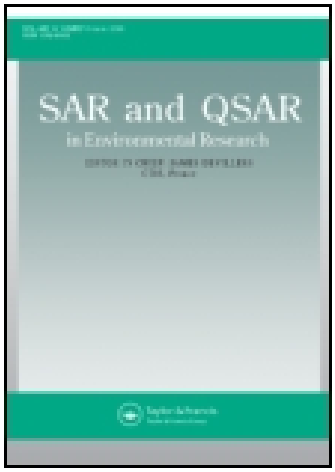
2002

Reactivity Profiles of Ligands of Mammalian Retinoic Acid Receptors: A Preliminary COREPA Analysis

Steven P. Bradbury
G. T. Ankley
O. G. Mekenyan
V. . Kamenska
P. K. Schmieder



SAR and QSAR in Environmental Research
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/gsar20

Reactivity profiles of ligands of mammalian retinoic acid receptors: A preliminary COREPA analysis

G.T. Ankley [a] , O.G. Mekenyan [b] , V.B. Kamenska [b] , P.K. Schmieder [a] & S.P. Bradbury [a]
[a] National Health and Environmental Effects Research Laboratory, Mid-Continent Ecology Division , U.S. Environmental Protection Agency , 6201 Congdon Boulevard, Duluth, MN, 55804, USA
[b] Laboratory of Mathematical Chemistry, Department of Physical Chemistry , Bourgas University "Prof. As. Zlatarov" , Bourgas, 118010, Bulgaria
Published online: 29 Oct 2010.

To cite this article: G.T. Ankley , O.G. Mekenyan , V.B. Kamenska , P.K. Schmieder & S.P. Bradbury (2002) Reactivity profiles of ligands of mammalian retinoic acid receptors: A preliminary COREPA analysis, SAR and QSAR in Environmental Research, 13:2, 365-377, DOI: 10.1080/10629360290002839

To link to this article: http://dx.doi.org/10.1080/10629360290002839

SAR and QSAR in Environmental Research, 2002 Vol. 13 (2), pp. 365–377

REACTIVITY PROFILES OF LIGANDS OF MAMMALIAN RETINOIC ACID RECEPTORS: A PRELIMINARY COREPA ANALYSIS*

G.T. ANKLEY[a,†], O.G. MEKENYAN[b], V.B. KAMENSKA[b], P.K. SCHMIEDER[a] and S.P. BRADBURY[a]

[a]*U.S. Environmental Protection Agency, National Health and Environmental Effects Research Laboratory, Mid-Continent Ecology Division, 6201 Congdon Boulevard, Duluth, MN 55804 USA;* [b]*Laboratory of Mathematical Chemistry, Department of Physical Chemistry, Bourgas University "Prof. As. Zlatarov", 118010 Bourgas, Bulgaria*

(Received 19 September 2000; In final form 28 April 2001)

Retinoic acid and associated derivatives comprise a class of endogenous hormones that bind to and activate different families of retinoic acid receptors (RARs, RXRs), and control many aspects of vertebrate development. Identification of potential RAR and RXR ligands is of interest both from a pharmaceutical and toxicological perspective. The recently developed COREPA (COmmon REactivity PAttern) algorithm was used to establish reactivity profiles for a limited data set of retinoid receptor ligands in terms of activation of three RARs (α, β, γ) and an RXR (α). Conformational analysis of a training set of retinoids and related analogues in terms of thermodynamic stability of conformers and rotational barriers showed that these chemicals tend to be quite flexible. This flexibility, and the observation that relatively small energy differences between conformers can result in significant variations in electronic structure, highlighted the necessity of considering all energetically reasonable conformers in defining common reactivity profiles. The derived reactivity patterns for three different subclasses of the RAR (α, β, γ) were similar in terms of their global electrophilicity (nucleophilicity) and steric parameters. However, the profile of active chemicals with respect to interaction with the RXR-α differed qualitatively from that of the RARs. Variations in reactivity profiles for the RAR versus RXR families would be consistent with established differences in their affinity for endogenous retinoids, likely reflecting functional differences in the receptors.

Keywords: Retinoid; Receptor; Transactivation; Model

INTRODUCTION

Vitamin A (retinol) is the parent molecule for a number of physiologically important retinoids, including all-*trans* retinoic acid (RA) and 9-*cis* RA. These compounds act as hormones, interacting with different retinoid receptors which serve as transcription factors controlling a number of critical processes in early vertebrate development, including cellular differentiation and axial patterning [1]. Two retinoid receptor families have been identified: the retinoic acid receptors (RARs), which bind all-*trans* RA and 9-*cis* RA, and the retinoid-X

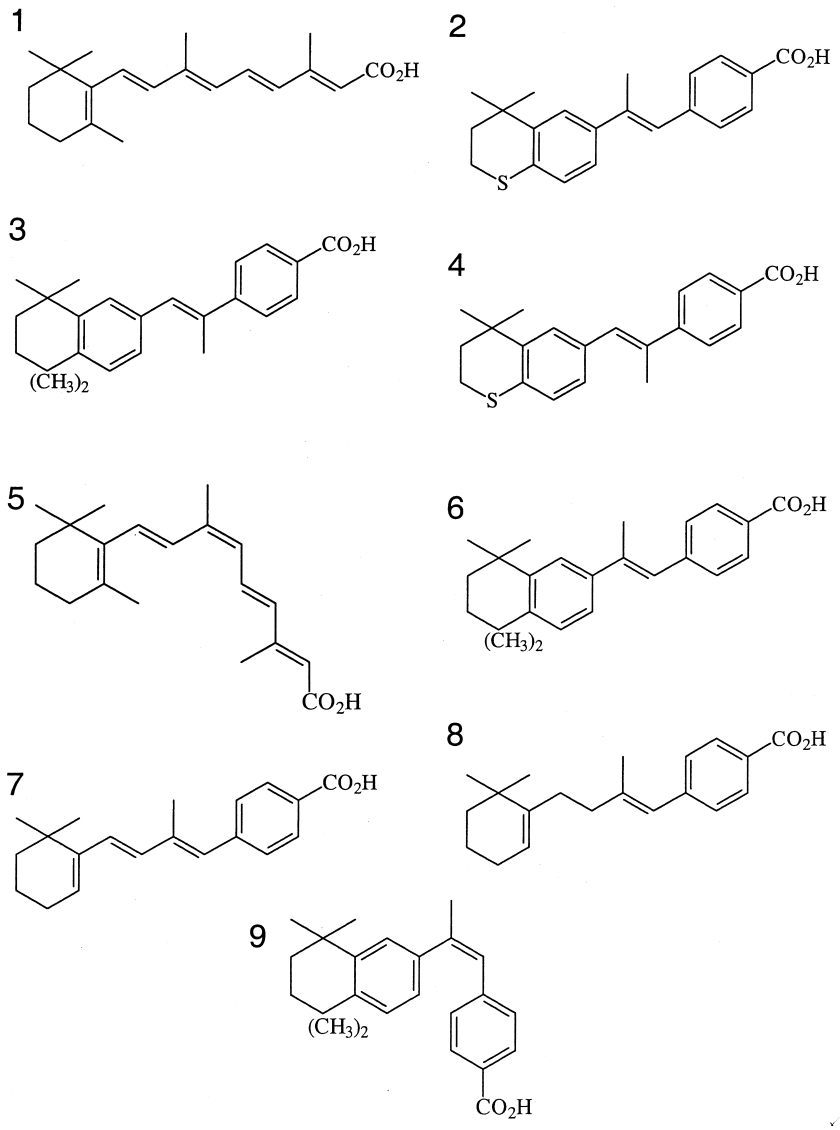
*Presented at the 9th International Workshop on Quantitative Structure–Activity Relationships in Environmental Sciences (QSAR 2000), 16–20 September 2000, Bourgas, Bulgaria.

