring analogue). To investigate the hypothesized impact of the phenolic A ring on ER binding, the generation of the above-mentioned local reactivity indices was restricted to this fragment. In addition, the averaged sums of the local electronic indices (lei) over the aromatic fragment atoms were calculated. These fragment indices were denoted as AF(lei).

Using the dynamic SAR method to isolate the most active PCHB conformers, the most statistically robust models (confidence level of 99%) were obtained when conformers having the lowest absolute values for VolP (more polarizable 3D structures) were incorporated in the correlation samples. In general, the most polarizable unoptimized PCHBs conformers tended to have lower conjugation and lower planarity than the lowest-energy gas-phase conformers. The apparent greater affinity of the more polarizable conformers to the ER under the in vitro conditions of the assay may be reflective of partitioning and distribution characteristics of the lipophilic PCHBs in the aqueous buffer [36].

Regressions for the most polarizable nonoptimized PCHB conformers (i.e., those conformers with the lowest absolute VolP values) can be summarized by Equation 5 (see Fig. 2 and Table 2):

$$\text{pEC50} = -6.46\ (\pm 0.32) + 0.51\ (\pm 0.063)\ \log \text{P} + 5.65\ (\pm 1.04)\ \text{AF}(f^{HOMO})$$

$$n = 30\ (12), \quad r^2 = 0.79, \quad s^2 = 0.11, \quad F = 51.58 \tag{5}$$

Significant regressions also were obtained when log P was combined with $+ f_3^{HOMO}$ ($r^2 = 0.79$, $s^2 = 0.12$), $+ f_{O3}^{HOMO}$ ($r^2 = 0.78$, $s^2 = 0.12$), $- f_3^{LUMO}$ ($r^2 = 0.75$, $s^2 = 0.13$), $- f_{O3}^{LUMO}$ ($r^2 = 0.74$, $s^2 = 0.13$); $-\text{AF}(f^{LUMO})$ ($r^2 = 0.74$, $s^2 = 0.13$), and $-\text{AF}(S^N)$ ($r^2 = 0.67$, $s^2 = 0.17$), where a negatively correlated parameter is denoted by $-$, and $+$ indicates a positively correlated parameter. As can be seen in these analyses, for the most polarizable, nonoptimized conformers, the frontier charges on the HOMO and LUMO orbitals, related to the ability of the phenolic fragment to exchange or donate electrons, appeared in most of the regressions. To provide a point of reference, the best regression derived with the lowest-energy gas-phase PCHB conformers was based on log P and S_{O3}^E and had an r^2 of 0.60.

Optimization of the PCHB conformers resulted in a set of 3D structures that were inherently more polarizable than the original set of nonoptimized compounds [36]. Thus, after geometry optimization, the range of VolP values for the conformers of a given PCHB congener tended to shift toward less negative values. Consistent with this observation, the differences in specific torsional angles for the most polarizable optimized conformers compared with the lowest-energy gas-phase conformers were not as large as those noted with the unoptimized conformers. However, a general tendency of increased polarizability and decreased planarity was noted when the lowest-energy gas-phase and most polarizable optimized conformers were examined.

Based on the results obtained with nonoptimized conformers, the optimized PCHB conformers were subsequently screened to select those 3D geometries associated with the most polarizable structures. The regression with the highest correlation for the polarizable subset of optimized PCHB conformers is summarized in Equation 6 (see Table 2):

$$\text{pEC50} = -43.6\ (\pm 6.2) + 3.13\ (\pm 0.58)\ D^2 + 3.25\ (\pm 0.52)\ E_{HOMO\text{-}LUMO}$$

$$n = 18\ (12), \quad r^2 = 0.74, \quad s^2 = 0.16, \quad F = 21.42 \tag{6}$$

where D^2 is the Balaban mean square distance topological index. This index reflects structural characteristics such as molecular size and branching; high values of D^2 correspond to large molecular size and a low degree of branching [57]. D^2 was correlated with log P, which is consistent with commonly acknowledged relationships between size, degree of branching, and hydrophobicity. With this data set, it appears that $E_{HOMO\text{-}LUMO}$ is representing an integral descriptor associated with the electron-donating character of the conformers (see Bradbury et al. [36]). These findings also suggest that comparable results could be obtained when using optimized or unoptimized conformations of PCHBs to model ER binding affinity.

