Prediction of Ligand Binding Affinity and Orientation of Xenoestrogens to the Estrogen Receptor by Molecular Dynamics Simulations and the Linear Interaction Energy Method

Marola M. H. van Lipzig,[†] Antonius M. ter Laak,[†] Aldo Jongejan,[‡] Nico P. E. Vermeulen,[†] Mirjam Wamelink,[†] Daan Geerke,[†] and John. H. N. Meerman*,[†]

Leiden/Amsterdam Center for Drug Research, Division of Molecular Toxicology and Division of Molecular Pharmacology, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

Received July 17, 2003

Exposure to environmental estrogens has been proposed as a risk factor for disruption of reproductive development and tumorigenesis of humans and wildlife (McLachlan, J. A.; Korach, K. S.; Newbold, R. R.; Degen, G. H. Diethylstilbestrol and other estrogens in the environment. *Fundam. Appl. Toxicol.* **1984**, *4*, 686−691). In recent years, many structurally diverse environmental compounds have been identified as estrogens. A reliable computational method for determining estrogen receptor (ER) binding affinity is of great value for the prediction of estrogenic activity of such compounds and their metabolites. In the presented study, a computational model was developed for prediction of binding affinities of ligands to the ERα isoform, using MD simulations in combination with the linear interaction energy (LIE) approach. The linear interaction energy approximation was first described by Åqvist et al. (Åqvist, J.; Medina, C.; Samuelsson, J. E. A new method for predicting binding affinity in computer-aided drug design. *Protein Eng.* **1994**, *7*, 385−391) and relies on the assumption that the binding free energy (ΔG) depends linearly on changes in the van der Waals and electrostatic energy of the system. In the present study, MD simulations of ligands in the ERα ligand binding domain (LBD) (Shiau, A. K.; Barstad, D.; Loria, P. M.; Cheng, L.; Kushner, P. J.; Agard, D. A.; Greene, G. L. The structural basis of estrogen receptor/coactivator recognition and the antagonism of this interaction by tamoxifen. *Cell* **1998**, *95*, 927−937), as well as ligands free in water, were carried out using the Amber 6.0 force field (http://amber.scripps.edu/). Contrary to previous LIE methods, we took into account every possible orientation of the ligands in the LBD and weighted the contribution of each orientation to the total binding affinity according to a Boltzman distribution. The training set ($n = 19$) contained estradiol (E2), the synthetic estrogens diethylstilbestrol (DES) and 11β-chloroethylestradiol (E2-Cl), 16α-hydroxy-E2 (estriol, EST), the phytoestrogens genistein (GEN), 8-prenylnaringenin (8PN), and zearalenon (ZEA), four derivatives of benz[*a*]antracene-3,9-diol, and eight estrogenic monohydroxylated PAH metabolites. We obtained an excellent linear correlation ($r^2 = 0.94$) between experimental (competitive ER binding assay) and calculated binding energies, with K_d values ranging from 0.15 mM to 30 pM, a 5 000 000-fold difference in binding affinity. Subsequently, a test set ($n = 12$) was used to examine the predictive value of our model. This set consisted of the synthetic estrogen 5,11-*cis*-diethyl-5,6,11,12-tetrahydrochrysene-2,8-diol (THC), daidzein (DAI), equol (EQU) and apigenin (API), chlordecone (KEP), progesterone (PRG), several mono- and dihydroxylated PAH metabolites, and two brominated biphenyls. The predicted binding affinities of these estrogenic compounds were in very good agreement with the experimental values (average deviation of 0.61 ± 0.4 kcal/mol). In conclusion, our LIE model provides a very good method for prediction of absolute ligand binding affinities, as well as binding orientation of ligands.

Introduction

Estrogens play a critical role in the growth, development, and maintenance of a diverse range of tissues. The effects of estrogens are mediated by the estrogen receptor (ER), a ligand-activated transcription factor.[5] Binding of agonists induces a conformational change of the ER, enabling the receptor to homodimerize. The dimer is then translocated to the nucleus, where it enhances gene transcription.[6] Two ER isoforms have been identified and crystallized, ERα and ERβ, that show great resemblance in overall structure. Both bind the endogenous estrogen, 17β-estradiol (E2), but show different tissue expression and distinct responses to various agonist and antagonists.[7,8] Binding of agonists induces a specific orientation of helix 12 (H12), thereby closing the narrow binding site and allowing coregulators to bind.[9] Antagonist binding, however, induces a different orientation of H12, preventing alignment over the binding site and transactivation.[6,10]

Exposure to environmental estrogens has been proposed as a risk factor for disruption of reproductive

* To whom correspondence should be addressed. Address: Leiden/Amsterdam Center for Drug Research (LACDR), Division of Molecular Toxicology, Faculty of Sciences, Vrije Universiteit, Amsterdam, De Boelelaan 1083, 1081 HV Amsterdam, The Netherlands. Phone: +31 20 4447606. Fax: +31 20 4447610. E-mail: meerman@chem.vu.nl.
[†] Division of Molecular Toxicology.
[‡] Division of Molecular Pharmacology.

10.1021/jm0309607 CCC: $27.50
Published on Web 01/13/2004

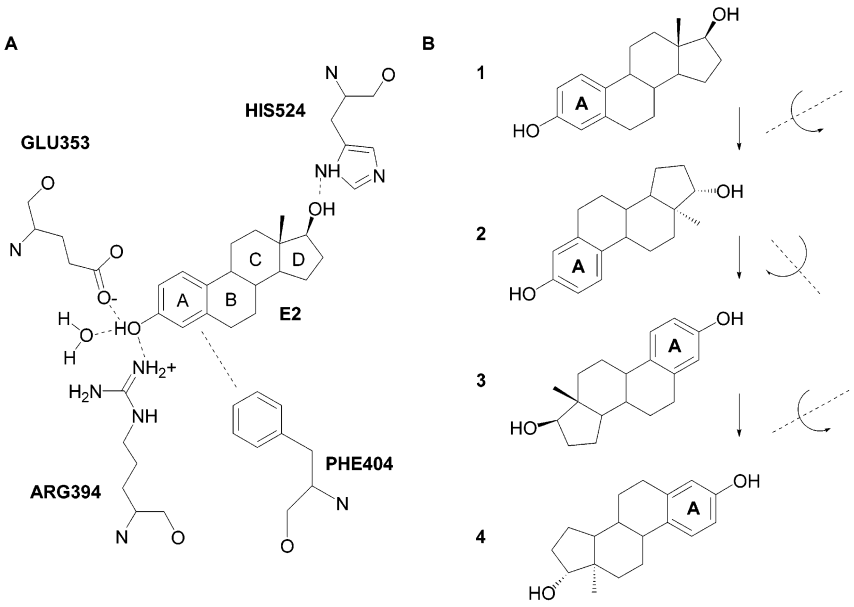

Figure 1. (A) Schematic view of hydrogen bonding and aromatic interactions (dashed lines) between estradiol (E2) and several residues (Glu353, Arg394, His524, and Phe404) of the ERα and a water molecule. E2 is depicted in orientation 1. (B) Four ligand docking orientations used in the present study, with E2 as an example.

development and tumorigenesis of humans and wildlife.[1] In recent years, a wide variety of structurally diverse environmental compounds have been identified as estrogens. Among these compounds are natural steroid hormones, pharmaceuticals, industrial chemicals including polychlorinated biphenyls (PCBs), pesticides, and polycyclic aromatic hydrocarbons (PAHs). PAHs have been considered a health risk for wild life and humans for a long time.[11,12] This is mainly due to their carcinogenic properties, which result from covalent binding to DNA and other macromolecules after bioactivation to reactive metabolites.[13−15] However, it has been recently shown that bioactivation of the PAHs benzo[*a*]pyrene (BaP) and chrysene (CHN) may also lead to the formation of hydroxylated metabolites with estrogenic activity.[16,17] Similarly, formation of hydroxylated metabolites with estrogenic activity has also been observed for chlorinated biphenyls[17,18] and metabolites of pesticides such as DDT and methoxychlor.[19,20]

In light of the large variety of structurally diverse compounds that bind to the ERα, it is of great value to develop a reliable method to predict the ERα binding affinity of compounds. For the ER, several approaches to calculate ligand binding affinities have been used over the past decade, ranging from ligand-based comparative molecular field analysis (CoMFA) studies[21] to protein-based methods using empirical scoring functions[22] and molecular dynamics simulations (MD), using free energy perturbation methods (FEP).[23] However, accurate prediction of binding affinity of novel estrogens remains difficult, not in the least since most of the approaches are limited in their applicability to compounds that have a common template structure. Furthermore, conclusive prediction of favored binding orientation is generally ignored in these studies.