†Corresponding author. E-mail: ankley.gerald@epa.gov

ISSN 1062-936X
DOI: 10.1080/10629360290002839

receptors (RXRs) which bind only 9-*cis* RA [1]. At least three different isoforms, α, β, γ, have been identified for both the RAR and RXR [1]. Although RARs and RXRs are similar from a structural perspective, they differ in spatial and temporal expression, and DNA binding affinities, indicating unique roles for the different receptors during development [1].

Because of their role in cellular proliferation and differentiation, retinoids have received considerable attention from the standpoint of therapeutic treatment of different diseases, including cancer [2,3]. In addition, there recently has been concern that contaminants acting as retinoid mimics in the environment might be responsible for adverse effects in wildlife, such as malformations in amphibians [4,5]. Consequently, identification of chemicals that interact with the retinoid receptors is important both from a pharmaceutical and toxicological perspective. The purpose of this study, therefore, was to conduct an evaluation of ligands which activate RARs and RXRs using the COREPA (COmmon REactivity PAttern)



algorithm [6,7]. The COREPA approach attempts to define chemical similarity in terms of reactivity patterns based on global and local chemical descriptors potentially associated with the biological activity of concern (in this case, RAR and RXR activation). The approach does not depend on identification of a pre-defined pharmacophore and it explicitly considers conformational flexibility of xenobiotics, which is a critical aspect of modeling interactions of ligands with hormone receptors. To complement results of previous studies concerned with structure–activity relationships of retinoids [8–11], we applied the COREPA technique in an attempt to assess the role of flexibility in predicting active ligands. The data set used for this analysis [12] was somewhat limited with respect to structural diversity of the test chemicals; hence, our results should be regarded as exploratory.

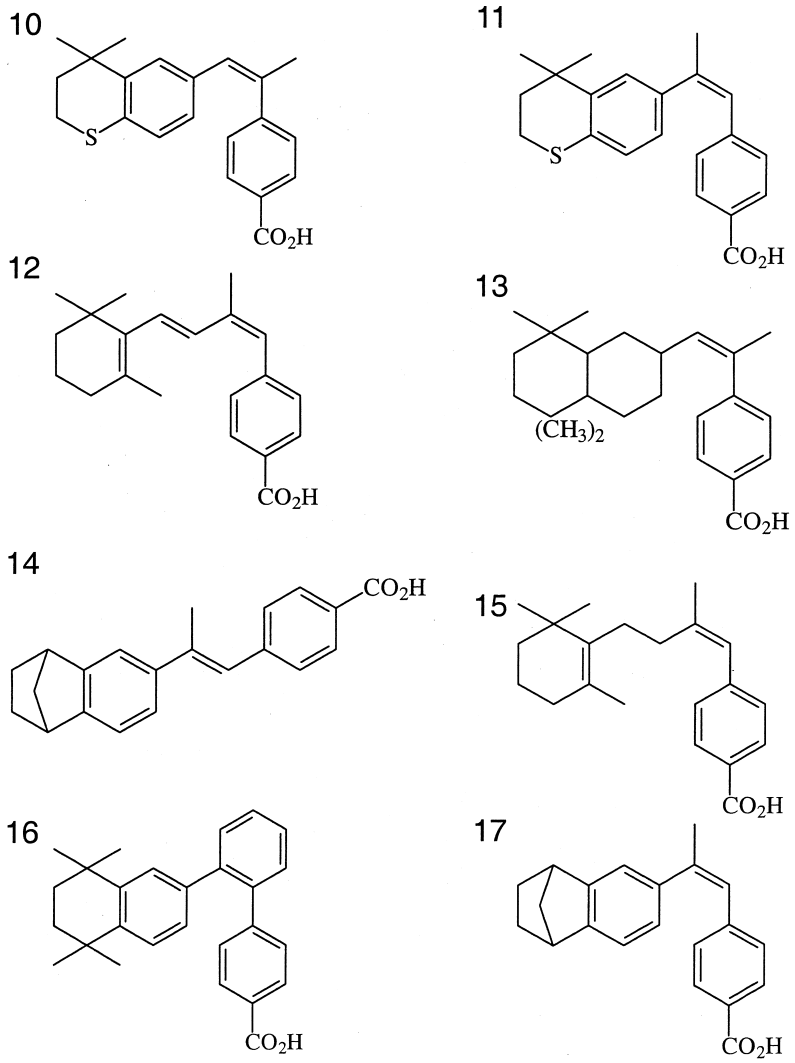


FIGURE 1 Structures of ligands used to establish training sets for COREPA analysis of activation of retinoic acid receptors (see Table I for compound names [12]).

MATERIAL AND METHODS

The data set used for this analysis was from Jong *et al.* [12], who evaluated transcriptional activity in CV1 cells treated with 17 different potential retinoid receptor agonists (Fig. 1; Table I). The cells were transiently transfected with a plasmid containing a reporter (chloramphenicol acetyl transferase)-receptor gene construct with either the RAR-α, RAR-β, RAR-γ, or RXR-α. Relative activity of the ligands was expressed as the concentration necessary to elicit 50% of maximal activity (EC50) compared with 10^{-5} M all-*trans* RA (RARs) or 9-*cis* RA (RXR). For modeling purposes, transactivation data were expressed as log(1/EC50). In generating common reactivity profiles for chemicals of interest it is necessary to categorize them as to relative activity. In the present study, different thresholds for defining relative activity of the test chemicals were explored, but herein we report results only of analyses in which chemicals with a $\log(1/\mathrm{EC50}) > -3$ were designated as highly active, while those with values lower than this were defined as having low activity. Selection of this threshold is arbitrary in the sense that there is no inferred biological significance. The purpose of this modeling effort, however, was to demonstrate a conceptual application of the COREPA method rather than develop a biologically relevant screening model such as described for ligands of the estrogen receptor [13–15].