The trends of reactivity parameters in the QSARs derived in this study supported the hypothesis that PCHBs with greater binding affinity should have higher electron-donating abilities associated with their corresponding phenolic ring. These results are consistent with an electrostatic modeling study of E_2 and related derivatives by VanderKuur et al. [50] that suggested a requirement for a distinct negative isopotential pattern around C2 and C3 oxygens of estratrienes and the C3 oxygen of androstane, and with a hypothesis for the existence of electrophilic sites (protonated amino acid residues) in the ER receptor cavity [53,54]. In keeping with these conclusions, Anstead et al. [42] has proposed that residues in the hydrophobic ligand binding site likely interacts with the A ring of E_2 through

predominately hydrogen bonding and weak polar interactions with a slightly positively charged receptor residue.

An active analogue algorithm based on conformational flexibility: AR binding affinity

Analogue selection techniques are used when screening untested industrial organic chemicals for possible testing [5]. In addition, analogue selection, or chemical similarity, techniques aid in selection of mode-of-action-specific QSARs [9]. Chemical similarity assessment approaches, if linked to a toxicodynamically credible knowledge base, help reduce effect characterization uncertainties that can be propagated through the selection of incorrect QSARs [6,7,9]. Because QSAR models required to predict biological activity associated with hormone receptors invariably require 3D analyses and, as a result, intensive computational resources, it is also necessary to devise computationally efficient 2D and 3D similarity assessment techniques to initially identify those chemicals with a low probability of being highly active. With such techniques, large numbers of compounds can be screened to establish smaller data sets that can then be subjected to more rigorous and computationally intensive QSAR analyses.

Typical approaches to quantify 3D similarity in the context of ligand–receptor interactions encompass pharmacophore search methods and receptor-site mapping. The pharmacophore search method attempts to quantify the essential 3D arrangement of functional groups that a molecule must possess to be recognized by the receptor of concern [58]. Field-based approaches, such as comparative molecular field analysis (CoMFA) [59,60] represent the lowest-energy gas-phase shape of the pharmacophore implicitly (e.g., steric and electrostatic interaction energies derived from a probe atom). In a related search method, the active analogue approach developed by Marshall [61] can be used as an implicit description of the ligand-binding site. Receptor-site models are more specific than pharmacophore approaches in that they rely on assumptions of the complementary relationship between the shape and properties of the receptor and active ligands [62]. Recently, receptor surface models have also been developed to explicitly represent the receptor's shape [58,63].

Pharmacophore search methods and receptor-site mapping face significant challenges, including selection of appropriate conformations and obtaining appropriate template alignments. A number of good techniques for superimposing molecules [64–67] are available, but developing a robust alignment model is not trivial. Typically, hundreds of alignments are explored to reach an optimal outcome, which must be carefully evaluated and explained in the context of a presumed set of ligand–receptor interaction mechanisms. As discussed previously, the use of the lowest-energy conformers to model hormone–ligand interactions may be inappropriate, which is also of direct concern to pharmacophore search/receptor-site mapping exercises [37]. In an attempt to address the issue of conformational flexibility, Prendergast et al. [68] described an approach to identify specific conformers of ligands acting as antagonists to angiotensin II by assessing interatomic distances. In their study, all geometrically reasonable conformers were assessed; however, conformational energies were not evaluated, and energy minimization was not performed during the search. In a sense, the methodology developed by Prendergast et al. [68] can be viewed as an augmented version of the active analogue approach because it accounts for conformational flexibility and eliminates the necessity of conformer alignment.

Recently, we reported a technique to further generalize the use of multiple conformers in an active analogue approach [37]. The approach circumvents the problems of conformer alignment and selection, and initial assumptions concerning specific atoms or fragments in a pharmacophore (e.g., identification and assignment of A and/or D rings in steroid derivatives) are not an obligatory step. In addition, interatomic distances are considered as only one of many components of molecular stereoelectronic structure. The principle steps of the algorithm can be summarized as follows.

Step 1: Definition of the training set of chemicals. A defined subset of chemicals in the reaction series under investigation is selected as the training set. The training set can include either the most or least active chemicals, as defined by a user-imposed threshold of biological activity. This initial step establishes the extent of biological similarity among the chemicals from which stereoelectronic similarity will be discerned in the subsequent steps of the algorithm. Conformations for each of the compounds are then generated; note that the term ''molecule'' in steps 2 and 3 denotes a set of energetically reasonable conformations for each compound.

Step 2: Evaluation of stereoelectronic parameters hypothesized to be associated with biologically similar compounds. A restricted set of parameters, hypothesized to be associated with biological activity, are evaluated on the basis of the normalized sum of dynamic similarity indices between each pair of molecules in the training set (see Mekenyan et al. [69] for a detailed description of the similarity indices used). The stereoelectronic parameters that provide the maximal measure of similarity among the chemicals in the training set are assumed to be most closely associated with the activity under consideration and are used in the subsequent step of the algorithm. Again, to minimize the potential for creating chance similarity relationships, or similarity relationships that are unique to the training set, the algorithm does not evaluate all possible stereoelectronic parameters but only those with a mechanistically plausible basis.

