Comparative Molecular Field Analysis of Polyhalogenated Dibenzo-*p*-dioxins, Dibenzofurans, and Biphenyls

Chris L. Waller and James D. McKinney*,†

Environmental Toxicology Division, Health Effects Research Laboratory, U.S. Environmental Protection Agency, Research Triangle Park, North Carolina 27711, and Division of Medicinal Chemistry and Natural Products, School of Pharmacy, University of North Carolina, Chapel Hill, North Carolina 27599-7360

Received January 24, 1992

Comparative molecular field analysis (CoMFA) was performed on polyhalogenated dibenzo-*p*-dioxins, dibenzofurans, and biphenyls for Ah (dioxin) receptor binding and associated enzyme inducing activities determined by others using in vitro assays. Since various members of all three classes of compounds have been shown to produce qualitatively similar toxicities, a separate CoMFA was performed on each class of compounds and combinations of the different classes for each bioactivity which included combining all three classes of molecules in one CoMFA study. For the Ah receptor binding, the CoMFA-derived QSARs for all three classes of compounds and combinations thereof showed strong crossvalidated correlations indicating that they are highly predictive. For enzyme induction, the CoMFA-derived QSARs were highly predictive for the dibenzofurans but were only partially successful for the dioxins. For the biphenyls, the results were clearly unpredictive. The overall results of these CoMFA studies which include both steric and electrostatic considerations are compared and contrasted to other SAR models that have met with some success in making qualitative predictions about the potential for receptor binding and associated toxicity in these classes of compounds. The CoMFA-derived QSAR for the dioxin series of molecules in most cases significantly overestimates the enzyme inducing ability of the ortho-substituted biphenyls. This weak inducing activity of the *o*-biphenyls is, however, consistent with their relatively low dioxin-like toxicity as measured in other biological systems. Fundamentally different mechanisms may be operating in the expression of dioxin-like toxic responses for the *o*-biphenyls, and their direct, dioxin-like toxic equivalency perhaps needs to be reconsidered in this light.

Introduction

Halogenated aromatic hydrocarbons, typified by the polychlorinated dibenzo-*p*-dioxins (PCDDs), dibenzofurans (PCDFs), and biphenyls (PCBs), have been identified in almost every compartment of the global ecosystem. Because of their lipophilic nature, these compounds have also been detected in fish, wildlife, and various human body fluids and tissues. There is considerable public and regulatory concern over the potential adverse human health effects and envirnomental damage associated with exposure to these chemicals. The prototypical halogenated aromatic hydrocarbon, 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD or dioxin), and related compounds elicit similar biochemical and toxic responses in humans, laboratory animals, and mammalian cells in culture. There is a substantial body of evidence that many, if not all, of their important biological responses are mediated by a common (Ah or dioxin) receptor mechanism of action. The activity of the different halogenated aromatics is structure dependent, and a number of studies have delineated the various structure–activity relationships (SARs).[1-4] These relatively well studied SARs have provided a mechanistic basis for the development of toxic equivalency factors (TEFs) for certain classes of halogenated aromatics. Since these compounds are invariably present in environmental and biological samples as highly complex mixtures of isomers and congeners, TEF values (expressed as TCDD equivalents) have been proposed as a way to predict the toxicity of their mixtures and facilitate hazard and risk assessment activities. However, because there are many data gaps in deriving such values, there is a continuing controversy over the use of such an approach in hazard and risk assessment.

There is a need for a more quantitative predictive model based on SAR that would also permit combined consideration of some of the more important classes of halogenated aromatic hydrocarbons of concern. Several relatively simple models have been proposed in the literature over the years that have met with some success in at least making some qualitative predictions about the potential for receptor binding and associated toxicity in these classes of compounds. Perhaps the simplest was the original 3 × 10 Å box model proposed by Poland et al.[1] In this model the most active compounds were planar and substituted with halogens in lateral positions affording a rectangular shape. This model was extended by Gillner et al.[2] to include nonhalogenated aromatic hydrocarbons that also approximated a rectangular shape when the van der Waals surface was considered. We later proposed a stacking model[3] in which the important molecular parameters were molecular polarizability and the receptor to ligand separation distance. In this model molecular size and shape and halogenation were important only to the extent that they affected polarizability and the stacking distance. Safe

† U.S. Environmental Protection Agency.

(1) Poland, A.; Knutson, J. C. 2,3,7,8-Tetrachlorodibenzo-p-dioxin and Related Halogenated Aromatic-hydrocarbons: Examination of the Mechanism of Toxicity. *Annu. Rev. Pharmacol. Toxicol.* **1982**, *22*, 517–554.

(2) Gillner, M.; Bergman, J.; Cambillau, C.; Fernstrom, B.; Gustafsson, J.-A. Interactions of Indoles with Specific Binding Sites for 2,3,7,8-Tetrachlorodibenzo-p-dioxin in Rat Liver. *Mol. Pharmacol.* **1985**, *28*, 357–363.

(3) McKinney, J. D.; Darden, T.; Lyerly, M. A.; Pederson, L. G. PCB and Related Compound Binding to the Ah Receptor(s). Theoretical Model Based on Molecular Parameters and Molecular Mechanics. *Quant. Struct.-Act. Relat.* **1985**, *4*, 166–172.

0022-2623/92/1835-3660$03.00/0

and co-workers[4] also derived various quantitative structure–activity relationships (QSARs) in which such molecular descriptors as electronegativity, hydrophobicity, and hydrogen-bonding capacity for substituents were shown to be important.