A recent method for estimation of binding affinities is the linear interaction energy (LIE) approximation, which was first described and further extended by Åqvist et al.[2,24−26] The LIE model relies on the assumption that the ΔG depends linearly on changes in the van der Waals and electrostatic energy of the system. This is supported by observations that the free energy of solvation on nonpolar moieties often scale linearly with respect to variables characterizing the size of the solute. The concept of the LIE approach is to separately evaluate the electrostatic and van der Waals interaction energies of the ligand in the bound and free states. For this purpose, two MD simulations are carried out: one with the ligand bound to the (solvated) receptor and one with the unbound ligand in solvent. Subsequently, averages of interaction energies between the ligand and its surroundings are analyzed. The calculated binding free energy (ΔG_{calc}) is obtained from the averages as[24]

$$\Delta G_{calc} = \alpha\,\Delta E_{INT}^{VDW} + \beta\,\Delta E_{INT}^{EL} = \alpha\,(E_{INT_{BOUND}}^{VDW} - E_{INT_{FREE}}^{VDW}) + \beta\,(E_{INT_{BOUND}}^{EL} - E_{INT_{FREE}}^{EL}) \quad (1)$$

where E_{INT}^{VDW} denotes the van der Waals interaction energy and E_{INT}^{EL} denotes the electrostatic interaction energy of the ligand bound to the receptor (BOUND) and the ligand free in solution (FREE).

Obtaining suitable values for α and β has been subject of several investigations described recently. The α value strongly depends on the system of interest, the force field employed, and the computational methods that are applied. A proper α value must therefore first be determined by comparing calculated and experimentally estimated binding affinities. Originally, the LIE method was introduced with a fixed value of 0.5 for β[2]. However, an extensive study on the solvation energy of various series of small substrates with different number of electronegative groups showed that β significantly differs between substrates that vary in the number of strong electronegative groups. For example, β was shown to decrease with the number of hydroxyl groups.[24] This might be due to an overestimation of the force field of electrostatic interactions between these groups and their surroundings possibly because polarization effects are not properly described by force fields. Therefore, like

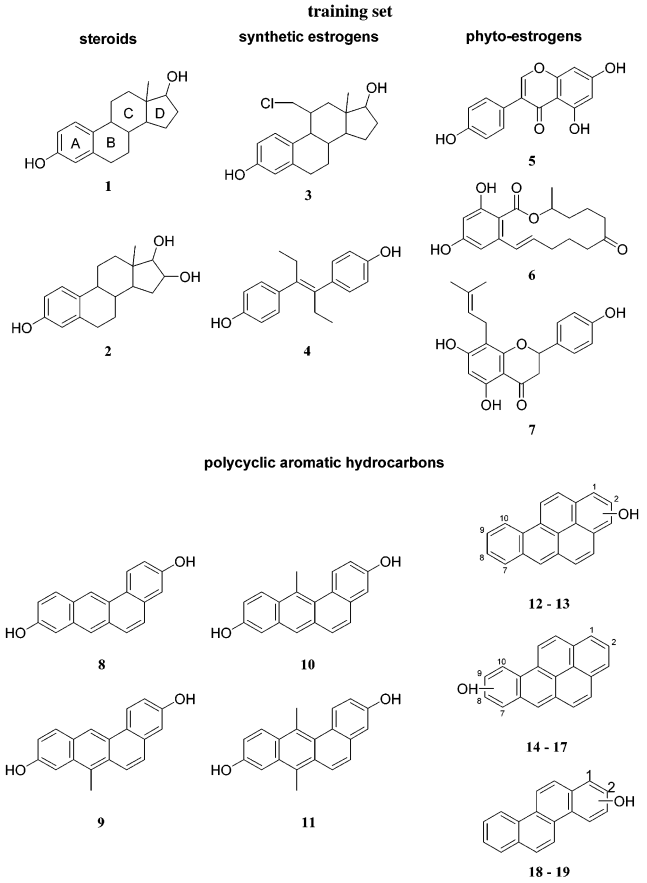

Figure 2. Molecular structures and names of estrogenic compounds of the training set ($n = 19$) used in the present study: **1**, estradiol (E2); **2**, estriol (EST); **3**, 11β-chloroestradiol (E2-Cl); **4**, DES; **5**, genisteine (GEN); **6**, zearalenone (ZEA); **7**, 8-prenylnaringenin (8PN); **8**, benz[*a*]anthracene-3,9-diol (BA-diol); **9**, 7-methylbenz[*a*]anthracene-3,9-diol (7-MBA-diol); **10**, 12-methylbenz[*a*]anthracene-3,9-diol (12-MBA-diol); **11**, 7,12-dimethylbenz[*a*]anthracene-3,9-diol (7,12-DMBA-diol); **12** and **13**, 1- and 2-hydroxybenzo[*a*]pyrene (1-OHBaP, 2-OHBaP); **14**−**17**, 7-, 8-, 9-, and 10-hydroxybenzo[*a*]pyrene (7-OHBaP, 8-OHBaP, 9-OHBaP, 10-OHBaP); **18**−**20**, 1- and 2-hydroxychrysene (1-OHCHN, 2-OHCHN). For benzo[*a*]pyrene and chrysene derivatives, the numbering indicates the positions of the hydroxyl groups.

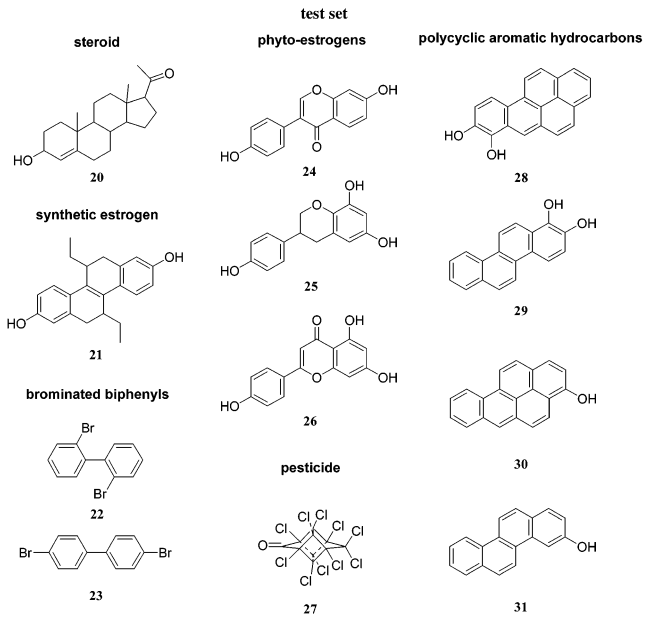

Figure 3. Molecular structures and names of estrogenic compounds of the test set ($n = 12$) used in the present study: **20**, progesterone (PRG); **21**, 5,11-*cis*-diethyl-5,6,11,12-tetrahydrochrysene-2,8-diol (THC); **22**, 2,2′-dibromobiphenyl (2DBB); **23**, 4,4′-dibromobiphenyl (4DBB); **24**, daidzein (DAI); **25**, equol (EQU); **26**, apigenin (API); **27**, kepone (KEP); **28**, 7,8-dihydroxybenzo[*a*]pyrene (7,8-diOHBaP); **29**, 1,2-dihydroxychrysene (1,2-diOHCHN); **30**, 3-hydroxybenzo[*a*]pyrene (3-OHBaP); **31**, 3-hydroxychrysene (3-OHCHN). All structures are shown in orientation 1, as illustrated for E2 in Figure 1B.