Step 3: Recognition of the common reactivity pattern. For each stereoelectronic parameter identified in step 2, the conformer distributions of the compounds from the training set are superimposed, and the parameter ranges common for conformers from all of the compounds are identified. The distribution intersections (i.e., commonly populated parameter ranges) can be either discrete or continuous. The collection of common stereoelectronic parameter ranges define the common reactivity pattern. In the context of local stereoelectronic ranges, it is essential to subsequently elucidate the associated specific atom types, fragments, and interatomic distances.

To illustrate the algorithm, the stereoelectronic requirements associated with the binding of a diverse set of ligands to the AR were defined. The AR ligands examined in this study consisted of nine steroids, two synthetic steroids, and 17 nonsteroids (Fig. 3) Androgen receptor binding affinities were obtained from Kelce et al. [70] and Waller et al. [71] and are based on a competitive binding assay using [^{3}H]RI881 (a radiolabled synthetic androgen; see Kelce et al. [70]). The pK_i values are also provided in Figure 3.

After establishing training sets (step 1) based on the four, six, and eight most active ligands and the eight least active ligands as defined by pK_i values (see Fig. 3), pairwise similarity was assessed between chemicals within each training set (step 2). This analysis suggested that the most active and inactive AR ligands have the greatest similarity to members

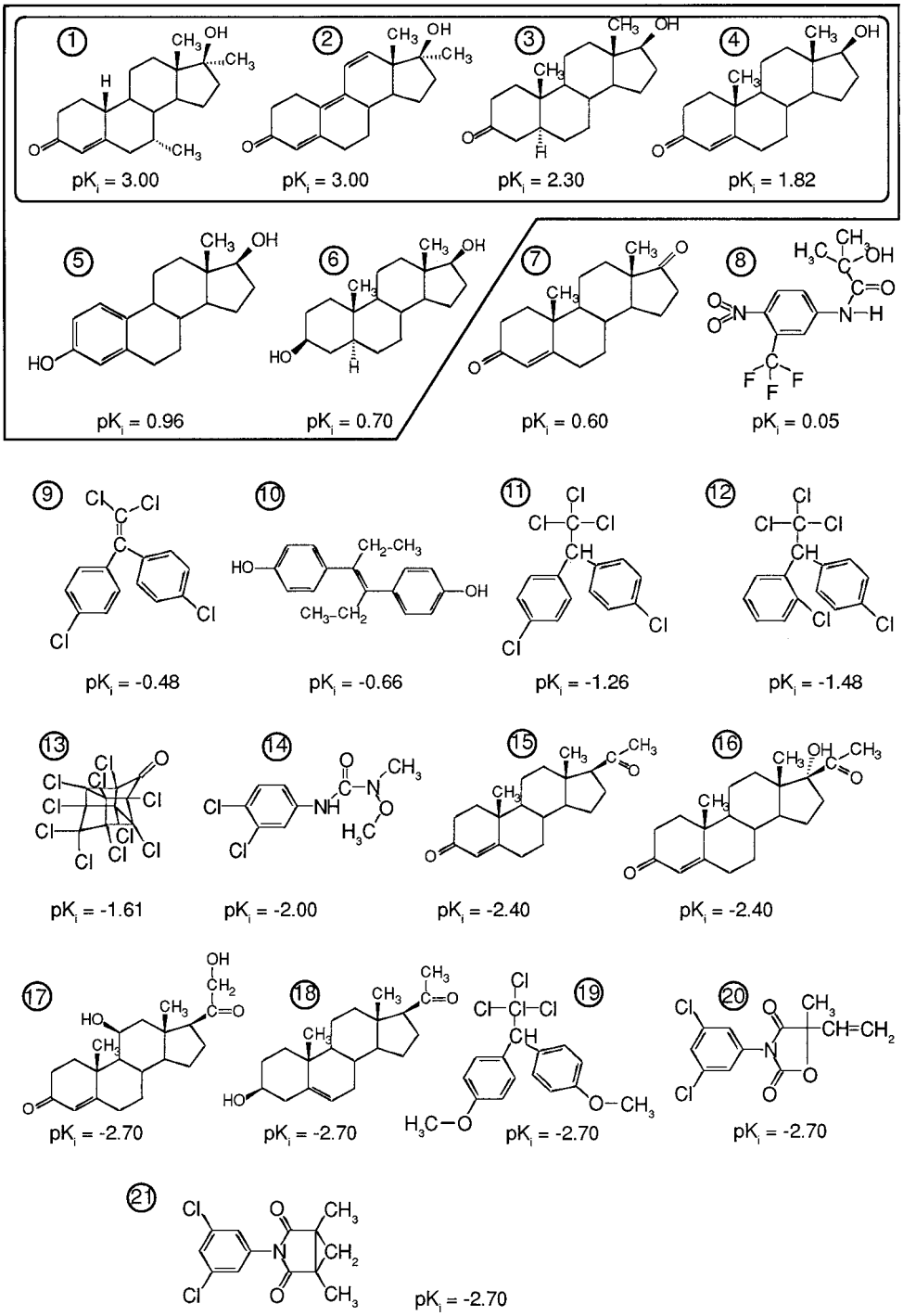

Fig. 3. Structures of ligands, with experimental binding affinities [71], used as training and validation sets to elucidate stereoelectronic requirements for androgen receptor binding affinity. The structures within the boxes were used to define the common reactivity patterns associated with andogren receptor ligands with pK_i values greater than or equal to 1.82 or 0.70 (modified from Mekenyan et al. [37]).

within each respective training subset in terms of distributions of charges (q_i), steric distances (d_{ij}), donor and acceptor delocalizabilities (S_i^E, S_i^N), frontier charges on HOMO (f_i^{HOMO}) and LUMO (f_i^{LUMO}), and atom polarizabilities (π_i).

Previous modeling studies [e.g., 36,42,45,53,71] and structural studies [e.g., 14–17,72] of nuclear steroid receptors have indicated the importance of steric constraints and size in predicting and interpreting ligand binding. For example, Waller et al. [71] reported that with steroid derivatives, increases in steric bulk (size) in the vicinity of the B ring off the C6 and C7 atoms, the C ring off the C11 and C12 atoms, and the D ring off the C17 atom are associated with increased AR binding affinity. Consistent with these observations, 3D models of steroid receptor ligand-binding domains suggest specific spatial arrangements of amino acid residues thought to be associated with A and D ring ligand interactions [14–17,72]. Thus, our finding that interatomic distances are similar for active ligands likely is associated with size constraints in the binding domain and/or indirectly associated with local steric characteristics [37]. The observation that active AR ligands have specific local