Comparative molecular field analysis (CoMFA)[5,6] is a promising new approach to structure–activity correlation. CoMFA permits representation of ligand molecules by their steric and electrostatic fields, field fitting to optimize mutual alignment within a series of molecules, crossvalidation[6] to indicate the likelihood that the results have predictive validity, and graphic representation of the results. Such an approach appeared to be particularly suited to study of the binding of these structurally related compounds to the Ah receptor and associated biological responses. However, the results of these QSAR studies obtained from the analyses of rather specific sets of related compounds do not alone necessarily give information of general utility for understanding mechanisms of molecular interactions, but they can be used as an element in a quantitative approach to toxicity prediction. In this work we describe the use of CoMFA to model various PCDDs, PCDFs, and PCBs and to correlate binding affinities and enzyme induction activities with ligand molecules' three-dimensional steric and electrostatic fields. The binding and enzyme induction data previously reported by Safe et al.[4,7,8] were used in our study since they represented the most homogeneous data sets across the compound classes of interest. It was of interest to see if all three classes of compounds could be combined into a single predictive model since various members of all three classes have been shown to produce qualitatively similar toxicities. Such a model might have value in extending efforts to develop TEFs for these compounds for possible use in hazard and risk assessment.

Computational Methods

Materials and Approaches. The molecular modeling and CoMFA studies were carried out on an Evans and Sutherland ESV workstation running SYBYL 5.41.[9] CoMFA was used as a three-dimensional QSAR method to analyze the correlations between the measured biological activity data and the molecules' steric and electrostatic properties as evaluated by a probe atom. A computationally efficient statistical algorithm, partial least squares (PLS) analysis,[10] was used in conjunction with crossvalidation[6] to measure the predictability of the dataset. Bootstrap analysis[6] of the dataset provided a means by which statistical confidence limits could be placed on the results. A QSAR equation derived from these datasets via PLS analysis typically contains hundreds of terms and is best portrayed as a three-dimensional contour plot on a molecular modeling workstations' video monitor. This map-like representation of the QSAR equation can then be used to qualitatively predict the structures of novel agents. After being subjected to field probing, the biological activities of these untested molecules may then be quantitatively predicted within statistical confidence limits.

Table I. Dibenzo-*p*-dioxins Included in QSAR Analyses

name	compd	pEC_{50}-(bind)[c]	pEC_{50}-(AHH)[c]	pEC_{50}-(EROD)[c]
Chlorinated Congeners[a]				
TCDD	2378	8.000	9.721	10.143
PCDD2	12378	7.102	7.770	7.959
PCDD3	2367	6.796	7.959	7.215
PCDD4	236	6.658		
PCDD5	123478	6.553	8.387	8.678
PCDD6	1378	6.102	6.495	6.229
PCDD7	12478	5.959	7.959	7.678
PCDD8	1234	5.886	5.620	5.432
PCDD9	237	7.149	6.854	6.444
PCDD10	28	5.495	4.000	4.000
PCDD11	12347	5.194	6.086	6.180
PCDD12	124	4.886	5.658	4.319
PCDD13	OCDD	5.000	6.155	6.509
PCDD14	1	4.000	4.000	4.000
Brominated (B) and Chlorinated (C) Congeners[b]				
PCDD18	2B3B7B8B	8.824		
PCDD19	2B3B7C8C	8.830		
PCDD20	2B3C7C8B	9.350		
PCDD21	2B3C7C8C	7.939		
PCDD22	1B3B789B	7.032		
PCDD23	1B3B7B8B	8.699		
PCDD24	1B2B4B7B8B	7.770		
PCDD25	1B2B3B7B8B	8.180		
PCDD26	2B3B7B	8.932		
PCDD27	2B7B	7.810		
PCDD28	2B	6.530		

[a] Safe, S. *Annu. Rev. Pharmacol. Toxicol.* **1986**, *26*, 371–399. [b] Safe, S. *Crit. Rev. Toxicol.* **1990**, *21* (1), 51–88. [c] Standard deviation data are not available.

Polychlorinated and polybrominated dibenzo-*p*-dioxins, polychlorinated dibenzofurans, and polychlorinated biphenyls have been analyzed for their ability to bind to the cytosolic Ah (dioxin) receptor and to stimulate the induction of aryl hydrocarbon hydroxylase (AHH) and ethoxyresorufin *O*-deethylase (EROD) in in vitro rat hepatocyte assays. The compounds and biological activity data used in the QSAR analyses reported herein are found in Tables I–III.

Molecular Modeling and Alignment Rules. All of the compounds were built de novo using the sketch option in SYBYL 5.41 and fully geometry optimized using the standard TRIPOS molecular mechanics force field with a 0.001 kcal/mol energy gradient convergence criterion and a distance-dependent dielectric. Partial atomic charges were used as determined by MOPAC[11] calculations using the AM1[12,13] model Hamiltonian.

TCDD the prototypical Ah receptor ligand was chosen as the template molecule on which to align the various dibenzo-*p*-dioxin, dibenzofuran, and biphenyl analogs (see

(4) Safe, S. H. Polychlorinated Biphenyls (PCBs), Dibenzo-p-Dioxins (PCDFs), Dibenzofurans (PCDFs), and Related Compounds: Environmental and Mechanistic Considerations Which Support the Development of Toxic Equivalency Factors (TEFs). *Crit. Rev. Toxicol.* **1990**, *21* (1), 51–88.

(5) Cramer, R. D.; Patterson, D. E.; Bunce, J. D. Comparative Molecular Field Analysis (CoMFA). 1. Effect of Shape on Binding of Steroids to Carrier Proteins. *J. Am. Chem. Soc.* **1988**, *110*, 5959–5967.

(6) Cramer, R. D.; Bunce, J. D.; Patterson, D. E.; Frank, I. E. Crossvalidation, Bootstrapping, and Partial Least Squares Compared with Multiple Regression in Conventional QSAR Studies. *Quant. Struct.–Act. Relat.* **1988**, *7*, 18–25.

(7) Safe, S. H. Comparative Toxicology and Mechanism of Action of Polychlorinated Dibenzo-p-dioxins and Dibenzofurans. *Annu. Rev. Pharmacol. Toxicol.* **1986**, *26*, 371–399.

(8) Bandiera, S.; Safe, S.; Okey, A. B. Binding of Polychlorinated Biphenyls Classified as Either Phenobarbitone-, 3-Methylcholanthrene-, or Mixed-type Inducers to Cytosolic Ah Receptor. *Chem.-Biol. Interact.* **1982**, *39*, 259–277.