α, the value of β also depends on the system under investigation and the force field used. In conclusion, to set up an LIE model for the prediction of ligand binding affinities using a particular force field, it is necessary to perform a double regression (α and β) analysis of a set of compounds and their experimental ΔG values.

In the present study, MD simulations (100 ps) of receptor-bound ligands as well as ligands free in water were carried out. In addition, to investigate the favored binding mode, ligands were manually docked in the crystal structure of the ERα LBD in all possible orientations, as illustrated for E2 in Figure 1. To calibrate our model, we used the endogenous estrogen E2, the synthetic estrogens diethylstilbestrol (DES) and 11β-chloroethylestradiol (E2-Cl), the steroid 16α-hydroxy-E2 (estriol, EST), the phytoestrogens genistein (GEN), 8-prenylnaringenin (8PN), and zearalenon (ZEA), four derivatives of benz[*a*]antracene-3,9-diol, and eight estrogenic monohydroxylated PAH metabolites (Figure 2. The binding affinities for the ERα in this training set ($n = 19$) range from 0.15 mM to 30 pM. For the prediction of binding affinities of compounds for the ERα, the synthetic estrogen 5,11-*cis*-diethyl-5,6,11,12-tetrahydrochrysene-2,8-diol (THC), the phytoestrogens daidzein (DAI), equol (EQU), and apigenin (API), the pesticide chlordecone (kepone, KEP), the endogenous hormone progesterone (PRG), the PAH metabolites 3-hydroxybenzo[*a*]pyrene (3-OHBaP), 3-hydroxychrysene (3-OHCHN), 7,8-dihydroxy-BaP (7,8-diOHBaP), and 1,2-dihydroxy-CHN (1,2-diOHCHN), and two brominated biphenyls (2,2′-dibromobiphenyl, 2DBB, and 4,4′-dibromobiphenyl, 4DBB) were used as a test set ($n = 12$, Figure 3).

Results

Structural Analysis of MD Simulations. The use of relatively short MD simulations of 100 ps allowed us to investigate many different binding orientations at relatively low computational cost. Because 100 ps is rather short for extensive studies on protein dynamics, equilibration and minimization were carried out before MD simulation in order to obtain stable receptor−ligand complexes for each ligand in each binding orientation tested, using positional restraints on the C_α's of the backbone. Inspection of the trajectories revealed that secondary and tertiary structures of the receptor−ligand complex remained stable during the 100 ps of the MD simulation with this strategy. The averaged root-mean-square deviation (rmsd) values of the backbone C_α's compared to the initial conformation were found to be stable after the first 30 ps of the MD run for almost all simulations. The rmsd values of the backbone C_α's for all MD simulations compared to that of the DES-bound crystal structure were rather small, i.e., 1.26 ± 0.12 Å.

For some of the biphenyl structures, the protein structure was still slightly diverging after 30 ps of simulation (Δ rmsd > 0.0025 Å/ps). Extension by 50 ps led to satisfactory stable (Δ rmsd < 0.001 Å/ps) and small rmsd values (C_α's, 1.14 ± 0.10 Å), which were subsequently analyzed over the last 85 ps. These results indicate that during 100−150 ps of simulation, the ligand−protein complex remained stable and that substitution of DES by other ligands does not affect the conformation of the protein.

Analysis of the rmsd values of the heavy atoms of several selected amino acid residues of the binding cavity (343, 346, 349, 350, 353, 384, 387, 388, 391, 394, 404, 424, 521, 524, 525) showed even smaller variations: average rmsd = 0.98 ± 0.1 Å. This indicates that the orientation and conformation of the amino acids directly involved in ligand binding was also well-preserved.

Free Energy of Binding and Hydrogen Bonding. Variation in the van der Waals interaction energy (E_{INT}^{VDW}) was smaller during the MD runs with an average standard deviation of ca. 10% over 35 frames compared with the electrostatic interaction energy (E_{INT}^{EL}), showing a standard deviation of more than 20%.

Hydrogen bonding during MD simulations between the ligand and the receptor or the surrounding water molecules was analyzed with Carnal/Anal (Amber module for Cartesian coordinates analysis). This revealed that next to the "regular" hydrogen bonds with the residues Glu353, Arg394, and His524 and a water molecule as seen for DES and E2 in the crystal structures, many ligands also interacted by hydrogen bonding with Leu525, Leu391, Gly521, and Met343. Sporadically, a hydrogen bond interaction was found with Ala350, Met388, Leu346, Leu387, and Met421 (data not shown). The occurrence of hydrogen bonds during the last 70 ps of the MD simulation is shown in Table 1. Hydrogen bonding with residues other than Glu353, Arg394, and His524 is summarized by "other". For the residues Glu353, Arg394, and His524 and the water molecule, a percentage of more than 100% indicates that more than one hydrogen bond between the ligand and the amino acid is formed during the simulation. This implies a more stable interaction. In most cases, hydrogen bonding with Glu353, Arg394, and His524 was observed for the orientation with the best Amber interaction energy.

Binding Orientations. For each ligand, generally four different orientations were evaluated for binding to the ERα. A good steric complementarity with the binding site was not achieved for all orientations. In some cases certain orientations showed very large overlap and steric hindrance with the protein atoms, causing these complexes to be unusable for MD simulations. This occurred with certain orientations of DES, ZEA, and 8PN. For all successfully docked orientations, MD simulations were performed and the results of their binding properties are presented in Table 2.

In the case of DES and E2, we found that the most favorable orientation in the ERα, i.e., the orientation with the best interaction energy, matched the orientation found in the crystal structures for both ligands. Interestingly, our studies showed that the energetically favorable orientation of GEN in the ERα (similar to orientation 3 of E2 in Figure 1) is different from its orientation in the crystal structure of ERβ.[10] This may be related to the fact that GEN is an agonist for ERα but a partial agonist for ERβ. Moreover, GEN induces an antagonized receptor conformation by a different alignment of helix 12 than seen for agonists.[10]

Linear Interaction Energy Calculations. The experimental binding energies (ΔG_{exp}) were calculated from the experimentally determined ligand binding affinities (K_d) by eq 2, using the gas constant (R) and the temperature (T):

$$\Delta G_{exp} = -RT \ln K_d \tag{2}$$

The K_d values were determined in a radioligand receptor binding assay with sheep uterus cytosol as a source for the ER.