Detailed descriptions of the conceptual basis and mathematical formalism of the COREPA method, and its application to analysis of data of the type used in the present study can be found elsewhere [6,7,13,14]. Briefly, 3D structures of conformers of the test chemicals were generated with the method of Ivanov *et al.* [16] using parameters such as torsion resolution, distance between non-bonded atoms, and ring closure [6,7,13]. Geometry optimization of the conformers was obtained with MOPAC 93 using the AM1 Hamiltonian [17,18]. For any given ligand, only conformers with ΔH_f^0 values within 20 kcal/mol of the value of the minimum energy level conformer were utilized [6,7,13,19]. Based on experience from previous analyses [6,7,13,14], molecular descriptors potentially associated with receptor binding/gene activation were calculated for each conformer of interest. These included global and local electronic parameters (e.g. electronegativity, energy of frontier orbitals, atomic charges, and self-polarizabilites, etc.) and steric descriptors (e.g. interatomic distances, planarity, steric distances between atoms, etc.).

The COREPA algorithm is comprised of three steps. First, two subsets of chemicals for each receptor were identified as training sets, representing high versus low activity (i.e. for the chemicals in this study, $\log(1/\mathrm{EC50}) > -3$ and < -3, respectively). Second, a subset of the molecular descriptors that served as the best predictors of retinoid receptor activation were established by evaluating the degree of overlap between the distributions of descriptor values associated with conformers of chemicals with high versus low activity. The descriptors were evaluated based on the normalized sum of dynamic similarity indices between each pair of molecules in the training set [7]. Those molecular descriptors that resulted in the maximal measure of similarity among chemicals in a training set, and least overlap between the high versus low activity training sets, were assumed to be related to transcriptional activation of the retinoid receptor(s). In the final step of the algorithm, common reactivity patterns for biologically similar molecules were obtained as products of the probabilistic distributions for specific molecular descriptors associated with chemicals with high versus low retinoid receptor activation. The width of these ranges is dependent upon values of Γ, which is related to the half-width of the gamma function and confidence limits chosen around the probability maxima. Based on previous studies [7,13], default Γ values of 0.1–0.125 and 0.01–0.05 of the respective variation ranges for global and local descriptors were utilized. Stability and specificity of the resultant patterns were assessed

TABLE I Retinoic acid receptor ligands, transactivation values for RAR-α, RAR-β, RAR-γ, and RXR-α [12], conformer numbers (N) and parameter ranges for the stereoelectronic parameters E_{LUMO}, E_{HOMO} (energy of the lowest unoccupied and highest occupied molecular orbitals), and maximum distance (L_{max}). Also shown are root mean square (RMS) values indicating the relative degree of flexibility of the test chemical, and the $\Delta\Delta H_f^0$ range for multiple conformers of a given chemical

No.	*Test chemical*	*N*	*RAR-α*	*RAR-β*	*RAR-γ*	*RXR-α*	E_{LUMO} (eV)	E_{HOMO} (eV)	L_{max} (Å)	*RMS*	$\Delta\Delta H_f^0$ (kcal/mol)
1	(*E*)-*trans*-Retinoic acid	31	−1.431	−1.079	−0.903	−2.633	−1.1162 to −0.8540	−8.9275 to −8.2853	14.6893–15.8759	0.341–6.863	−61.4298 to −52.6660
2	(*E*)-6-[1-(4-carboxyphenyl)propen-2-yl]-3,4-dihydro-4,4-dimethyl-2H-1-benzothiopyrans	12	−1.431	−1.176	−1.477	−4.000	−0.8779 to −0.7928	−8.3500 to −8.1197	14.8410–15.1634	0.884–2.968	−42.3241 to −33.8797
3	(*E*)-4-[1-(5,6,7,8-tetrahydro-5,5,8,8-tetramethyl-2-naphthalenyl)propen-2-yl] benzoic acid	17	−1.568	−1.398	−1.279	−4.000	−0.8402 to −0.6598	−8.9141 to −8.7198	14.8227–15.1717	0.420–3.814	−59.8331 to −47.9838
4	(*E*)-6-[2-(4-carboxyphenyl)propen-1-yl]-3,4-dihydro-4,4-dimethyl-2H-1-benzothiopyrans	24	−1.653	−1.602	−1.672	−4.000	−0.9579 to −0.7597	−8.3745 to −8.0901	14.7234–15.2003	0.493–2.663	−42.3106 to −29.5416
5	(9*Z*)-Retinoic acid	16	−1.653	−1.342	−1.079	−1.114	−0.9032 to −0.4566	−8.9191 to −8.6924	13.4985–14.8644	0.563–9.068	−55.3704 to −47.5900
6	(*E*)-4-[2-(5,6,7,8-tetrahydro-5,5,8,8-tetramethyl-2-naphthalenyl)propen-1-yl] benzoic acid	9	−1.929	−0.778	−1.380	−4.000	−0.7801 to −0.6484	−9.0038 to −8.9185	14.9748–15.0458	0.286–3.262	−59.9323 to −58.0525
7	(*E*)-4-[2-methyl-4-(2,6,6-trimethylcyclohexen-1-yl)buta-1,3-dien-1-yl] benzoic acid	11	−2.380	−0.699	−1.544	−4.000	−0.9243 to −0.7894	−9.0091 to −8.6939	14.8669–15.2096	0.462–3.749	−58.1576 to −51.5342
8	(*E*)-4-[2-methyl-4-(2,6,6-trimethylcyclohexen-1-yl) but-1-en-1-yl] benzoic acid	24	−2.740	−1.301	−2.322	−3.176	−0.8460 to −0.5744	−9.2251 to −9.0942	14.7207–15.1247	0.000–4.739	−84.2581 to −73.2076
9	(*E*)-4-[2-(5,6,7,8-tetrahydro-5,5,8,8-tetramethyl-2-naphthalenyl)propen-1-yl] benzoic acid	9	−3.000	−3.342	−2.845	−3.431	−0.7418 to −0.6224	−9.2325 to −8.9440	11.5412–12.6841	0.310–10.602	−59.0121 to −53.8529
10	(*Z*)-6-[2-(4-carboxyphenyl)propen-2-yl]-3,4-dihydro-4,4-dimethyl-2H-1-benzothiopyrans	10	−3.322	−3.176	−3.176	−3.447	−0.6880 to −0.5667	−8.2481 to −8.0130	11.9012–12.7928	0.327–4.603	−41.2522 to −35.4526
11	(*Z*)-6-[1-(4-carboxyphenyl)propen-2-yl]-3,4-dihydro-4,4-dimethyl-2H-1-benzothiopyrans	12	−3.362	−2.398	−3.041	−4.000	−0.7498 to −0.6873	−8.2880 to −8.1158	10.7995–12.4613	1.394–10.594	−41.5042 to −36.6936
12	(1*Z*,3*E*)-4-[2-methyl-4-(2,6,6-trimethylcyclohexen-1-yl)buta-1,3-dien-1-yl] benzoic acid	21	−3.699	−3.255	−3.301	−4.000	−0.8299 to −0.6945	−8.8867 to −8.6476	11.5709–12.5671	0.321–9.800	−58.4543 to −53.6223
13	(*Z*)-4-[1-(5,6,7,8-tetrahydro-5,5,8,8-tetramethyl-2-naphthalenyl)propen-2-yl] benzoic acid	9	−4.000	−3.279	−3.431	−3.041	−0.6209 to −0.5587	−8.9574 to −8.8488	12.1454–12.7513	0.300–6.848	−58.7794 to −55.8783
14	(*E*)-4-[2-(1,4-methano-1,2,3,4-tetrahydro-naphthalen-6-yl)propen-1-yl]-benzoic acid	4	−4.000	−3.041	−2.740	−4.000	−0.7879 to −0.7752	−8.9626 to −8.9213	14.5568–14.6100	0.527–4.166	−17.3145 to −17.1504
15	(*Z*)-4-[2-methyl-4-(2,6,6-trimethylcyclohexen-1-yl) but-1-en-1-yl] benzoic acid	10	−4.000	−3.301	−3.114	−4.000	−0.8434 to −0.6165	−9.1791 to −9.0753	10.6710–13.5267	1.030–7.948	−83.8797 to −73.2739
16	(*E*)-4-[2-(5,6,7,8-tetrahydro-5,5,8,8-tetramethyl-2-naphtalenyl)-phenyl]benzoic acid	25	−4.000	−3.505	−2.000	−3.447	−0.6956 to −0.6633	−9.2680 to −9.0336	11.4553–12.0623	0.277–14.280	−39.3447 to −30.1206
17	(*E*)-4-[2-(1,4-methano-1,2,3,4-tetrahydro-naphthalen-6-yl)propen-1-yl]-benzoic acid	6	−4.000	−4.000	−1.845	−4.000	−0.7336 to −0.6490	−9.1432 to −8.9432	11.1558–12.0532	0.538–6.205	−16.4324 to −15.3742