(9) The program SYBIL 5.41 is available from TRIPOS Associates, St. Louis, MO, 1991.

(10) Wold, S.; Ruhe, A.; Wold, H.; Dunn, W. J. The covariance problem in linear regression. The partial least squares (PLS) approach to generalized inverses. *SIAM J. Sci. Stat. Comp.* **1984**, *5* (3), 735–743.

(11) MOPAC 6.0 available from Quantum Chemistry Program Exchange.

(12) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. AM1: A New General Purpose Quantum Mechanical Molecular Model. *J. Am. Chem. Soc.* **1985**, *107*, 3902–3909.

(13) Dewar, M. J. S.; Zoebisch, E. G. Extension of AM1 to the Halogens. *THEOCHEM* **1988**, *49*, 1–21.

Table II. Dibenzofurans Included in OSAR Analyses

name	compd	pEC_{50} (bind)[c]	pEC_{50} (AHH)[c]	pEC_{50} (EROD)[c]
PCDF1[a]	2	3.553		
PCDF2	3	4.377 ± 0.058		
PCDF3	4	3.000	5.000	4.767
PCDF4	23	5.326	5.660	5.315
PCDF5	26	3.609	4.210	4.200
PCDF6	28	3.590	4.403	4.398
PCDF7	136	5.357	5.597	5.472
PCDF8	138	4.071	4.712	4.520
PCDF9	234	4.721	6.821	6.606
PCDF10	238	6.000 + 0.041	5.604	5.807
PCDF11	267	6.347	5.553	5.505
PCDF12	2346	6.456	5.879	5.947
PCDF13	2348	6.699	7.383	7.425
PCDF14	1368	6.658	5.983	6.109
PCDF15	2378	7.387 ± 0.059	8.401	8.695
PCDF16	1248	5.000	4.921	4.033
PCDF17	12467	7.169	6.488	6.458
PCDF18	12479	4.699	7.424	7.416
PCDF19	12348	6.921	6.680	6.788
PCDF20	12378	7.128 ± 0.105	8.595	8.514
PCDF21	12478	5.886	6.975	6.830
PCDF22	23478	7.824 ± 0.028	9.952	9.873
PCDF23	123478	6.638	9.449	9.421
PCDF24	123678	6.569 ± 0.137	7.833	8.907
PCDF25	124678	5.081	7.373	8.533
PCDF26	234678	7.328 ± 0.036	9.163	9.240
PCDF27[b]	2368	6.658		
PCDF28	1236	6.456		
PCDF29	1237	6.959		
PCDF30	13478	6.699		
PCDF31	23479	6.699		
PCDF32	12379	6.398		
PCDF33	non	3.000		
PCDF34	2347	7.602	7.824	7.745
PCDF35	1237	6.959	4.201	4.569
PCDF36	13478	6.699	8.854	8.796
PCDF37	23479	6.699	8.237	8.102
PCDF38	12379	6.398	7.066	7.066
PCDF39	12468	5.509	4.921	5.000

[a] Safe, S. *Annu. Rev. Pharmacol. Toxicol.* **1986**, *26*, 371–399 (PCDD1–PCDD26). [b] Safe, S. *Crit. Rev. Toxicol.* **1990**, *21* (1), 51–88 (PCDD27–PCDD39). [c] When not reported, standard deviation data are not available.

Table III. Biphenyls Included in QSAR Analyses

name	compd	pEC_{50} (bind)[b]	pEC_{50} (AHH)[b]	pEC_{50} (EROD)[b]	(2–1–1′–2′) torsion angle, deg
PCB1	33′44′	6.149	7.55	7.05	128.9
PCB2[a]	344′5	4.553			129.0
PCB3	33′44′5	6.886	9.62	9.61	128.4
PCB4[a]	2′344′5	4.854	5.41	5.95	119.8
PCB5	233′44′	5.367	7.06	6.92	121.9
PCB6	23′44′5	5.041	4.94	5.05	60.1
PCB7[a]	2344′5	5.387	6.02	6.25	118.4
PCB8[a]	233′44′5	5.301	6.15	5.90	119.6
PCB9[a]	23′44′55′	4.796	4.88	5.05	61.7
PCB10[a]	233′44′5′	5.149	5.68	6.05	117.0
PCB11	22′44′	3.886			−68.3
PCB12	22′44′55′	4.102			−116.2
PCB13[a]	2345	3.854			−119.1
PCB14[a]	23′44′5′6	4.004			−110.8

Bandiera, S.; Safe, S.; Okey, A. B. *Chem.-Biol. Interact.* **1982**, *39*, 259–277. [a] Ambiguous torsion measurement due to ring-substitution symmetry. [b] Standard deviation data are not available.

Figure 1). Due to the symmetrical nature of dibenzo-*p*-dioxin and dibenzofuran molecules, they were aligned via root mean square (RMS) fit of their 2,3,7,8 carbons to the corresponding carbons of the TCDD template molecule. To avoid redundancy in the dataset, i.e. the inclusion of the same molecule into the analysis in a second or third alignment, the conventional (IUPAC) substituent numbering scheme was used to determine the 2,3,7,8 carbons of the molecules. Biphenyl molecules were aligned initially

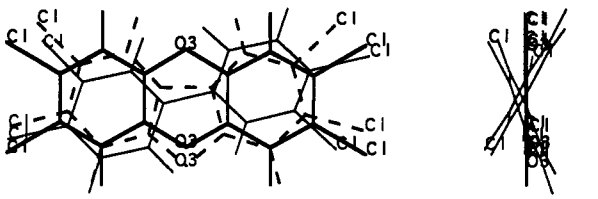

Figure 1. Orthogonal view of alignment rule. Molecules are TCDD, 2,3,7,8-tetrachlorodibenzofuran, and 3,3′,4,4′-tetrachlorobiphenyl.

via RMS fit of their 3,3′,4,4′ carbons to the 3,8,2,7 carbons of the TCDD template molecule, respectively, followed by field fit optimization to the template molecule. Field fitting typically forced the biphenyl molecules into energetically unreasonable coplanar conformations in order to obtain maximal similarity between the steric and electrostatic fields of template and test molecules. The resulting high-energy structures were subsequently reoptimized without the field fit option to allow for relaxation of the molecule about the torsion angle.