To obtain values for ΔG_{calc}, van der Waals (E_{INT}^{VDW}) and electrostatic (E_{INT}^{EL}) interaction energies between the ligand and its environment were calculated for all trajectory files from the MD simulation. The interaction energy (E_{INT}) between a ligand and its surroundings, i.e., receptor−water (eq 3a) or water only (eq 3b), is equal to the difference between the energy of the total system (E_T) and the sum of the energy of the individual components of the system: water (E_W), receptor (E_R), and ligand (E_L) for the ligand−receptor complex; water (E_W) and ligand for the free ligand (E_L). Thus, the interaction energies (E_{INT}) are described as follows:

$$E_{INT_{BOUND}}^{VDW} + E_{INT_{BOUND}}^{EL} = (E^{VDW} + E^{EL})_{T_{BOUND}} - (E^{VDW} + E^{EL})_W - (E^{VDW} + E^{EL})_R - (E^{VDW} + E^{EL})_L \tag{3a}$$

$$E_{INT_{FREE}}^{VDW} + E_{INT_{FREE}}^{EL} = (E^{VDW} + E^{EL})_{T_{FREE}} - (E^{VDW} + E^{EL})_W - (E^{VDW} + E^{EL})_L \tag{3b}$$

Energy terms were calculated in the Sander module of Amber (acronym for simulated annealing with NMR-derived energy restraints), by a single-step energy calculation of the total system and all the separate components, under the same conditions as the MD simulations. Energies were averaged over frames taken every 2 ps from the last 70 ps of the trajectory. Subsequently, Amber interaction energies (ΔE_{AMBER}) were calculated for any orientation i by

$$\Delta E_{AMBER_i} = (E_{INT_{bound,i}}^{VDW} - E_{INT_{free}}^{EL}) + (E_{INT_{bound,i}}^{VDW} - E_{INT_{free}}^{EL}) = E_{INT_i}^{VDW} + E_{INT_i}^{EL} \tag{4}$$

To include all orientations of each compound in our model, ΔE_{AMBER} values for the different orientations were used for calculation of a Boltzman weight factor (p_i) for orientation i by

$$p_i = \frac{e^{-\Delta E_{AMBERi}/k_B T}}{\sum_i e^{-\Delta E_{AMBERi}/k_B T}} \tag{5}$$

where k_B is the Boltzman constant (kcal/(mol·K)) and T is the temperature (K). The experimental ΔG_{exp}

Table 1. Presence of Hydrogen Bonds between Ligands and the Residues Glu353, Arg394, and His524 of the ERα LBD or (any) Water Molecule during 70 ps of MD Simulation[a]

ligand	orientation	GLU353	ARG394	HIS524	other	water	total
DES	1	102.9	28.6	40.0	17.1	62.9	251.4
	2		28.6	5.7	105.7	100.0	240.0
	4	102.9	17.1	91.4	105.7	2.9	320.0
E2	1	114.3	14.3	97.1	11.4	88.6	325.7
	2	100.0	14.3	94.3	88.6	82.9	380.0
	3	100.0	28.6	100.0	100.0	105.7	434.3
	4	100.0		97.1	100.0	14.3	311.4
E2-Cl	1	111.4	25.7	97.1	20.0	91.4	345.7
	2	105.7	31.4	100.0	105.7	51.4	394.3
EST	1	117.1	8.6	40.0	45.7	91.4	302.9
	3	97.1	80.0		88.6	2.9	268.6
GEN	1	94.3	60.0		108.6	65.7	328.6
	2	102.9	31.4	91.4	100.0	77.1	402.9
	3	120.0	31.4	91.4	154.3	57.1	454.3
	4	100.0	8.6	94.3		91.4	294.3
8PN	1			2.9	154.3		157.1
	4	111.4		68.6	17.1	91.4	288.6
ZEA	2	100.0					100.0
BA-diol	1	100.0	23.1	73.1	42.3	92.3	330.8
	2	14.3	14.3	100.0	108.6	97.1	334.3
	3	105.7	45.7	51.4	8.6	91.4	302.9
	4	100.0	31.4	97.1	100.0		328.6
7-MBA-diol	1		2.9	68.6	5.7	108.6	185.7
	2			45.7	100.0	54.3	200.0
	3		14.3	57.1	102.9	51.4	225.7
	4	111.4	31.4	88.6	28.6	85.7	345.7
12-MBA-diol	1	100.0	46.2		76.9	69.2	292.3
	2	100.0	11.4	94.3	22.9		228.6
	3	103.9	19.2	50.0		88.5	261.5
	4	71.4	8.6	62.9	31.4	65.7	240.0
	1a	91.4	37.1	94.3	94.3	88.6	405.7
	2a	100.0	25.7	100.0	105.7		331.4
	3a	107.7	30.8	96.2	100.0	84.6	419.2
	4a	102.9		62.9	2.9		168.6
7,12-DMBA-diol	1	108.6	22.9	94.3	17.2	85.7	328.6
	2	100.0	105.7		11.4	68.6	285.7
	3	100.0	25.7	34.3	42.9	94.3	297.2
	4	100.0	22.9	94.3		85.7	302.9
	1a	100.0	17.1	68.6		74.3	260.0
	2a	105.7	37.1	100.0	91.4	97.1	431.4
	3a		11.4	37.1	97.1	34.3	180.0
	4a	8.6		51.4		97.1	157.1
1-OHBaP	1				5.7	100.0	105.7
	2				114.3		114.3
	3			91.4			91.4
	4					97.1	97.1
2-OHBaP	1	108.6	57.1			60.0	225.7
	2	100.0	8.6				108.6
	3			100.0			100.0
	4			14.3	31.4		45.7
7-OHBaP	1			97.1			97.1
	2				11.4		11.4
	3					100.0	100.0
	4				114.3		114.3
8-OHBaP	1			82.9			82.9
	2			57.1	2.9		60.0
	3	108.6	71.4			94.3	274.3
	4	100.0	2.9				102.9
9-OHBaP	1					100.0	100.0
	2			97.1			97.1
	3				100.0		100.0
	4	45.7				125.7	171.4
10-OHBaP	1						
	2				2.9		2.9
	3				88.6		88.6
	4				100.0		100.0
1-OHCHN	1	74.3	2.9			48.6	125.7
	2	100.0					100.0
	3			97.1			97.1
	4				14.3		14.3
2-OHCHN	1	100.0	28.6			88.6	217.1
	2	11.4	11.4			102.9	125.7
	3			85.7	5.7		91.4
	4			11.4	22.9		34.3

[a] The values represent the percentages of the total time of MD simulation that a particular hydrogen bond was present. Hydrogen bonding with other residues of the ERα (i.e., Leu525, Leu391, Gly521, Met343, Ala350, Met388, Leu346, Leu387, and Met421) is summarized by "other". The table presents estrogens from the training set with two (or more) hydroxyl groups only. All simulated orientations were analyzed. For the residues Glu353, Arg394, and His524 and the water molecule, a percentage of more than 100% indicates that more than one hydrogen bond between the ligand and the amino acid is formed during the simulation. For an explanatio of the orientation numbers, see Figure 1B.