TABLE II Averaged within-group similarity between chemicals classified as high ($\log(1/EC50) > -3$) versus low ($\log(1/EC50) < -3$) activity with respect to activation of four retinoic acid receptors. Between-group similarity for this same activity range also is indicated. Representative descriptors include E_{HOMO}, E_{LUMO}, electronic gap ($E_{HOMO-LUMO}$), L_{max}, sum of geometric distances (GW), electronegativity (EN), dipole moment (D), distance between carboxylic oxygen and an unsaturated double bond in the cyclic moiety (d(RX_O;RX:R{scy}=R{scy},R{scy}*R{scy}), and the distance between carboxylic oxygen and sp^3-hybridized C-atoms (d(O_C{sp^3}))

		Within-group similarity (%)											
		RAR-α		*RAR-β*		*RAR-γ*		*RXR-α*		*Between-group similarity (%)*			
Descriptor	Γ	*High*	*Low*	*High*	*Low*	*High*	*Low*	*High*	*Low*	*RAR-α*	*RAR-β*	*RAR-γ*	*RXR-α*
E_{HOMO} (eV)	0.157	36.98	15.10	34.56	17.29	41.32	11.28	48.78	14.74	64.19	29.29	38.82	14.11
E_{LUMO} (eV)	0.082	55.34	11.46	55.88	11.49	56.56	11.26	34.29	11.55	17.27	12.07	21.18	7.92
$E_{HOMO-LUMO}$ (eV)	0.18	37.56	17.36	35.39	19.74	41.01	13.43	34.88	16.83	55.38	23.30	34.45	11.29
L_{max} (Å)	0.65	67.36	55.62	67.24	56.01	52.18	74.55	44.48	42.52	0.0	0.0	0.0	46.08
GW (Å)	536.8	38.03	100	33.96	100	30.24	100	43.30	100	0.61	17.40	3.53	57.58
EN (eV)	0.084	47.0	8.96	43.38	10.33	51.61	7.71	94.83	9.57	74.03	50.83	19.76	31.98
D (μ)	0.35	64.26	39.75	63.45	42.00	62.71	39.53	51.62	41.68	18.67	24.66	11.31	25.84
d(RX_O) (Å)	1.6	61.94	95.29	61.94	95.53	63.43	98.22	54.94	91.67	42.86	42.77	1.48	0.069
d(O_C{sp^3}) (Å)	1.86	78.71	92.99	74.86	93.65	67.79	98.77	79.60	86.05	1.57	1.57	4.42	32.0

TABLE III Variation of parameter ranges (10% confidence limits) of relevant molecular descriptors for activation of RAR-α, RAR-β, RAR-γ, and RXR-α, associated with patterns of active chemicals ($\log(1/EC50 > -3)$)

		Parameter ranges			
Descriptor	Γ	*RAR-α*	*RAR-β*	*RAR-γ*	*RXR-α*
E_{LUMO}	0.090 eV	−0.792 to −0.765	−0.779 to − 0.771	−0.763 to −0.756	−0.886 to −0.868
E_{HOMO}	0.338 eV	−8.834 to −8.811	−8.806 to −8.780	−8.937 to −8.920	−8.757 to −8.716
L_{max}	0.929 Å	14.99–15.03	14.97–15.01	14.91–14.94	14.85–14.96
Distance, d(RX_O); RX:R{scy}=R{scy},R{scy}=R{scy}	1.448 Å	11.22–11.30	11.17–11.27	11.19–11.26*	11.58–11.75
Distance, d(O_C{sp^3})	1.699 Å	12. 93–13.11	12.10–12.26	11.97–12.10	12.64–12.79
Distance, d(O_C{sp^3})	1.000 Å	12.04–12.23	11.97–12.09	11.94–12.05	
		13.25–13.39	13.11–13.37	12.86–13.30	12.63–12.95

* Distance range >5 Å was analyzed.

using statistical estimates based on an Euclidean distance metric and "leave-one-out" analyses [7,13].

RESULTS AND DISCUSSION

Table I indicates the number of conformers generated for each of the test chemicals within the specified ΔH_f^0 range of 20 kcal/mol. The number of conformers in this range varied from four for chemical **14** to 31 for chemical **1**. For any given chemical, the range in descriptor values for its ensemble of conformers was often quite large. For example, conformers of chemical **1** (all-*trans* RA), that differed within a ΔH_f^0 range of 8.764 kcal/mol, had a range of 0.262 eV for E_{LUMO} and 0.642 eV for E_{HOMO} (energies of the lowest unoccupied and highest occupied molecular orbitals, respectively). The observation of significant variations in descriptor values of thermodynamically and kinetically reasonable conformers is similar to