electronic descriptor parameter ranges associated with greater negative charge [37] is also consistent with previous research for steroid receptors. For example, analyses by Waller et al. [71] suggested that increased negative charge in the vicinity of the C3 atoms of the A ring and C17 substituents of the D ring are associated with increased ligand AR binding affinity of steroidal compounds. These findings are consistent with those of Bradbury et al. [36] and VanderKuur et al. [50], which indicated similar characteristics in ER ligands. Although recent structural studies of the ligand-binding domains of nuclear receptors suggest that these sites are largely hydrophobic [14–17,72], there is also evidence of specific residues that may be associated with hydrogen bonding and weak polar interactions [16,42]. In turn, different residues within the binding pockets of different nuclear receptors are presumed to be responsible for determining ligand specificity [16].

Once appropriate descriptors were identified (step 2), common reactivity patterns based on the conformer frequency distributions of the AR ligands in training sets were then examined (step 3). The reactivity patterns described below were derived from interatomic distances and atomic charges. Common reactivity patterns based on the analyses of the four and six most active ligands identified two distance ranges due to oxygen–oxygen interatomic distances. A less restrictive range of 10.2 to 11.1 Å was derived from the six most active ligands, whereas a more restrictive range of 10.7 to 11.0 Å was derived from the four most active ligands. These distance ranges were combined with the charge requirements for the oxygen atoms associated with ranges derived from the six or four most active ligands. A less restrictive charge range of −0.333 to −0.303 atomic units (a.u.) was derived from the set of the six most active ligands, whereas a range of −0.324 to −0.304 a.u. was derived from the four most active ligands.

With the simultaneous fulfillment of these criteria, our objective was to determine whether the algorithm could identify ligands within validation sets with pK_i values less than −0.48 or −2.00, respectively. All distance and charge ranges were assessed against ''wild card'' atoms in the validation sets (i.e., atom types were not specified in the compounds when performing the screenings), even though these parameter ranges were associated with oxygen–oxygen distances, or oxygen charges. A screening based on wild card atoms was undertaken to illustrate the ability of the algorithm to assess similarity without the need to predetermine a pharmacophore or to establish an alignment against a lead, or template, molecule. An initial evaluation of the ability of the algorithm to screen unknown compounds for AR binding affinity was obtained through an analysis of compounds 9 through 21 in Figure 3. These 13 compounds were divided into two subsets. The first subset included compounds 9 through 13 (pK_i values of −0.48 to −1.61) and was represented by 45 conformers. The second subset was comprised of compounds 14 through 21 (pK_i values of −2.00 to −2.70) and was represented by 157 conformers. A one-step screening that simultaneously incorporated charge and distance components of the reactivity patterns was used.