Table 2. Overview of the van der Waals (E^{VDW}) and Electrostatic (E^{EL}) Interaction Energy Contributions to the Interaction Energy (ΔE_{AMBER}), the Calculated (ΔG_{calc}) and the Experimentally (ΔG_{exp}) Derived Free Energies of Binding for Binding of 19 Ligands to the ERα, in All Orientations Simulated with MD[a]

		receptor-bound		free						
ligand	orientation	E^{VDW}	E^{EL}	E^{VDW}	E^{EL}	ΔE_{AMBER}	p_i	$p_i\Delta G_{calc,i}$	ΔG_{calc}	ΔG_{exp}
(A) Ligands of Training Set with ≥Two OH Groups										
DES	1	−40.44	−23.04	−21.53	−31.51	−10.43	0.27	−3.78	−12.54	−12.55
	2	−38.22	−19.69	−21.53	−31.51	−4.87	0.00	0.00		
	4	−37.32	−26.74	−21.53	−31.51	−11.02	0.73	−8.76		
E2	1	−39.60	−27.12	−23.12	−32.74	−10.86	0.79	−9.78	−12.24	−12.40
	2	−38.36	−27.57	−23.12	−32.74	−10.08	0.21	−2.46		
	3	−37.59	−23.39	−23.12	−32.74	−5.12	0.00	0.00		
	4	−38.76	−19.88	−23.12	−32.74	−2.78	0.00	0.00		
E2-Cl	1	−44.69	−25.64	−26.40	−32.11	−11.83	0.99	−13.66	−13.74	−14.42
	2	−41.30	−26.15	−26.40	−32.11	−8.94	0.01	−0.09		
EST	1	−42.70	−24.23	−23.21	−43.23	−0.49	1.00	−12.23	−12.23	−12.20
	3	−41.31	−18.02	−23.21	−43.23	7.12	0.00	0.00		
GEN	1	−36.12	−25.43	−20.14	−40.22	−1.19	0.00	−0.01	−9.78	−9.39
	2	−31.72	−27.07	−20.14	−40.22	1.58	0.00	0.00		
	3	−34.21	−31.37	−20.14	−40.22	−5.21	0.81	−7.89		
	4	−34.29	−30.44	−20.14	−40.22	−4.36	0.19	−1.87		
8PN	1	−48.51	−9.61	−26.87	−42.03	10.79	0.00	0.00	−10.19	−11.02
	4	−43.83	−23.28	−26.87	−42.03	1.80	1.00	−10.19		
ZEA	2	−44.29	−17.70	−26.27	−35.85	0.12	1.00	−11.19	−11.19	−11.04
BA-diol	1	−38.55	−17.08	−21.66	−34.94	−5.83	0.85	−7.44	−8.66	−9.15
	2	−37.23	−14.19	−21.66	−34.94	−2.42	0.00	−0.02		
	3	−38.77	−13.17	−21.66	−34.94	−2.04	0.00	−0.01		
	4	−38.87	−24.76	−21.66	−34.94	−4.78	0.15	−1.18		
7-MBA-diol	1	−33.84	−28.59	−23.12	−32.88	0.38	0.00	0.00	−11.33	−11.59
	2	−34.25	−24.77	−23.12	−32.88	4.59	0.00	0.00		
	3	−34.54	−24.09	−23.12	−32.88	4.07	0.00	0.00		
	4	−33.20	−28.17	−23.12	−32.88	−7.63	1.00	−11.33		
12-MBA-diol	1	−37.19	−22.58	−21.44	−37.32	−1.01	0.00	−0.01	−10.42	−9.36
	2	−39.34	−23.48	−21.44	−37.32	−4.06	0.11	−1.34		
	3	−38.19	−23.03	−21.44	−37.32	−2.46	0.01	−0.08		
	4	−36.60	−18.64	−21.44	−37.32	3.53	0.00	0.00		
	1a	−36.25	−26.26	−21.44	−37.32	−3.74	0.07	−0.66		
	2a	−36.34	−27.59	−21.44	−37.32	−5.16	0.72	−7.41		
	3a	−35.70	−27.02	−21.44	−37.32	−3.96	0.09	−0.91		
	4a	−37.18	−20.93	−21.44	−37.32	0.66	0.00	0.00		
7,12-DMBA-diol	1	−36.73	−27.00	−24.28	−31.92	−7.53	0.48	−4.46	−9.65	−8.81
	2	−39.83	−23.44	−24.28	−31.92	−7.08	0.23	−2.54		
	3	−39.97	−21.32	−24.28	−31.92	−5.10	0.01	−0.09		
	4	−37.44	−25.56	−24.28	−31.92	−6.80	0.14	−1.36		
	1a	−40.13	−21.41	−24.28	−31.92	−5.35	0.01	−0.14		
	2a	−35.65	−27.28	−24.28	−31.92	−6.73	0.13	−1.07		
	3a	−42.02	−13.03	−24.28	−31.92	1.15	0.00	0.00		
	4a	−39.05	−13.23	−24.28	−31.92	3.92	0.00	0.00		
(B) Ligands of Training Set with One OH Group										
1-OHBaP	1	−37.02	−14.34	−24.48	−21.28	−5.61	0.99	−7.33	−7.37	−6.58
	2	−37.19	−8.31	−24.48	−21.28	0.25	0.00	0.00		
	3	−36.71	−11.53	−24.48	−21.28	−2.49	0.01	−0.03		
	4	−38.39	−8.43	−24.48	−21.28	−1.07	0.00	0.00		
2-OHBaP	1	−34.87	−23.40	−24.71	−23.35	−10.20	1.00	−8.38	−8.38	−7.51
	2	−33.93	−17.21	−24.71	−23.35	−3.08	0.00	0.00		
	3	−36.42	−11.83	−24.71	−23.35	−0.19	0.00	0.00		
	4	−39.32	−6.70	−24.71	−23.35	2.05	0.00	0.00		
7-OHBaP	1	−35.83	−12.39	−24.06	−22.21	−1.95	0.10	−0.54	−6.13	−5.87
	2	−38.21	−5.65	−24.06	−22.21	2.41	0.00	0.00		
	3	−36.14	−13.39	−24.06	−22.21	−3.26	0.90	−5.57		
	4	−38.44	−7.53	−24.06	−22.21	0.31	0.00	−0.01		
8-OHBaP	1	−35.70	−10.04	−23.44	−22.93	0.63	0.00	0.00	−8.29	−7.85
	2	−37.23	−10.48	−23.44	−22.93	−1.35	0.00	0.00		
	3	−34.25	−21.50	−23.44	−22.93	−9.38	1.00	−8.28		
	4	−36.79	−14.83	−23.44	−22.93	−5.25	0.00	−0.01		
9-OHBaP	1	−36.31	−9.55	−24.54	−23.03	1.72	0.00	0.00	−6.56	−6.85
	2	−35.91	−11.70	−24.54	−23.03	−0.04	0.01	−0.02		
	3	−37.41	−8.45	−24.54	−23.03	1.72	0.00	0.00		
	4	−37.68	−13.07	−24.54	−23.03	−3.17	0.99	−6.54		
10-OHBaP	1	−37.76	−3.91	−25.51	−18.95	2.79	0.00	0.00	−6.31	−5.26
	2	−38.06	−5.42	−25.51	−18.95	0.98	0.00	−0.01		
	3	−36.71	−7.57	−25.51	−18.95	0.18	0.01	−0.03		
	4	−38.46	−8.75	−25.51	−18.95	−2.75	0.99	−6.27		
1-OHCHN	1	−32.45	−17.95	−24.48	−21.28	−4.65	0.36	−1.84	−4.41	−6.09
	2	−29.04	−21.67	−24.48	−21.28	−4.96	0.61	−2.39		
	3	−33.94	−15.06	−24.48	−21.28	−3.25	0.03	−0.18		
	4	−34.69	−5.70	−24.48	−21.28	5.37	0.00	0.00		

Table 2 (Continued)

ligand	orientation	receptor-bound E^{VDW}	receptor-bound E^{EL}	free E^{VDW}	free E^{EL}	ΔE_{AMBER}	p_i	$p_i\Delta G_{calc,i}$	ΔG_{calc}	ΔG_{exp}
(B) Ligands of Training Set with One OH Group (Continued)										
2-OHCHN	1	−33.17	−19.60	−24.71	−23.35	−4.70	1.00	−5.36	−5.36	−8.51
	2	−32.97	−12.03	−24.71	−23.35	3.06	0.00	0.00		
	3	−33.97	−10.44	−24.71	−23.35	3.65	0.00	0.00		
	4	−34.26	−6.46	−24.71	−23.35	7.35	0.00	0.00		