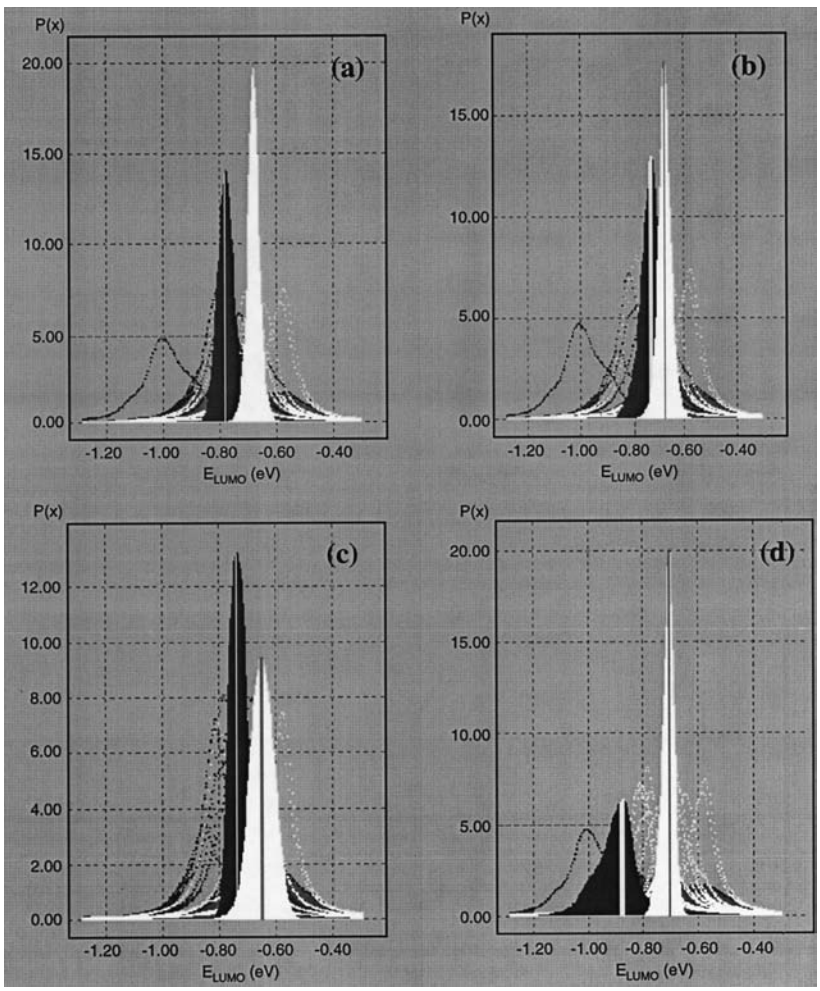


FIGURE 2 The reactivity patterns based on E_{LUMO}, at Γ(default) = 0.090 eV, for: RAR-α (a), RAR-β (b), RAR-γ (c), and RXR-α (d). The integral reactivity pattern of highly active ligands is dark, whereas the pattern of low activity ligands is white.

what we have found in other studies with ligands of steroid hormone receptors [6,7,13,19,20], and is of critical importance from a modeling perspective. Specifically, our ΔH_f^0 cut-off of 20 kcal/mol was selected based on experimental data indicating that the free energy of binding of hormones to steroid receptors is in the range of -10 to -20 kcal/mol [21–23], which can provide the necessary energy to elevate conformer(s) from the lowest energy state during binding. As a consequence, modeling approaches that focus only on the lowest energy conformers as those responsible for biological activity could yield erroneous results.

The degree of conformational flexibility of chemicals in this study can be interpreted in terms of the magnitude of the range of root mean square (RMS) differences between atoms of the conformers for each chemical within the configuration providing maximal alignment, based on comparisons to conformers with the lowest energy structure (Table I). Consistent with their comparatively lesser rigidity, larger RMS ranges were associated with the mono-cyclic than the multi-cyclic molecules. For example, RMS ranges of 0.341–6.863 and 0.563–9.068 were derived for all-*trans* RA (**1**) and 9-*cis* RA (**5**), respectively, while ranges of

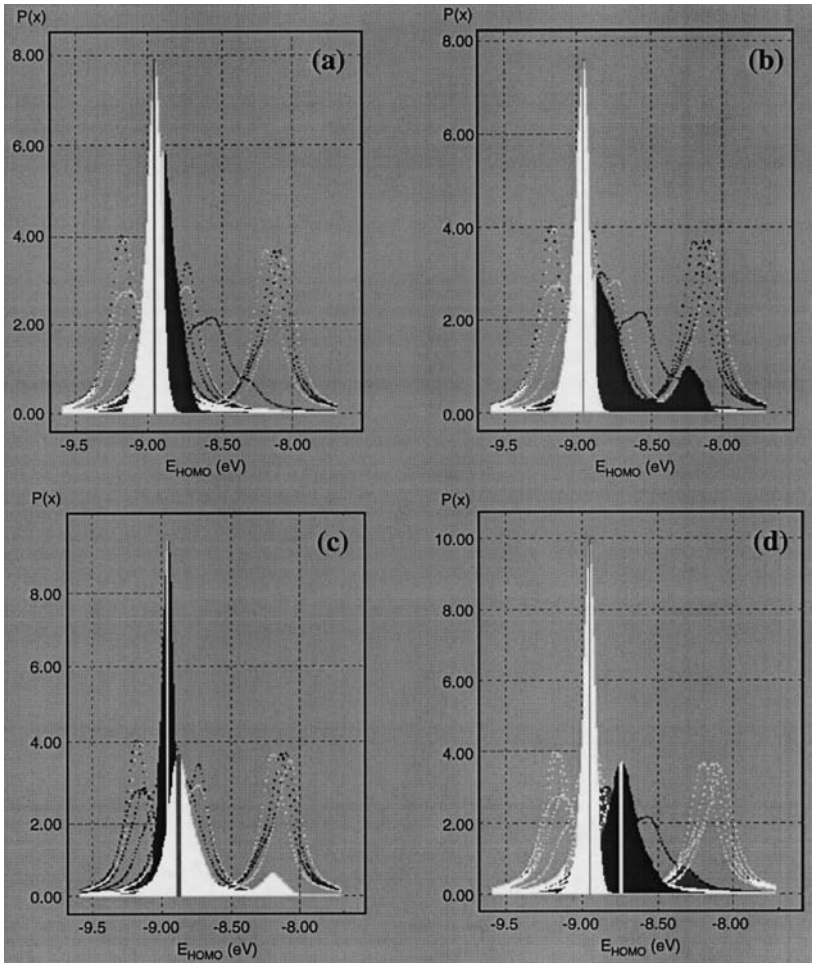


FIGURE 3 The reactivity patterns based on E_{HOMO}, at Γ(default) = 0.338 eV, for: RAR-α (a), RAR-β (b), RAR-γ (c), and RXR-α (d). The integral reactivity pattern of highly active ligands is dark, whereas the pattern of low activity ligands is white.

0.286–3.262 and 0.527–4.166 were obtained for (*E*)-4-[2-(5,6,7,8-tetrahydro-5,5,8,8-tetramethyl-2-naphthalenyl)propen-1-yl] benzoic acid (**6**) and (*E*)-4-[2-(1,4-methano-1,2,3,4-tetrahydronaphthalen-6-yl)propen-1-yl] benzoic acid (**14**).

Table II summarizes, for the four receptors, an evaluation of the average within-group similarity of chemicals classified as having low versus high activity for a representative set of descriptors. Also shown in Table II are between-group similarity values between chemicals in the high versus low activity ranges. Based on these results, the global parameters E_{LUMO}, E_{HOMO}, and L_{max} (greatest interatomic distance), and two local distance parameters—distance between carboxylic oxygen and an unsaturated bond in the cyclic moiety (d(RX_O);RX:R{scy} = R{scy},R{scy}∗R{scy}) and distance between carboxylic oxygen and sp^3-hybridized C-atoms (d(O_C{sp^3})—were selected as model descriptors.