This particular alignment rule was chosen because it provided for the most direct comparison of substituent effects between the three series of molecules and yielded the most predictive model. However, due to the ambiguous nature of this alignment, i.e. molecular symmetry, and the dependence of a CoMFA study on such, this alignment rule may represent a potential source of noise in the QSAR analyses presented herein.

CoMFA Interaction Energy Calculations. The steric, in terms of the van der Waals (6–12) interactions, and electrostatic, Coulombic with a $1/r$ distance-dependent dielectric, potential energy fields were calculated at each lattice intersection on a regularly-spaced grid. The grid spacing was 2.0 Å units in each direction with the grid extending 4.0 Å units in every direction away from the molecule. An sp^3 carbon probe atom with a van der Waals radius of 1.52 Å units and a +1.0 charge was used in the calculations. The steric and electrostatic energy values were truncated to 30 kcal/mol. The electrostatic energy term was ignored at lattice intersections yielding maximal (30 kcal/mol) steric values.

Partial Least Squares (PLS) Analysis. Initially, the PLS algorithm was used in conjunction with the cross-validation option to obtain the optimal number of components to be used in the subsequent analyses of the datatable. The PLS analysis was then repeated with the number of crossvalidation groups set equal to zero and the optimal number of components designated as that which yielded the highest crossvalidated r^2 value, in order to generate conventional r^2 values. A final PLS analysis with 10 bootstrap groups and the optimal number of components was performed on each datatable to place statistical confidence limits on the analysis.

A separate CoMFA was performed on each class of compounds for each measured bioactivity type. The dibenzo-*p*-dioxin and dibenzofuran datatables were then combined and analyzed followed by the inclusion of the biphenyl molecules, thus combining three classes of molecules into one CoMFA study.

Results

CoMFA of Ah (Dioxin) Receptor Binding Affinity. The compounds and biological data used in these and the following analyses are found in Tables I–III. Initially, a QSAR analysis was performed on each class of compounds with the maximum number of components set equal to

Table IV. CoMFA QSAR Results of Ah (Dioxin) Receptor Binding Affinity

class(es)	obs	r^2_{cross}	r^2	SE	F test	p value	r^2_{bs}	SD
dioxins	25	0.715 (4)[a]	0.919	0.450	56.850	0.000	0.946	0.028
furans	39	0.742 (5)	0.858	0.539	39.866	0.000	0.872	0.036
biphenyls	14	0.534 (2)	0.823	0.399	25.531	0.000	0.879	0.080
dioxins and furans	64	0.766 (6)	0.884	0.525	72.489	0.000	0.915	0.017
dioxins and biphenyls	39	0.798 (4)	0.923	0.470	101.499	0.000	0.936	0.025
furans and biphenyls	53	0.711 (4)	0.850	0.520	68.055	0.000	0.888	0.028
dioxins, furans, and biphenyls	78	0.724 (6)	0.878	0.535	85.552	0.000	0.901	0.016

[a] Numbers in parentheses are optimal number of components.

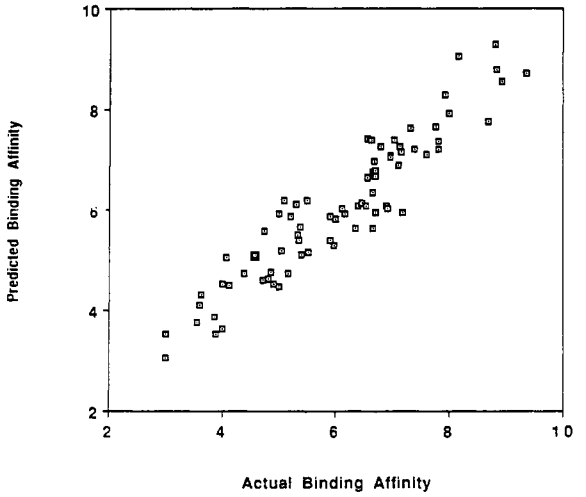

Figure 2. Predicted versus actual binding affinity values from CoMFA including all compounds.

ten and the crossvalidation groups set equal to the number of observations (rows) in the data table of interest. Each analysis was then repeated with the optimum number of components, as determined in the initial analysis, and no crossvalidation. In order to generate statistical confidence limits, this last analysis was repeated with ten bootstrap groups. The results from these analyses are found in Table IV.

The CoMFA-derived QSARs for every class of molecules and every combination of classes which were examined exhibited strong crossvalidated correlations, indicating that they were highly predictive. Crossvalidation, as employed, yields information concerning the predictive ability of the QSAR dataset by minimizing the occurrence of chance correlations in the QSAR model. Taken as such, the difference between crossvalidated and non-crossvalidated r^2 values provided for an indication of the amount of spurious correlations in the QSAR dataset. The small residual between crossvalidated and conventional r^2 values indicated only a small occurrence of spurious correlations within the datatable. The bootstrapped r^2 values, in every case, are very high with small standard deviations, thus adding a high degree of confidence to the analyses. The predicted binding affinities obtained from the analysis including all classes of compounds are plotted versus the actual values in Figure 2.