[a] The ΔG_{exp} values were calculated from K_d values determined using a sheep uterus cytosol radioligand binding assay. Structures and names of ligands and orientations are given in Figures 2 and 1, respectively. Section A presents ligands in the training set with two or more hydroxyl groups. Section B shows ligands with one hydroxyl group. p_i denotes the weight factor from the Boltzman distribution based on ΔE_{AMBER}. $p_i\Delta G_{calc,i}$ represents the binding free energy for orientation i calculated with the LIE method, using a common $\alpha = 0.82$ and $\beta_{OH\geq 2} = 0.20$ (number of OH groups ≥ 2) or $\beta_{OH=1} = 0.43$ (number of OH groups = 1). For each ligand, the final ΔG_{calc} equals the sum of $p_i\Delta G_{calc,i}$. All ΔE and ΔG values are given in kcal/mol.

values, the calculated ΔE_{INT}^{VDW} and ΔE_{INT}^{EL} values, and the Boltzman factors (weighting $\Delta E_{INT}^{VDW} + \Delta E_{INT}^{EL}$ for each individual orientation) were then applied in a double linear regression analysis using the fit function of the *XL*STAT tool (*XL*STAT4.4, http://www.xlstat.com) in Microsoft Excel 2000. By use of this tool, the Boltzman weight factors can be used in the double regression analysis. Optimal α and β values are calculated via

$$\Delta G_{exp} = \alpha\, \Delta E_{INT_i}^{VDW} + \beta\, \Delta E_{INT_i}^{EL} \qquad (6)$$

First, ligands with two or more OH groups were used for the calculation of α and β specific for two or more OH groups. Subsequently, a specific β was calculated for ligands with one OH group using the previously obtained α. By application of the optimal values for α and β to the averaged van der Waals and electrostatic energy terms in eq 1, ΔG_{calc_i} for each orientation was obtained. The final binding energy (ΔG_{calc}) of a ligand was obtained by summation of the binding energies of each individual orientation multiplied by the Boltzman weight factor (p_i):

$$\Delta G_{calc} = \Sigma(p_i\, \Delta G_{calc_i}) \qquad (7)$$

Training Set. For estrogens with two or more OH groups, the best fitting values were $\alpha = 0.82$ and $\beta_{OH\geq 2} = 0.20$. For estrogens with one OH group, single regression analysis using $\alpha = 0.82$ yielded a value for $\beta_{OH=1}$ of 0.43. When no fixed value for α was used, double regression analysis yielded optimal values of $\alpha = 0.84$ and $\beta_{OH=1} = 0.46$. This indicates the LIE method results in similar parameters, independent of the chosen data set. Nevertheless, one value for α (0.82) was used for all compounds.

Upon correlation of ΔG_{calc} with ΔG_{exp}, a linear correlation coefficient (r^2) of 0.94 was obtained (Figure 4) when all orientations of each ligand were taken into account and Boltzman weighting based on the Amber interaction energy was applied. If the regression analysis was performed, also taking all orientations in account but weighted evenly, the correlation between calculated and experimental ΔG values drops ($r^2 = 0.87$). This is to be expected because there is a large bias due to the overestimated influence of less favorable orientations on the regression analysis. When ΔG_{calc} values were calculated on the basis of only the orientation with the best interaction energy of each ligand in the ER, an excellent correlation was also found between the calculated ΔG_{calc} value and experimental binding energy ($r^2 = 0.94$).

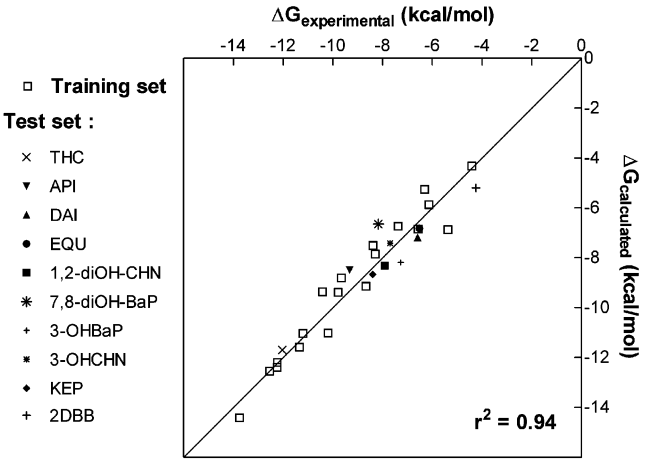

Figure 4. Graph presenting the relationship of calculated ΔG values (ΔG_{calc}) versus experimentally obtained ΔG values (ΔG_{exp}) in kcal/mol for binding energies of the ligands of the training set ($n = 19$), shown by the open squares. The straight line represents the line $\Delta G_{calc} = \Delta G_{exp}$ ($y = x$). The linear correlation coefficient (r^2) reflects the deviations of the experimental and calculated values from this line. The other symbols represent the ΔG_{calc} versus ΔG_{exp} of the ligands in the test set ($n = 12$). Data for PRG and 4DBB are not shown. The highest concentration that could be tested was 0.3 mM, which showed no competition with radiolabeled E2. The linear correlation coefficient (r^2) for the compounds of the test set is 0.85.

Prediction. The optimized LIE parameters ($\alpha = 0.82$, $\beta_{OH\geq 2} = 0.20$ and $\beta_{OH=1} = 0.43$) were applied to a test set of compounds with known estrogenic activity, namely, THC, API, DAI, and EQU and the PAH metabolites 1,2-di-OHCHN and 7,8-di-OHBaP, all compounds with two or more OH groups. Two monohydroxylated compounds, 3-OHBaP and 3-OHCHN, were also investigated. Finally, the binding of four ligands without hydroxyl groups, KEP, 2DBB, 4DBB, and PRG, was investigated (see Figure 3 for structures). Interaction energies were obtained from similar MD simulations as described for the other compounds. The results of the interaction energy and ΔG_{calc} calculations are shown in Table 3. The $\Delta G_{calc,i}$ values were calculated using the optimal α (0.82) and the two OH-dependent β values (0.20 and 0.43) obtained by linear regression analysis as described above. For ligands without hydroxyl group, no β parameter could be optimized; therefore, the theoretical value for β was used, i.e., $\beta_{OH=0} = 0.5$. The calculated ΔG values are plotted against the experimental ΔG values in Figure 4, with the open symbols representing the training set of compounds. The predicted ΔG values for

Table 3. Overview of the van der Waals (E^{VDW}) and Electrostatic (E^{EL}) Interaction Energy Contributions to the Interaction Energy (ΔE_{AMBER}) (A), the Calculated (ΔG_{calc}) and the Experimentally (ΔG_{exp}) Derived Free Energies of Binding (B) of the Test Set ($n = 12$) for Binding to the ERα, in All Orientations Simulated with MD[a]