Figures 2–7(a)–(d) depict reactivity patterns associated with the descriptors that provided the best separation between chemicals with high versus low activity in the four retinoid receptor systems, while Table III summarizes ranges in descriptor values for those chemicals

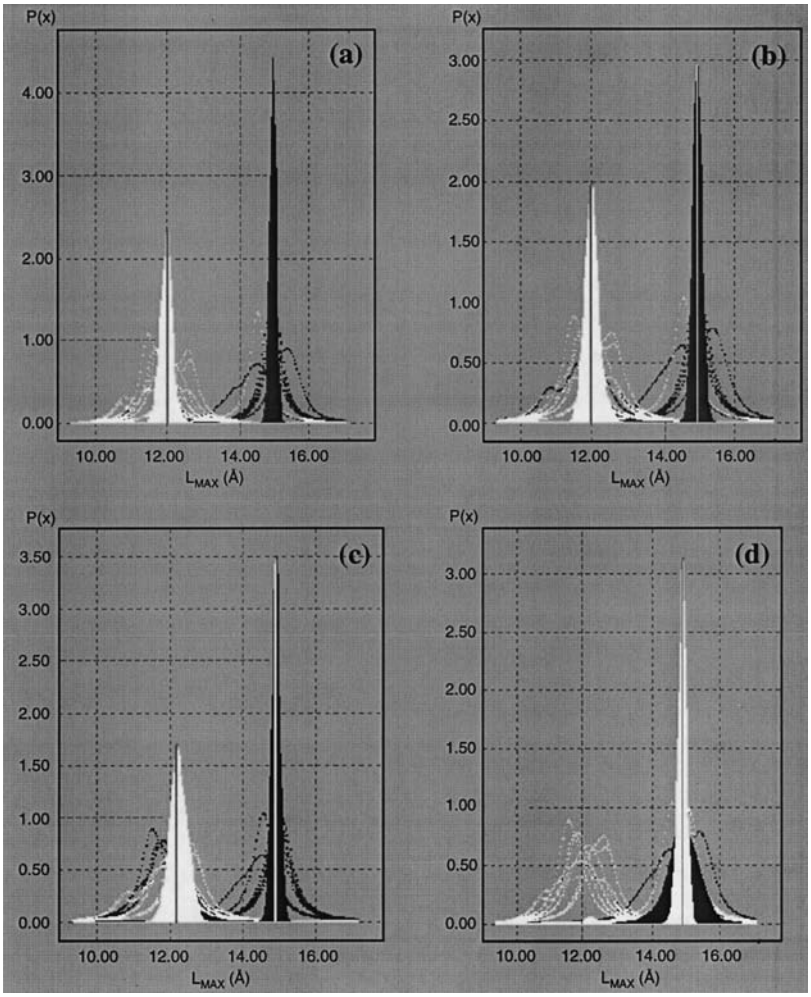


FIGURE 4 The reactivity patterns based on L_{max}, at Γ(default) = 0.929 Å, for: RAR-α (a), RAR-β (b), RAR-γ (c), and RXR-α (d). The integral reactivity pattern of highly active ligands is dark, whereas the pattern of low activity ligands is white.

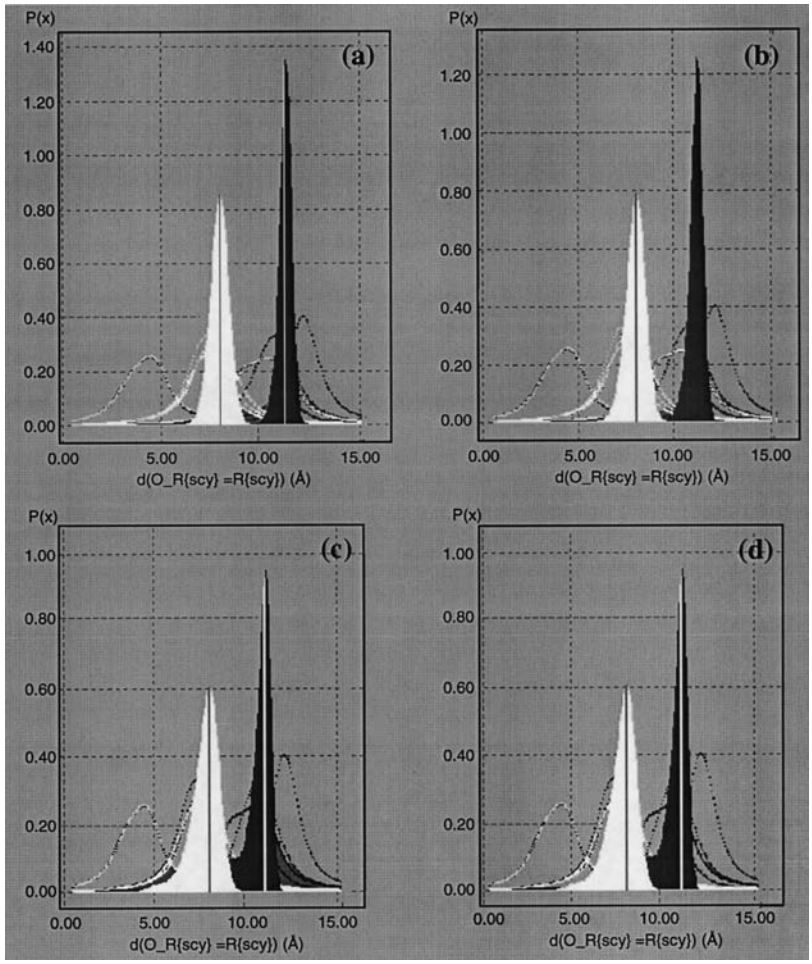


FIGURE 5 The reactivity patterns based on local steric distances between carboxylic oxygen and unsaturated (aromatic or double) bond in cyclic moiety, at Γ(default) = 1.448 Å, for: RAR-α (a), RAR-β (b), RAR-γ (c), and RXR-α (d). The integral reactivity pattern of highly active ligands is dark, whereas the pattern of low activity ligands is white.

classified as most active. Reactivity patterns are shown for the electronic parameters E_{LUMO} and E_{HOMO} (Figs. 2 and 3(a)–(d)), and the steric descriptor L_{max} (Fig. 4(a)–(d)). The relevance of L_{max} to receptor activation would seem to be reflected in the identification of the local steric descriptors as predictive of chemicals with high versus low activity, specifically, *d*(RX_O);RX:R{scy} = R{scy},R{scy}*R{scy} (Fig. 5(a)–(d)) and *d*(O_C{sp^3}) (Fig. 6(a)–(d)). In the case of the latter variable, it was of interest to note that as resolution of Γ increased (from 1.699 to 1.000 A), two ranges/peaks were noted for the chemicals classified as most reactive with the RARs (Fig. 7(a)–(d); Table II).