The results of the above analyses are best viewed as three-dimensional color-coded contour plots. The CoMFAs for all classes of compounds with respect to binding affinity for the Ah receptor yielded similar correlation values and qualitatively comparable contour plots. In Figures 3–6, the non-crossvalidated results of the CoMFA study including all compounds are plotted as a black and white alternative to the color-coded plot. The polyhedrons in Figure 3 represent areas in space around the molecules where increases in steric bulk have led to increases in binding affinity. The van der Waals volume of the lateral

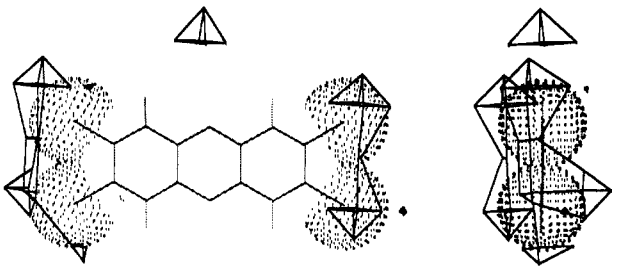

Figure 3. Orthogonal view of positive steric contribution contour plot. Contour level = +0.1. Molecule is TCDD with van der Waals radii of lateral chlorines represented as dot surfaces.

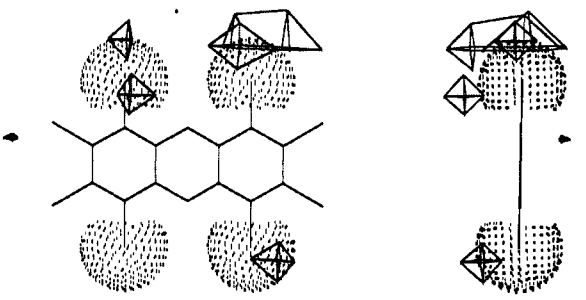

Figure 4. Orthogonal view of negative steric contribution contour plot. Contour level = −0.05. Molecule is OCDD with van der Waals radii of nonlateral chlorines represented as dot surfaces.

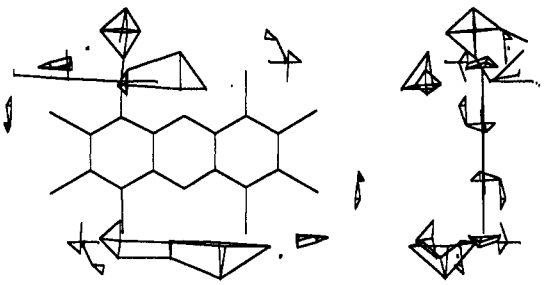

Figure 5. Orthogonal view of positive electrostatic contribution contour plot. Contour level = +0.04. Molecule is OCDD with isoelectrostatic potential plot contoured to +4 kcal/mol.

halogens of TCDD can be seen to extend into these desired areas. These contour plots were generated via an interpolation procedure which denotes spatial locations corresponding to particular columns in the datatable where the scalar products of the QSAR coefficient and the standard deviation associated with that point were a user-specified positive value of +0.1. The converse situation is plotted in Figure 4. The contours, generated at a negative 0.05 value in this case, designate areas in space where increases in steric bulk associated with the 1,4,6,9 halogens of octachlorodibenzo-*p*-dioxin (OCDD) detract from the binding affinity of the molecule.

Contour plots were also used to analyze the electrostatic contribution made to binding affinity for this same series of analogs. In Figures 5 and 6, the polyhedrons represent spatial areas where positive electrostatic potential is desired, specifically associated with the nonlateral posi-

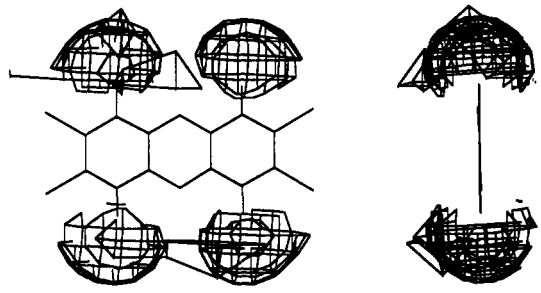

Figure 6. Orthogonal view of positive electrostatic contribution contour plot. Contour level = +0.04. Molecule is TCDD with isoelectrostatic potential plot contoured to +4 kcal/mol.

tions on the parent ring systems. The isoelectrostatic potential (+4.0 kcal/mol) map for OCDD is superimposed in Figure 5 to clearly illustrate the deficiency of positive potential in these areas, while the TCDD map demonstrates the presence of positive potential in the required areas in Figure 6.

CoMFA of in Vitro Aryl Hydrocarbon Hydroxylase (AHH) Induction. The compounds included in the following analyses were subsets of the series of compounds used to generate QSARs for binding to the Ah receptor due to the unavailability of in vitro induction data (see Tables I–III) for some of the compounds. The analyses were performed as described above, and the results are found in Table V.

The largest occurrence of spurious correlations appeared to exist with the biphenyl dataset as evidenced by the sizeable variation between crossvalidated and conventionally-derived r^2 values. Furthermore, the crossvalidated correlation was very weak, suggesting almost no predictive value.

The correlation values obtained for the dibenzo-*p*-dioxin and dibenzofuran datasets also decreased significantly in all cases. Although the crossvalidated value appears to represent predictability, the low conventional and bootstrapped r^2 values suggested that the correlations drawn between the changes in the fields of these molecules with respect to the changes in measured AHH induction should be of limited use. As expected, the inclusion of biphenyls into either of these analyses was detrimental to the model.

When all of the compounds were considered in one analysis, the resulting correlations, although weaker than those for the Ah receptor binding data, revealed that this dataset still retained a significant ability to predict the induction of AHH activity obtained from in vitro assays. The predicted AHH induction values are plotted versus the actual values in Figure 7.

CoMFA of in Vitro Ethoxyresorufin *O*-Deethylase (EROD) Induction. The following analyses were performed on the same series of compounds as used in the QSAR studies on AHH induction (see Tables I–III). The results are found in Table VI.

As is the case with the CoMFA of Ah receptor binding affinity and in vitro induction of AHH activity, the dibenzofuran dataset retained high predictive value. The values obtained for the compounds of the dibenzo-*p*-dioxin and biphenyl datasets were, however, much less significant and were classified as nonpredictive.

The CoMFA including all compounds produced a statistically significant but less predictive model than the QSAR developed based on the combination of dibenzo-*p*-dioxin and dibenzofuran molecules. This was expected since the biphenyl compounds were included in the calculations. In Figure 8, the predicted EROD induction values obtained from this analysis are plotted versus the actual values.