After applying the less restrictive screens for the distances and charges (i.e., which determined whether there are two wild card atoms within the range of −0.333 to −0.303 a.u. and within an interatomic distance range of 10.2 to 11.1 Å) to compounds 9 through 13, no conformers were identified, which is consistent with the observed pK_i values for these ligands. By applying the same distance and charge requirement to the second set of compounds (14–21), one conformer was identified for compound 17 (pK_i = −2.70). However, use of the more restrictive distance and charge ranges eliminated this conformer.

The reactivity-pattern-based screening approach was also assessed against an external validation set comprised of seven compounds, including vinclozolin, linuron, and methoxychlor metabolites, as well PCB 153 and its hydoxylated analogue (pK_i values of −1.63 to −2.54 for the seven compounds). The conformer generation routine and subsequent quantum chemical optimization produced 132 conformers for the seven compounds (all within 20 kcal/mol of the lowest-energy geometries). None of these compounds were identified when the less restrictive distance and charge range requirements were imposed. Thus, all seven compounds were properly identified as ligands likely to exhibit a pK_i of less than 0.70.

SUMMARY AND CONCLUSIONS

Assuming that conformational flexibility could be of importance in ligand–receptor interactions, we used a dynamic SAR method to model the relative AhR, ER, and AR binding affinity of a series of diverse ligands. As opposed to conventional SAR methods, the exploratory approaches reviewed here do not assume a priori that 3D molecular structures represented by single lowest-energy gas-phase conformers are appropriate to model physicochemical properties, toxicological endpoints, or the interactions of xenobiotics with receptors in a solvated environment. Use of this approach permits a systematic assessment of the conformational space associated with the compounds and endpoints of interest, either in the context of testing specific hypotheses or as a means of exploring possible ligand–receptor interactions. Research in progress with more diverse chemical structures will evaluate the predictability of this method when applied to large data sets of industrial organic chemicals.

Because the energy for transitions among certain conformations of a ligand can be less than the free energy of binding to a receptor, and because the electronic characteristics across conformers of a molecule can vary widely, hypotheses concerning the role of specific electronic properties in predicting binding affinities of xenobiotics should be tested with careful attention to the appropriate use of specific geometries. In addition, ligand binding to a receptor can alter the geometry or energy of these components in the resulting activated complex and modulate subsequent effects (see Katzenellenbogen et al. [13] and references cited therein). Thus, reports that the regulation of estrogenic responses by E_2 derivatives [38,49–51] and PCHBs [73] is not directly related to ER binding affinity may reflect, at least in part, the role of ligand flexibility [38,49–51], the geometry and energy of the activated ER–ligand complex, and associated cooperative conformational changes, and the resulting interactions of the receptor with the effector [13]. In the context of developing SARs for nonsteroidal compounds, the ability to assess conformational flexibility also appears to be critical, not only for predicting hormone binding affinity but also for predicting coupling of the receptor with nuclear effectors and the subsequent induction of products from responsive genes.

Acknowledgement—This work was supported in part by U.S. Environmental Protection Agency Cooperative Agreement CR822306-01-0 with the Higher Institute of Chemical Technology, Bulgaria, as well as a Bulgarian Science Foundation grant X-409. The participation of S. Karabunarliev, J. Ivanov, D. Nikolov, and D. Bonchev in the development of the OASIS system is greatly appreciated.