There were systematic differences between reactivity patterns derived for the three RAR variants and that for RXR-α. Reactivity patterns for RAR-α, RAR-β, and RAR-γ were quite similar with respect to ranges of both the electronic and steric descriptors (Table III). Active ligands had common ranges of −0.79 to −0.76 eV for E_{LUMO}, −8.94 to −8.81 eV for E_{HOMO}, 14.9–15.0 Å for L_{max}, 11.2–11.3 Å for *d*(RX_O);RX:R{scy} = R{scy},R{scy}*R{scy}, and a doublet pattern of 12.0–12.1 Å and 12.9–13.4 Å for *d*(O_C{sp^3}) at Γ = 1.00 Å. The

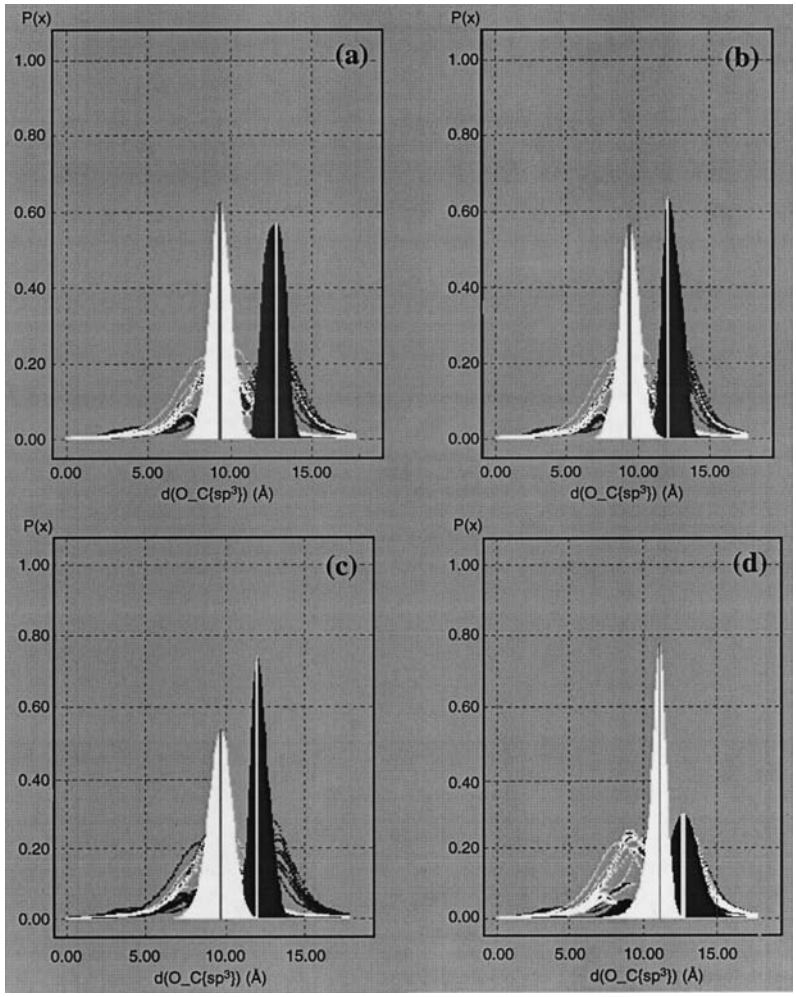


FIGURE 6 The reactivity patterns based on local steric distances between carboxylic oxygen and C{sp^3}, at Γ(default) = 1.669 Å, for: RAR-α (a), RAR-β (b), RAR-γ (c), and RXR-α (d). The integral reactivity pattern of highly active ligands is dark, whereas the pattern of low activity ligands is white.

profile of active chemicals for the RXR differed from this, with shifts toward lower electrophilicity (−0.87 to −0.89 eV for E_{LUMO}), higher nucleophilicity (−8.76 to −8.72 eV for E_{HOMO}), and larger distances for d(O_(R{scy} = R{scy}) (11.6–11.8 Å) and d(O_C{sp^3}) (12.6–13.0 Å), the latter of which further differed from the RAR pattern in that it was comprised of a single peak. Caution must be exercised in the interpretation of RAR versus RXR reactivity patterns due to the small number of active chemicals in this data set for the latter receptor. However, relatively fewer ligands have been identified for RXRs than RARs, perhaps reflecting a greater specificity of receptors in the RXR family for discrete structural characteristics in ligands.

Despite a high degree of structural similarity, differences in ligand affinity/specificity between RARs and RXRs, as well as variations in their biological function are well established [1,12]. Use of the COREPA algorithm to generate different reactivity patterns for the two receptor families for a common set of relatively flexible test chemicals could provide a basis for better understanding and, perhaps, exploiting differences in activity of retinoid

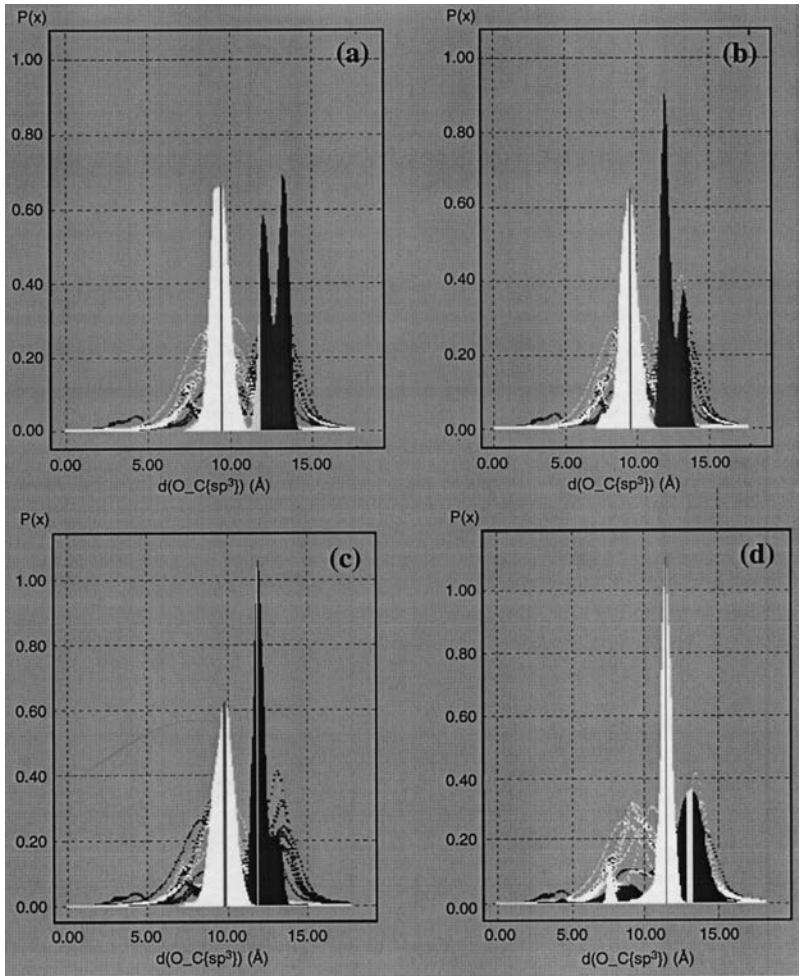


FIGURE 7 The reactivity patterns based on local steric distances between carboxylic oxygen and C{sp^3}, at Γ(default) = 1.000 Å, for: RAR-α (a), RAR-β (b), RAR-γ (c), and RXR-α (d). The integral reactivity pattern of highly active ligands is dark, whereas the pattern of low activity ligands is white.

ligands that control specific biological processes. This type of analysis could be enhanced through utilization of larger, structurally more diverse data sets, such as that described by Douguet *et al.* [11] for natural and synthetic retinoid receptor agonists.