(A) Calculated Binding Energies per Orientation

			receptor-bound		free				
ligand	no. of OH	orientation	E^{VDW}	E^{EL}	E^{VDW}	E^{EL}	ΔE_{AMBER}	p_i	$\Delta G_{calc,i}$
THC	2	1	−44.23	−22.83	−27.82	−30.23	−9.02	1.00	−12.03
		2	−42.90	−20.12	−27.82	−30.23	−4.97	0.00	−10.38
API	3	1	−35.53	−20.36	−20.30	−37.88	2.30	0.47	−4.26
		2	−36.45	−19.49	−20.30	−37.88	2.23	0.53	−5.05
		3	−36.57	−15.77	−20.30	−37.88	5.84	0.00	−0.01
		4	−35.63	−14.78	−20.30	−37.88	7.77	0.00	0.00
DAI	3	1	−31.47	−22.87	−19.88	−42.34	7.88	0.00	0.00
		2	−30.70	−30.79	−19.88	−42.34	0.72	0.99	−6.54
		3	−31.03	−27.42	−19.88	−42.34	3.77	0.01	−0.04
		4	−31.08	−26.52	−19.88	−42.34	4.62	0.00	−0.01
EQU	2	1	−33.46	−19.62	−20.22	−50.30	17.44	0.00	−0.01
		2	−34.96	−16.27	−20.22	−50.30	19.29	0.00	0.00
		3	−34.89	−21.90	−20.22	−50.30	13.74	0.87	−5.52
		4	−37.05	−18.61	−20.22	−50.30	14.86	0.13	−0.98
1,2-diOHCHN	2	1	−33.33	−22.78	−23.07	−25.43	−7.61	1.00	−7.91
		2	−37.20	−10.75	−23.07	−25.43	0.55	0.00	0.00
		3	−37.67	−6.60	−23.07	−25.43	4.23	0.00	0.00
		4	−36.10	−9.85	−23.07	−25.43	2.55	0.00	0.00
7,8-diOHBaP	2	1	−38.68	−17.81	−24.89	−34.22	2.63	0.00	0.00
		2	−36.91	−25.66	−24.89	−34.22	−3.45	1.00	−8.17
		3	−40.57	−15.51	−24.89	−34.22	3.04	0.00	0.00
		4	−35.35	−22.38	−24.89	−34.22	1.39	0.00	0.00
3-OHBaP	1	1	−37.62	−8.26	−24.06	−21.92	0.09	0.00	0.00
		2	−34.87	−18.07	−24.06	−21.92	−6.97	1.00	−7.27
		3	−36.35	−12.08	−24.06	−21.92	−2.45	0.00	0.00
		4	−38.69	−8.83	−24.06	−21.92	−1.55	0.00	0.00
3-OHCHN	1	1	−33.99	−20.79	−24.06	−21.92	−8.81	1.00	−7.70
		2	−32.26	−18.20	−24.06	−21.92	−4.49	0.00	0.00
		3	−35.87	−7.58	−24.06	−21.92	2.52	0.00	0.00
		4	−31.27	−15.42	−24.06	−21.92	−0.72	0.00	0.00
KEP	0	1	−40.98	0.12	−28.36	−5.79	−6.71	0.00	−0.02
		2	−42.77	−1.59	−28.36	−5.79	−10.21	1.00	−9.73
2DBB	0	1a	−28.29	−2.23	−20.37	−6.94	−3.22	0.03	−0.13
		1b	−28.22	−1.03	−20.37	−6.94	−1.95	0.00	−0.01
		1c	−28.42	−0.70	−20.37	−6.94	−1.82	0.00	−0.01
		2	−29.40	−3.16	−20.37	−6.94	−5.25	0.96	−5.33
4DBB	0	1a	−32.19	−1.62	−19.78	−5.52	−8.52	0.28	−2.36
		1b	−31.65	−1.69	−19.78	−5.52	−8.04	0.13	−1.00
		1c	−30.91	−1.78	−19.78	−5.52	−7.39	0.04	−0.31
		1d	−32.31	−1.90	−19.78	−5.52	−8.90	0.54	−4.63
PRG	0	1	−49.75	−5.53	−29.54	−27.77	2.03	1.00	−5.52

(B) Binding Energies and Affinities

ligand	ΔG_{calc}, predicted	ΔG_{exp}, observed	K_d, predicted	K_d, observed
THC	−12.03	−11.70	1.7×10^{-9}	3.0×10^{-9}
API	−9.32	−8.49	1.6×10^{-7}	6.6×10^{-7}
DAI	−6.59	−7.18	1.6×10^{-5}	5.9×10^{-6}
EQU	−6.52	−6.82	1.8×10^{-5}	1.1×10^{-5}
1,2-diOHCHN	−7.91	−8.31	1.7×10^{-6}	8.8×10^{-7}
7,8-diOHBaP	−8.17	−6.65	1.1×10^{-6}	1.4×10^{-5}
3-OHBaP	−7.27	−8.19	5.1×10^{-6}	1.1×10^{-6}
3-OHCHN	−7.70	−7.41	2.5×10^{-6}	4.0×10^{-6}
KEP	−9.75	−8.66	7.8×10^{-8}	4.9×10^{-7}
2DBB	−5.49	−5.20	1.0×10^{-4}	1.6×10^{-4}
4DBB	−8.30	>−4.8	9.0×10^{-7}	$>3 \times 10^{-4}$
PRG	−5.52	>−4.8	1.0×10^{-4}	$>3 \times 10^{-4}$

[a] The ΔG_{exp} values were calculated from K_d values determined using a sheep uterus cytosol radioligand binding assay. Structures and names of ligands and orientations are shown in Figures 3 and Figure 1, respectively. $\Delta G_{calc,i}$ represents the binding free energy for orientation i calculated with the LIE method, using $\alpha = 0.82$ and $\beta_{OH\geq 2} = 0.20$ (number of OH groups ≥ 2), $\beta_{OH=1} = 0.43$ (number of OH groups = 1) or $\beta_{OH=0} = 0.5$ (no OH groups). For each ligand, the final ΔG_{calc} equals the sum of $p_i \Delta G_{calc,i}$. All ΔE and ΔG values are given in kcal/mol. K_d is expressed in mol/L. No ΔG_{exp} could be obtained for PRG and 4DBB (highest concentration that could be tested was 0.3 mM).

these compounds are all in very good agreement with the experimental values (average deviation of 0.61 ± 0.4 kcal/mol and linear correlation coefficient between experimental and calculated ΔG values, $r^2 = 0.85$, $n = 10$) (Table 3 and Figure 4). THC was originally docked in the ER in four orientations; however, analysis of the MD runs showed that the ethyl moieties of the centrosymmetrical molecule showed a large flexibility such that THC in orientations 1 and 4 adopted the same conformation, as well as orientations 2 and 3. Therefore, the interaction energy values for orientations 1 and 4, and 2 and 3 were averaged. The predicted orientation

of binding of THC (orientations 1 and 4) to the ER is identical to the observed binding orientation in the crystal structure published by Shiau et al.[27] The LIE model predicted very low ERα affinity (K_d = 10 mM) for PRG. However, because of solubility problems, the highest concentration that could be tested in the ER binding assay was 0.3 mM, which showed no significant displacement of radiolabeled estradiol. The LIE model predicted moderate ERα affinity for 4DBB (Table 3), while no affinity was observed in vitro. Inspection of the trajectory files showed that the binding cavity is much bigger than the planar 4DBB structure. Calculating interactions with extra vacuum in the binding site may be the reason for the observed high interaction energy. Taking into account extra water molecules in the binding cavity may represent reality better; however, we have not included this in our model.

Discussion

The primary aim of this study was to develop a MD model for the binding of agonists to the ER to accurately predict binding affinity and preferred binding orientation of xenoestrogens with an LIE approach.