REFERENCES

1. Kavlock RJ, et al. 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.
2. Ankley GT, Johnson RD, Detenbeck NE, Bradbury SP, Toth G, Folmar LC. 1997. Development of a research strategy for assessing the ecological risk of endocrine disruptors. Rev Toxicol 1:231–267.
3. Zeeman M, Auer CM, Clements RG, Boethling RS. 1995. U.S. EPA regulatory practices on the use of QSAR for new and existing chemical evaluations. *SAR QSAR Environ Res* 3:179–201.
4. Waller CL. 1997. Recent developments on the use of computer techniques to identify endocrine disruptors. *Abstracts*, 17th Annual Meeting, Society of Environmental Toxicology and Chemistry, Washington, DC, USA, November 17–21, p 2.
5. Auer CM, Nabholz JV, Baetcke KP. 1990. Mode of action and the assessment of chemical hazards in the presence of limited data: Use of structure–activity relationships (SARs) under TSCA section 5. *Environ Health Perspect* 87:183–197.
6. U.S. Environmental Protection Agency. 1991. Evaluation of research on expert systems for predicting the environmental fate and effects of chemicals. EPA-SAB-EPEC-91-004, Science Advisory Board, Washington, DC.
7. U.S. Environmental Protection Agency. 1991. Evaluation of EPA's Ecorisk assessment research program. EPA-SAB-EPEC-92-006. Science Advisory Board, Washington, DC.
8. Bradbury SP. 1995. Quantitative structure–activity relationships and ecological risk assessment: An overview of predictive aquatic toxicology research. *Toxicol Lett* 79:229–237.
9. Bradbury SP. 1994. Predicting modes of toxic action from chemical structure: An overview. *SAR QSAR Environ Res* 2:89–104.
10. Russom CL, Bradbury SP, Broderius SJ, Hammermeister DE, Drummond RA. 1997. Predicting modes of toxic action from chemical structure: Acute toxicity of industrial organic chemicals to the fathead minnow (*Pimephales promelas*). *Environ Toxicol Chem* 16:948–967.
11. Rickenbacher U, McKinney JD, Oatley SJ, Blake CCF. 1986. Structurally specific binding of halogenated biphenyls to thyroxine transport protein. *J Med Chem* 29:641–648.
12. de Groot MJ, van der Aar EM, Nieuwenhuizen RJ, van der Plas RM, den Kelder GMD, Commandeur JNM, Vermeulen NPE. 1995. A predictive substrate model for rat glutathione *S*-transferase 4-4. *Chem Res Toxicol* 8:649–658.
13. 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.
14. Wagner RL, Apriletti JW, McGrath ME, West BL, Baxter JD, Fletterick RJ. 1995. A structural role for hormone in the thyroid hormone receptor. *Nature* 378:690–697
15. Renaud JP, Rochel N, Ruff M, Vivat V, Chambon P, Gronemeyer H, Moras D. 1995. Crystal structure of the RAR–ligand binding domain bound to all trans retinoic acid. *Nature* 378:681–689.
16. 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. *Nature Struct Biol* 3:87–94.
17. Bourguet W, Ruff M, Chambon P, Gronemeyer H, Moras D. 1995. Crystal structure of the ligand binding domain to all-trans retinoic acid. *Nature* 375:377–382.
18. Eliel EL. 1993. Chemistry in three dimensions. In Warr WA, ed, *Chemical Structures*, Vol. 1. Springer, Berlin, Germany, pp 1–8.
19. Mekenyan OG, Ivanov JM, Veith GD, Bradbury SP. 1994. Dynamic QSAR: A new search for active conformations and significant stereoelectronic indices. *Quant Struct-Act Relat* 13:302–307.
20. Ivanov JM, Karabunarliev SH, Mekenyan OG. 1994. 3DGE: An algorithm for exhaustive generation of 3D isomers proceeding from molecular topology. *J Chem Inf Comput Sci* 34:234–243.
21. Mekenyan OG, Schultz TW, Veith GD, Kamenska VB. 1996. ''Dynamic'' QSAR for semicarbazide-induced mortality in frog embryo. *J Appl Toxicol* 16:355–363.
22. Veith GD, Mekenyan OG, Ankley GT, Call DJ. 1995. QSAR evaluation of α-terthienyl phototoxicity. *Environ Sci Technol* 29: 1267–1272.
23. Mekenyan OG, Veith GD, Call DJ, Ankley GT. 1996. A QSAR evaluation of Ah receptor binding of halogenated aromatic xenobiotics. *Environ Health Perspect* 104:1302–1310.