Acknowledgements

This paper has been reviewed in accordance with US Environmental Protection Agency (EPA) policy. Mention of tradenames, models, or modeling approaches does not indicate endorsement by the US government. Portions of this work were supported by a cooperative agreement (CR822306-01-0) between EPA and Bourgas University ("Prof. As. Zlatarov"). Helpful review comments on an earlier version of the manuscript were provided by Sig Degitz and Chris Russom. Roger LePage and Diane Spehar assisted with manuscript preparation.

References

[1] Sucov, M. and Evans, R.M. (1995) "Retinoic acid and retinoic acid receptors in development", *Mol. Neurobiol.* **10**, 169–184.

[2] Boyd, A.S. (1989) "An overview of retinoids", *Am. J. Med.* **86**, 568–574.

[3] Hong, W.K. and Itri, L.M. (1994) "Retinoids and human cancer", In: Sporn, M.B., Roberts, A.B. and Goodman, D.S., eds, *The Retinoids: Biology, Chemistry and Medicine* (Raven, New York), pp. 597–630.

[4] Gardiner, D.M. and Hoppe, D.M. (1999) "Environmentally induced limb malformations in mink frogs (*Rana septentrionalis*)", *J. Exp. Zool.* **284**, 207–216.

[5] Degitz, S.J., Kosian, P.A., Makynen, E.A., Jensen, K.M. and Ankley, G.T. (2000) "Stage- and species-specific developmental toxicity of all-*trans* retinoic acid in four native North American ranids and *Xenopus laevis*", *Toxicol. Sci.* **57**, 264–274.

[6] Mekenyan, O.G., Ivanov, J.M., Karabunarliev, S.H., Bradbury, S.P., Ankley, G.T. and Karcher, W. (1997) "A computationally-based hazard identification algorithm that incorporates ligand flexibility. 1. Identification of potential androgen receptor ligands", *Environ. Sci. Technol.* **31**, 3702–3711.

[7] Mekenyan, O.G., Nikolova, N., Karabunarliev, S.H., Bradbury, S.P., Ankley, G.T. and Hansen, B. (1999) "New developments in a hazard identification algorithm for hormone receptor ligands", *Quant. Struct.-Act. Relat.* **18**, 139–153.

[8] Bonchev, D., Mountain, C.F., Seitz, W.A. and Balaban, A.T. (1993) "Modelling the anticancer action of some retinoid compounds by making use of OASIS method", *J. Med. Chem.* **36**, 1562–1569.

[9] Jaeger, E.P., Jurs, P.C. and Stouch, T.R. (1993) "Structure–activity relationship studies of retinoid cancer inhibition", *Eur. J. Med. Chem.* **28**, 275–290.

[10] Klaholz, B.P., Renaud, J.P., Mitschler, A., Zusi, C., Chambon, P., Gronemeyer, H. and Moras, D. (1998) "Conformational adaptation of agonists to the human nuclear receptor RARα", *Nat. Struct. Biol.* **5**, 199–202.

[11] Douguet, D., Thoreau, E. and Grassy, G. (1999) "Quantitative structure–activity relationship studies of RARα, β, γ retinoid agonists", *Quant. Struct.-Act. Relat.* **18**, 107–122.

[12] Jong, L., Lehmann, J.M., Hobbs, P.D., Harlev, E., Huffman, J.C., Pfahl, M. and Dawson, M.I. (1993) "Conformational effects on retinoid receptor selectivity. 1. Effect of 9-double bond geometry on retinoid X receptor activity", *J. Med. Chem.* **36**, 2605–2613.

[13] Bradbury, S., Kamenska, V., Schmieder, P., Ankley, G. and Mekenyan, O. (2000) "A computationally-based identification algorithm for estrogen receptor ligands. Part I. Predicting hERα binding affinity", *Toxicol. Sci.* **58**, 253–269.

[14] Mekenyan, O.G., Kamenska, V., Schmieder, P.K., Ankley, G.T. and Bradbury, S.P. (2000) "A computationally-based identification algorithm for estrogen receptor ligands. Part II. Evaluation of a hERα binding affinity model", *Toxicol. Sci.* **58**, 270–281.

[15] Schmieder, P.K., Aptula, A.O., Routledge, E.J., Sumpter, J.P. and Mekenyan, O.G. (2000) "Estrogenicity of alkylphenolic compounds: a 3-D structure activity evaluation of gene activation", *Environ. Toxicol. Chem.* **19**, 1727–1740.

[16] Ivanov, J.M., Karabunarliev, S.H. and Mekenyan, O.G. (1994) "A system for an exhaustive 3D molecular design", *J. Chem. Inf. Comput. Sci.* **34**, 234–243.

[17] Stewart, J.J.P. (1990) "MOPAC: a semiempirical molecular orbital program", *J. Comput.-Aided Mol. Des.* **4**, 1–105.

[18] Stewart, J.J.P. (1993). MOPAC 93. Fujitsu Limited, Chiba-City, Chiba 261, Japan and Stewart Computational Chemistry, Colorado Springs, CO.

[19] Bradbury, S.P., Mekenyan, O.G. and Ankley, G.T. (1998) "The role of ligand flexibility in predicting biological activity: structure–activity relationships for aryl hydrocarbon, estrogen and androgen receptor binding affinity", *Environ. Toxicol. Chem.* **17**, 15–25.

[20] Ivanov, J.M., Mekenyan, O.G., Bradbury, S.P. and Schuurmann, G. (1998) "A kinetic analysis of the conformational flexibility of steroids", *Quant. Struct.-Act. Relat.* **17**, 437–449.

[21] Anstead, G.M., Wilson, S.R. and Katzenellebogen, J.A. (1989) "2-Arylindenes and 2-arilindenones: molecular structures and considerations in the binding orientation of unsymmetrical non-steroidal ligands to the estrogen receptor", *J. Med. Chem.* **32**, 2163–2171.

[22] Anstead, G.M., Carlson, K.E. and Katzenellenbogen, J.A. (1997) "The estradiol pharmacophore: ligand structure-estrogen receptor binding affinity relationships and a model for the receptor binding site", *Steroids* **62**, 268–303.

[23] Wiese, T. and Brooks, S.C. (1994) "Molecular modeling of steroidal estrogens: novel conformations and their role in biological activity", *J. Steroid Biochem. Mol. Biol.* **50**, 61–72.