Discussion

For Ah receptor binding, the CoMFA-derived QSARs for all three classes of compounds and combinations thereof afforded strong crossvalidated correlations, indicating that they are highly predictive. The CoMFA-derived QSARs can thus be useful for predicting the binding affinities of untested compounds in all three classes of halogenated aromatic hydrocarbons. We have used the CoMFA-derived QSARs to predict the binding of the two heptachlorodibenzo-*p*-dioxins for which binding data was not available (see Table VII). It can be seen that somewhat different predictions are obtained depending on whether one uses the dioxin model alone, or models based on a combination of both the dioxins and dibenzofurans or all three classes of compounds. The dioxin alone model might be anticipated to perform better when predicting the activity of other dioxins. The relatively low binding affinity predicted (as compared to TCDD) is consistent with the much lower toxicity that has been reported[14] for the hepta dioxins.

For enzyme induction, the CoMFA-derived QSARs were highly predictive for the dibenzofurans (the class containing the largest number of compounds) and partially successful for the dioxins (particularly for AHH induction). For the biphenyls, the results were clearly unpredictive. The combination CoMFA-derived QSARs retained some significant predictability for AHH induction, but were somewhat less predictive for EROD induction.

The failure of the approach to predict the enzyme-inducing ability of the biphenyls is not entirely surprising in view of the well-documented effect of *o*-chlorine substitution on the overall conformational properties of the biphenyl molecule, particularly in raising the energy barrier to attain a coplanar state.[15] While in this work we permitted the biphenyls to attain an energy-minimized structure (see Table III), we were not able to improve the CoMFA-derived QSARs by assigning torsional angles based on experimentally determined measurements on related compounds (data not shown). In fact, ortho-substituted PCBs may represent a special case in which the induction potency is only partially related to Ah receptor binding affinity. On examining the predicted values, the dioxin QSAR in most cases appears to significantly overestimate the enzyme-inducing ability of the ortho-substituted PCBs. This is consistent with the fact that ortho-substituted PCBs are in every case considerably less toxic than their comparably substituted non-ortho-containing congeners. In fact, significant lethality as seen in laboratory animal studies with certain dioxins, furans and non-ortho-substituted PCBs has never been reported for any ortho-substituted PCB. In the multifunctional binding model that we previously proposed[16] for dioxin and related compound toxic action, we suggested that the conformationally restricted, noncoplanar ortho-

(14) McKinney, J. D.; McConnell, E. Structural Specificity and the Dioxin Receptor. In *Chlorinated Dioxin and Related Compounds. Impact on the Environmental*; Hutzinger, O., Frei, R. W., Merian, E., Pocchiari, F., Eds.; Pergamon: New York, 1982; Pergamon Series on Environmental Science Vol. 5, pp 367–381.

(15) McKinney, J. D.; Chae, K.; McConnell, E. E.; Birnbaum, L. S. Structure–induction versus Structure–toxicity Relationships for Polychlorinated Biphenyls and Related Aromatic Hydrocarbons. *Environ. Health Perspect.* **1985**, *60*, 57–68.

(16) McKinney, J. D. Multifunctional Receptor Model for Dioxin and Related Compound Toxic Action: Possible Thyroid Hormone-responsive Effector-linked Site. *Environ. Health Perspect.* **1989**, *82*, 323–336.

Table V. CoMFA QSAR Results of in Vitro AHH Induction

class(es)	obs	r^2_{cross}	r^2	SE	F test	p value	r^2_{bs}	SD
dioxins	13	0.424 (1)[a]	0.698	1.056	25.442	0.000	0.779	0.094
furans	30	0.695 (3)	0.848	0.667	48.278	0.000	0.887	0.047
biphenyls	9	0.273 (3)	0.926	0.518	21.002	0.003	0.978	0.022
dioxins and furans	43	0.647 (2)	0.781	0.800	72.372	0.000	0.789	0.038
dioxins and biphenyls	22	0.399 (2)	0.754	0.874	29.156	0.000	0.803	0.041
furans and biphenyls	39	0.592 (3)	0.813	0.713	50.758	0.000	0.815	0.048
dioxins, furans, and biphenyls	52	0.572 (4)	0.788	0.783	43.674	0.000	0.856	0.041

[a] Numbers in parentheses are optimal number of components.

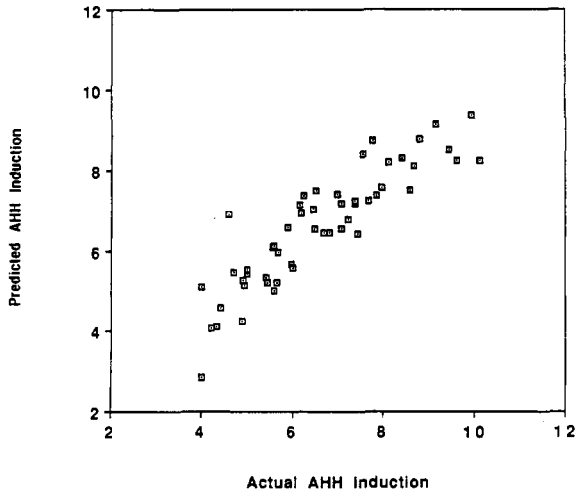

Figure 7. Predicted versus actual AHH induction values from CoMFA including all compounds.

substituted PCBs would experience more difficulty (in stereoelectronic and energetic terms) in accommodating simultaneously two different types of molecular recognition processes needed to effect a toxic response. Other explanations include the potential for partial agonist–antagonist activity as has been shown[17] for some PCBs, and the possibility that certain PCBs may potentiate the activity of some endogenous ligand such as thyroid hormone.[18] Thus, fundamentally different mechanisms may be operating in the expression of toxic responses for these ortho-substituted PCBs. Determining the dioxin-like toxic equivalency of any ortho-substituted PCB is not likely to be a straightforward process based on one molecular mechanism.