24. Poland A, Knutson JC. 1982. 2,3,7,8-Tetrachlorodibenzo-*p*-dioxin and related halogenated aromatic hydrocarbons: Examination of the mechanism of toxicity. *Ann Rev Pharmcol Toxicol* 22:517–554.
25. Bandiera S, Safe S, Okey AB. 1982. Binding of polychlorinated biphenyls classified as either phenobarbitone, 3-methylchlolanthrene- or mixed-type inducers to cytosolic Ah receptor. *Chem-Biol Interact* 39:259–277.
26. Bandiera S, Sawyer T, Campbell M, Fujita T, Safe S. 1983. Competitive binding to the cytosolic 2,3,7,8-tetrachlorobiphenyl-*p*-dioxin receptor. Effects of structure on the affinities of substituted halogenated biphenyls—A QSAR analysis. *Biochem Pharmacol* 32:3803–3813.
27. Safe S, et al. 1985. Effects of structure on binding to the 2,3,7,8-TCDD receptor protein and AHH induction—Halogenated biphenyls. *Environ Health Perspect* 61:21–23.
28. Nerbert DW. 1989. The Ah locus: Genetic differences in toxicity, cancer, mutation, and birth defects. *Toxicology* 20:153–174.
29. Okey AB. 1990. Enzyme induction in the cytochrome P-450 system. *Pharmacol Ther* 45:241–298.
30. Safe S. 1990. Polychlorinated biphenyls (PCBs), dibenzo-*p*-dioxins (PCDDs), dibenzofurans (PCDFs), and related compounds: Environmental and mechanistic considerations which support the development of toxic equivalency factors (TEFs). *Crit Rev Toxicol* 21:51–88.
31. Gallo MA, Scheuplein RJ, Van Der Hejiden KA, eds. 1991. *Biological Basis for Risk Assessment of Dioxins and Related Compounds.* Banbury Report 35. Cold Springs Harbor Laboratory Press, Cold Springs Harbor, NY, USA.
32. McKinney JD, Long GA, Pedersen LG. 1984. PCBs and dioxin binding to cytosol receptors: A theoretical model based on molecular parameters. *Quant Struct-Act Relat* 3:99–105.
33. Kafafi SA, Afeefy HY, Ahi AM, Said HK, Abd-Elazem IS, Kafafi AG. 1993. Affinities for the arylhydrocarbon receptor, potencies as arylhydrocarbon hydroxylase inducers and relative toxicities of polychlorinated biphenyls: A congener-specific approach. *Carcinogenesis* 14:2063–2071.
34. Kafafi SA, Afeefy HY, Said HK, Kafafi AG. 1993. Relationships between arylhydrocarbon hydroxylase and 7-ethoxyresofuran *O*-deethylase enzymes, and toxic activities of aromatic xenobiotics in animals: A new model. *Chem Res Toxicol* 6:328–334.
35. Waller CL, McKinney JD. 1995. Three-dimensional quantitative structure–activity relationships of dixoins and dioxin-like compounds: Model validation and Ah receptor characterization. *Chem Res Toxicol* 8:847–858.
36. 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.
37. Mekenyan OG, Ivanov JM, Karabunarliev S, Bradbury SP, Ankley GT, Karcher W. 1998. COREPA: A new approach for the elucidation of common reactivity patterns of chemicals. I. Stereoelectronic requirements for androgen receptor binding. *Environ Sci Technol* (in press).
38. Wiese T, Brooks SC. 1994. Molecular modeling of steroidal estrogens: Novel conformations and their role in biological activity. *J Steroid Biochem Mol Biol* 50:61–72.
39. McKinney JD, Pederson LG. 1986. Biological activity of polychlorinated biphenyls related to conformational structure. *Biochem J* 240:621–622.
40. Parr RG, Zhou Z. 1993. Absolute hardness: Unifying concept for identifying shells and subshells in nuclei, atoms, molecules, and metallic clusters. *Accounts Chem Res* 26:256–258.
41. Pearson RG. 1993. The principle of maximum hardness. *Accounts Chem Res* 26:250–255.
42. 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.
43. Hahnel R, Twaddle E, Ratajczak T. 1973. The specificities of the estrogen receptor of human uterus. *J Steroid Biochem* 4:21–31.
44. Hochberg RB, Zielinski JE, Duax WL, Strong P. 1986. The molecular structure of 16 α-iodo-17β-estradiol, a high affinity ligand for the estrogen receptor. *J Steroid Biochem* 25:615–618.

45. 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.
46. 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.
47. Palomino E, Heeg MJ, Horwitz JP, Polin L, Brooks SC. 1994. Skeletal conformations receptor binding of some 9,11-modified estradiols. *J Steroid Biochem Mol Biol* 50:75–84.
48. Ruh MF, Turner JW, Paulson CM, Ruh TS. 1990. Differences in the form of the salt-transformed estrogen receptor when bound by estrogen versus antiestrogen. *J Steroid Biochem* 32:509–516.
49. Pilat MJ, Hafner MS, Kral LG, Brooks SC. 1993. Differential inductin of pS2 and cathepsin D mRNAs by structurally altered estrogens. *Biochemistry* 32:7009–7015.
50. VanderKuur JA, Wiese T, SC Brooks. 1993. Influence of estrogen structure on nuclear binding and progesterone receptor induction by the receptor complex. *Biochemistry* 32:7002–7008.
51. VanderKuur JA, Hafner MS, Christman JK, Brooks SC. 1993. Effects of estradiol-17β analogues on estrogen response element regulated chloramphenicol acetyltransferase expression. *Biochemistry* 32:7016–7021.
52. Korach KS, Sarver P, Chae K, McLachlan JA, McKinney GD. 1987. Estrogen receptor-binding activity of polychlorinated hydroxybiphenyls: Conformationally restricted structural probes. *Mol Pharmcol* 33:120–126.
53. Pakdel F, Katzenellenbogen BS. 1992. Human estrogen receptor mutants with altered estrogen and antiestrogen ligand discrimination. *J Biol Chem* 267:3429–3437.
54. Amouri HE, Vessiers A, Vichard D, Top S, Grusselle M, Jaouen G. 1992. Syntheses and affinities of novel organometallic-labeled estradiol derivatives: A structure–affinity relationship. *J Med Chem* 35:3130–3135.
55. Lewis DFV. 1989 The calculation of molar polarizability by the CNDO/2 method: Correlation with the hydrophobic parameter, Log P. *J Comput Chem* 10:145–151.
56. Schüürmann G. 1990. Quantitative structure–property relationships for the polarization, solvatochromic parameters and lipophilicity. *Quant Struct-Act Relat* 9:326–333.
57. Balaban AT, Motoc I, Bonchev D, Mekenyan O. 1984. Topological indices for structure–activity correlations. *Top Curr Chem* 114:21–71.
58. Hahn M. 1995. Receptor surface models. 1. Definition and construction. *J Med Chem* 38:2080–2090.
59. Cramer RD III, Patterson DE, Bunce JD. 1988. Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins. *J Am Chem Soc* 110:5959–5967.
60. Goodford PJ. 1985. Computational procedure for determining energetically favorable binding sites on biologically important macromolecules. *J Med Chem* 28:849–857.
61. Marshall GR. 1993. Binding site modeling of unknown receptors. In Kubinyi H, ed., *3D QSAR in Drug Design: Theory, Methods and Applications.* Escom, Leiden, The Netherlands, pp 80–116.
62. Buchheit K-H, et al. 1995. The serotonin 5-HT_4 receptor. 1. Design of a new class of agonists and receptor map of the agonist recognition site. *J Med Chem* 38:2326–2330.
63. Hahn M, Rogers D. 1995. Receptor surface models. 2. Application to quantitative structure-activity relationships studies. *J Med Chem* 38:2091–2102.
64. Kearsely SK, GM Smith. 1990. An alternative method for the alignment of molecular structures: Maximizing electrostatic and steric overlap. *Tetrahedron Comput Methodol* 3:615–633.
65. Blaney JM, Dixon JS. 1993. A good ligand is hard to find: Automatic docking methods. *Perspect Drug Des* 1:301–319.
66. Perkins TD, Deam PM. 1993. An exploration of a novel strategy for superimposing several flexible molecules. *J Comput-Aided Mol Des* 7:155–172.
67. Martain YC, Bures MG, Danahar EA, DeLazzar J, Lico I, Pavlik PA. 1993. A fast new approach to pharmacophore mapping and its application to dopaminergic and benzodiazepine agonists. *J Comput-Aided Mol Des* 7:83–102.
68. Prendergast K, Adams K, Greenlee WJ, Nachbar RB, Patcett AA, Underwood DJ. 1994. Derivation of a 3D pharmacophore model for the angiotensin-II site one receptor. *J Comput-Aided Mol Des* 8:491–512.
69. Mekenyan OG, Ivanov JM, Karabunarliev SH, Hansen B, Ankley GT, Bradbury SP. 1998. A new approach for estimating three-dimensional similarity that incorporates molecular flexibility. In Chen F, ed, *Proceedings,* 7th International Workshop on QSAR in Environmental Sciences, Elsinore, Denmark, June 21–25, 1996. SETAC, Pensacola, FL, USA.
70. Kelce WR, Monosson E, Gamasik MP, Laws SC, Gray LE Jr. 1994. Environmental hormone disruptors: Evidence that vinclozolin developmental toxicity is mediated by antiandrogenic metabolites. *Toxicol Appl Pharmacol* 126:276–285.
71. Waller CL, Juma BW, Gray LE Jr, Kelce WR. 1996. Three-dimensional quantitative structure–activity relationships for androgen receptor ligands. *Toxicol Appl Pharmacol* 137:219–227.
72. 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 U S A* 90:9949–9953.
73. Moore M, et al. 1997. Antiestrogenic activity of hydroxylated polychlorinated biphenyl congeners identified in human serum. *Toxicol Appl Pharmacol* 142:160–168.