The contour plot of positive steric contribution shown in Figure 3 emphasizes the importance of lateral chlorine substitution in these classes of compounds for maximizing receptor binding and enzyme induction, whereas the contour plot of negative steric contribution (Figure 4) emphasizes the importance of nonlateral substitutions in lowering activity. These plots are qualitatively similar to the previous representations by Safe et al.,[7] indicating the importance (or lack thereof) of halogen substitutions in various positions on the ring system using a simple plus/minus notation. The overall structure alignments used for the three classes of compounds (see Figure 1) to maximize structural similarities is also similar to that we have previously suggested[19] on the basis of qualitative considerations of important structural features. The appropriateness of the alignment used is further supported by the success achieved with the QSAR for the binding data for all three compound classes.

It should be pointed out, however, that even though these CoMFA contour maps look like receptor maps, they are not directly comparable to them. In fact, the major contribution of lateral halogens may not be steric but more a reflection of the importance of lateral halogen substitution on overall stereoelectronic structure. In our previous stacking model for the Ah receptor interaction, the molecular polarizability of these basically rectangular shaped molecules is most affected by lateral halogens since the polarizability component in the lateral direction is usually dominant. The negative influence of nonlateral halogens as in 1,4,6,9 positions of dioxin may reflect a true repulsive steric interaction of some kind. Again in our multifunctional binding model[16] we had suggested that the 1,4,6,9 positions may be important in a perpendicular stacking interaction, and chlorine substitution in these positions would inhibit such an interaction. We also suggested that perpendicular stacking would require positive potential or "hole density" at the 1,4,6,9 positions as indicated in Figure 6. Our results complement those of previous workers[20] who reported that regions of negative electrostatic potential associated with the lateral positions of the molecule separated by an area of positive potential are necessary for activity. In the case of *ortho*-PCBs, additional steric considerations have to be made. One of the more important may be the substantial effect of ortho substitution on the energy barrier to achieve a coplanar condition needed to enhance a stacking interaction as a molecular recognition factor in the receptor binding domain. The CoMFA-derived QSARs are thus compatible with our previous models and provide a useful predictive ability for assessing toxic potential.

The above argument suggests a high degree of colinearity between the steric and electrostatic molecular fields of the molecules included in the analyses. With respect to Ah receptor binding affinity, the relative contributions of the steric and electrostatic terms to the QSAR equation (model) including all molecules were approximately equivalent. This indicates that both sets of regressors are important, but yields no information concerning colinearity. When either the steric or electrostatic terms are eliminated from the model, the crossvalidated regression coefficient is decreased while the complexity of the model (as measured by the optimal number of components) is increased. The exclusive steric term model yielded a value of 0.721 with nine components; the electrostatic model yielded 0.660 with eight components. Although the similarly high crossvalidated regression coefficients in-

(17) Davis, D.; Safe, S. Dose–response Immunotoxicities of Commercial Polychlorinated Biphenyls (PCBs) and Their Interaction with 2,3,7,8-Tetrachlorodibenzo-p-dioxin. *Toxicol. Lett.* **1989**, *48*, 35–43.

(18) Lamb, J. C., IV, Harris, M. W.; McKinney, J. D.; Birnbaum, L. S. Interaction of 2,3,7,8-Tetrachlorodibenzo-p-dioxin (TCDD) with Thyroid Hormones in the Induction of Cleft Palate in C57BL/6N Mice. *Toxicol. Appl. Pharmacol.* **1986**, *84*, 115–124.

(19) McKinney, J. D.; Singh, P. Structure–activity Relationships in Halogenated Biphenyls. Unifying Hypothesis for Structural Specificity. *Chem.-Biol. Interact.* **1981**, *33*, 271–283.

(20) Sjoberg, P.; Murray, J. S.; Brinck, T.; Evans, P.; Politzer, P. The Use of Electrostatic Potential at the Molecular Surface in Recognition Interactions: Dibenzo-p-dioxins and Related Systems. *J. Mol. Graphics* **1990**, *8*, 81–90.

Table VI. CoMFA QSAR Results of in Vitro EROD Induction

class(es)	obs	r^2_{cross}	r^2	SE	F test	p value	r^2_{bs}	SD
dioxins	13	0.232 (1)	0.633	1.062	19.009	0.001	0.767	0.098
furans	30	0.663 (3)	0.838	0.758	44.718	0.000	0.876	0.051
biphenyls	9	0.021 (1)	0.490	1.053	5.738	0.036	0.655	0.155
dioxins and furans	43	0.579 (2)	0.741	0.904	57.086	0.000	0.768	0.052
dioxins and biphenyls	22	0.250 (2)	0.688	0.901	20.944	0.000	0.781	0.080
furans and biphenyls	39	0.553 (3)	0.786	0.812	42.915	0.000	0.797	0.042
dioxins, furans, and biphenyls	52	0.489 (4)	0.745	0.876	34.416	0.000	0.832	0.042

[a] Numbers in parentheses are optimal number of components.

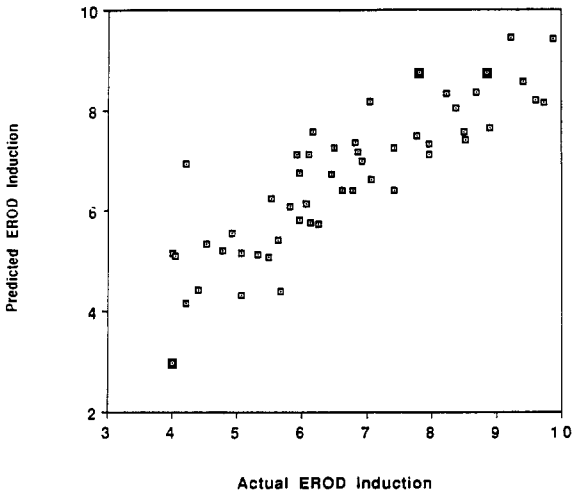

Figure 8. Predicted versus actual EROD induction values from CoMFA including all compounds.

Table VII. Predicted Binding Affinity Values for Heptachlorodibenzo-p-dioxins

compd	prediction 1[a]	prediction 2[b]	prediction 3[c]
1234678	5.696	6.223	6.351
1234679	4.123	4.442	4.549

[a] Prediction 1 based on CoMFA for dioxins. [b] Prediction 2 based on CoMFA for dioxins and furans. [c] Prediction 3 based on CoMFA for dioxins, furans, and biphenyls.

dicate that the terms are colinear and suggest the exclusion of either the steric or electrostatic set of terms from the model, the increased complexity in both cases supports the use of the overall model which yielded a crossvalidated regression coefficient of 0.724 from an equation based on only six components.

It is possible that QSAR models of this type could be expanded such that in vivo toxicity data could be predicted. This would necessitate the inclusion of additional physical parameters as regressors in the model. One such parameter would provide information concerning the partitioning of the molecules within particular body compartments, etc. Traditionally, octanol:water partition coefficients (P) have been used to estimate this physical characteristic.[21] However, experimental log P values are only available for a limited number of analogs within the series of interest. Therefore, a computational alternative, the clog P technique,[22] must be utilized to provide this data. In preliminary studies in our laboratory, clog P proved to be a detrimental regressor in that it decreased the significance of the QSAR model in every case (data not shown). The theory behind the generation of log P values using the clog P approach could provide one possible explanation for this. Since clog P is a fragment approach, it is incapable of calculating differences in log P values with respect to substituent patterns. Thus, all analogs within a particular series (PCDD, PCDF, or PCB) possessing the same number and type of halogen substituent are calculated to have the same log P. This is clearly not the case. For this reason, in the absence of experimental data, parameters such as polarizability and desolvation energies, as well as additional computational techniques for calculating log P, are being explored as possible alternatives to the clog P method to provide partitioning data.

In future studies, we plan to examine the potential of CoMFA to reveal important structural features for various Ah receptor antagonists and chlorinated biphenyl ether agonists. In the latter case, we would hope to reveal structural similarities of the active compounds to those found to be important in the expression of thyroid hormone activity.

Acknowledgment. We thank the TRIPOS Associates for a generous grant of the SYBYL 5.41 molecular modeling software to the Laboratory for Molecular Modeling located in the School of Pharmacy at the University of North Carolina in Chapel Hill.

Supplementary Material Available: Coordinates of all molecular model conformations and region definitions discussed in this work are available (88 pages). Ordering information is given on any current masthead page.

Registry No. TCDD, 1746-01-6; PCDD2, 40321-76-4; PCDD3, 34816-53-0; PCDD4, 82306-65-8; PCDD5, 39227-28-6; PCDD6, 50585-46-1; PCDD7, 58802-08-7; PCDD8, 30746-58-8; PCDD9, 33857-28-2; PCDD10, 38964-22-6; PCDD11, 39227-61-7; PCDD12, 39227-58-2; PCDD13, 39227-53-7; PCDD14, 3268-87-9; PCDD18, 50585-41-6; PCDD19, 50585-40-5; PCDD20, 109333-32-6; PCDD21, 109333-33-7; PCDD22, 143106-17-6; PCDD23, 109333-31-5; PCDD24, 109333-35-9; PCDD25, 109333-34-8; PCDD26, 51974-40-4; PCDD27, 39073-07-9; PCDD28, 105906-36-3; PCDF1, 51230-49-0; PCDF2, 25074-67-3; PCDF3, 74992-96-4; PCDF4, 64126-86-9; PCDF5, 60390-27-4; PCDF6, 5409-83-6; PCDF7, 83704-39-6; PCDF8, 76621-12-0; PCDF9, 57117-34-7; PCDF10, 57117-32-5; PCDF11, 83704-45-4; PCDF12, 83704-30-7; PCDF13, 83704-32-9; PCDF14, 71998-72-6; PCDF15, 51207-31-9; PCDF16, 64126-87-0; PCDF17, 83704-50-1; PCDF18, 71998-74-8; PCDF19, 67517-48-0; PCDF20, 57117-41-6; PCDF21, 58802-15-6; PCDF22, 57117-31-4; PCDF23, 70648-26-9; PCDF24, 57117-44-9; PCDF25, 67562-40-7; PCDF26, 60851-34-5; PCDF27, 57117-37-0; PCDF28, 83704-21-6; PCDF29, 83704-22-7; PCDF30, 58802-16-7; PCDF31, 70648-21-4; PCDF32, 83704-53-4; PCDF34, 83704-31-8; PCDF35, 83704-22-7; PCDF36, 58802-16-7; PCDF37, 70648-21-4; PCDF38, 83704-53-4; PCDF39, 69698-57-3; PCB2, 70362-50-4; PCB3, 57465-28-8; PCB4, 65510-44-3; PCB5, 32598-14-4; PCB6, 31508-00-6; PCB7, 74472-37-0; PCB8, 38380-08-4; PCB9, 52663-72-6; PCB10, 69782-90-7; PCB11, 2437-79-8; PCB12, 35065-27-1; PCB13, 33284-53-6; PCB14, 59291-65-5; octachlorodibenzofuran, 39001-02-0; 33'44'-tetrachlorobiphenyl, 32598-13-3; 1,2,3,4,6,7,8-heptachlorodibenzo-p-dioxin, 35822-46-9; 1,2,3,4,6,7,9-heptachlorodibenzo-p-dioxin, 58200-70-7; aromatic hydrocarbon hydroxylase, 9037-52-9; ethoxyresorufin O-deethylase, 59793-97-4.

(21) Coats, E. Role of the Hansch Partition Coefficient as a Descriptor of Drug Transport and Receptor Interaction in Correlation Analyses. *Physical Chemical Properties of Drugs*; Yalkowsky, S. H., Sinkula, A. A., Valvani, S. C., Eds.; Marcel Dekker, Inc.: New York, 1980; pp 111–137.
(22) Leo, A.; Hansch, C.; Elkins, D. Partition Coefficients and their uses. *Chem. Rev.* 1971, *71*, 525–616.