Differences between Sheep and Human Erα. For binding assays, we used ER from sheep uterus cytosol. Binding affinities observed for many ligands toward sheep ER, such as E2, DES, genistein, 8PN, and zearalenon, and also other ligands tested in our laboratory, which are not included in this study, were in excellent agreement with values for binding to human ERα (hERα) from the literature.[28,29] The sheep ERα (sERα) shows highly conserved identity with ER from other mammalian species. Overall homology with human ERα is 91%. However, all specific regions responsible for DNA binding, hormone binding, phosphorylation, nuclear localization, and transcription activation are even more conserved; in particular, the DNA binding domain (DBD) shows 100% similarity with hER and the ligand binding domain (LBD) shows 96.8% identity. In the LBD of human ERα, which was used for the MD simulation studies, 12 amino acids differ with sheep ERα (in helix 1 of hERα, 306, 307; in helix 2, 316, 321; in loop, 327; in helix 3, 348; at start of helix 5, 371; at start of helix 10, 467; in helix 11, 502, 507, 510, and 514).[30] Structural alignment of the LBD sequence of sheep ERα with the crystal structure coordinates of human ERα in InsightII showed that none of these amino acids are directly interacting with any of the ligands tested in the MD simulations. Furthermore, the differences between the interchanged residues are relatively small; i.e., sizes, charges, and especially hydrophobic properties of the residues are similar between human and sheep ERα. An exception is the neutral Gln 502 in hERα, which is substituted with a charged Arg residue in sheep ERα; nevertheless, its position is on the outer surface of helix 11 and distant from the binding cavity. Therefore, the differences in the amino acid sequence appear to be of no direct consequence to the observed binding affinities of the tested ligands. Thus, we conclude that the use of the ER from sheep uterus cytosol is appropriate for correlation with human ERα modeling studies.

Effect of Temperature on K_d. The MD simulations of ligand binding to the ER were performed at a temperature of 300 K, while in vitro binding studies were carried out at 277 K (4 °C). We did not observe significant differences in the ER binding affinity of E2 incubated at room temperature (293 K) compared to that from incubation at 4 °C (data not shown). Ligand binding to wild-type ER is known to be very temperature-independent.[31] Several mutations induce temperature sensitivity of ligand binding characteristics,[31−33] but very few naturally occurring temperature-unstable receptor mutants have been identified.[34] In addition, ΔG_{calc} is calculated from experimental K_d values using eq 2. Consequently, the deviation in ΔG_{calc} using T = 277 K instead of T = 300 K would be too small (8%) to be significantly resolved in ER binding studies.

Structural Analysis of MD Simulations. Evaluation of the MD simulations of the receptor−ligand complexes revealed that the secondary and tertiary structures including helix 12 remained stable during the MD simulation. The rmsd values, van der Waals and electrostatic interaction energies, and temperature were found to be constant over the last 70 ps of the MD runs for all simulations. Examination of the orientation and conformation of the amino acids directly involved in ligand binding showed that the architecture of the binding site was conserved very well during simulation. Hydrogen bond interactions between ligand, Glu353 and Arg394, and stabilizing water molecules showed good correlation with better interaction energy values. The six crystal water molecules swapped places for hydrogen bonding with the ligand and Glu353 and Arg394 during MD simulations. This emphasizes their important role in stabilizing the interactions between Glu353, Arg394, and the bound ligand. Furthermore, hydrogen bonding with these residues contributed more to the interacting energies of ligands than interaction with His524. This confirms the knowledge that the energy contribution of the interaction of the phenol ring in E2 with the ER is about 1.9 kcal/mol, while the energy contribution of the opposite hydroxyl group is about 0.6 kcal/mol.[29,35]

Aromatic Interactions. Deviation of the originally perpendicular orientation of the aromatic ring of Phe404 in the E2−ERα complex, resulting in loss of π stacking of the interacting phenol of the ligand during the MD simulations, resulted in less favorable interaction energies. Twisting of the Phe404 ring occurred especially in the absence of an aromatic ring, for example, in orientation 3 or 4 of E2. Perpendicular π-stacking conformation of Phe404 and the aromatic moiety of a ligand contributes significantly to the binding energy.[36,37] In general, molecular mechanical force fields do not perform specifically well in calculating aromatic interactions. However, this particular contribution is apparently estimated quite well using the Amber 6.0 force field because there was a good correlation between calculated and experimentally determined binding energies.

Parametrization of Ligands. In Amber 6.0, no suitable length and torsion parameters were available for a single bond between two sp^2-hybridized carbon atoms. Therefore, extra Amber parameters were developed in this study. Because of $\pi-\pi$ interactions over this bond, there is a tendency to adapt a coplanar conformation. However, from experimental studies (e.g., crystal structures) and various computational (DFT and ab initio) studies, it is known that in biphenyl, the rings

adapt a dihedral angle of ca. 40° because of repulsion of the adjacent H atoms. Ab initio calculations with genistein also showed that its phenolic ring favorably conforms to a nonplanar dihedral angle of 30° with the other rings. After optimization of the parameters, examination of the newly gained MD trajectories showed more flexibility of the biphenyls and genistein-like molecules, and the torsions over the single bond mainly resulted in dihedrals of 40° and 30°, respectively.

In general, the solubility of a ligand decreases with reduced flexibility. Consequently, the ΔG of binding to macromolecules, such as the ligand binding cavity of the ER, may be bigger for such ligands compared to more flexible ligands because of a smaller loss of entropy. In conclusion, our study shows that taking into account the flexibility of the biphenyl and genistein-like structures led to a more realistic calculation of interaction energies, and more crucially, the final results are in excellent agreement with the theory behind the LIE method.

Linear Interaction Energy Method. One of the main advantages of the LIE method over other computational methods calculating binding affinities is the fact that the LIE includes a solvent model. Since binding of the ligand to the receptor as well as binding to the solvent is taken into account, the LIE handles the desolvation free energy reasonably well. Initially, the electrostatic part was calculated by an electrostatic linear response approximation (parameter β), while the van der Waals contribution (parameter α) was calculated by calibration against experimental binding data. Since the latter is dependent on experimental data, it may implicitly take into account other contributions, such as force field errors and systematic entropy terms. However, in recent times, the electrostatic scaling factor (β) for the electrostatic contributions has been the subject of several studies and appeared to be (partially) system-dependent. Åqvist and Hansson[24] demonstrated that the presence of hydroxyl groups in solvent and solvated compound interfered with the electrostatic linear response. This was mainly attributed to the short-range character of dipolar fields and the existence of hydrogen bonding in and between the simulated states (solute and solvent in their studies). Marelius et al. improved the application of the LIE by applying different values for β for ligands with different polarity.[26] They divided compounds into different β classes: charged, dipolar without hydroxyl groups, dipolar with one hydroxyl group, dipolar with two hydroxyl groups, and dipolar with two or more hydroxyl groups. With the use of the different β's, they found good correlation between experimental and calculated binding energies.

The development of an accurate model for ligand binding to the ER is particularly susceptible to deviations such as those described above, since the hydrogen bonds between ER and a ligand provide interactions that are crucial for binding for most ligands and significantly enhance the binding affinity.[29] We simulated the binding to the ER with a training set consisting of 11 ligands with two or more hydroxyl groups and 10 ligands with only one hydroxyl group. Both groups were large enough and also the binding affinities covered a wide enough range to allow us to perform a double regression analysis to gain optimal values for α and β. The results are shown in the correlation graph (Figure 4). The optimal values for ligands with two hydroxyl groups were $\alpha = 0.82$ and $\beta_{OH\geq 2} = 0.20$ ($n = 11$). Subsequently, when $\alpha = 0.82$ was used, a specific value of $\beta_{OH=1} = 0.43$ was obtained for monohydroxylated compounds ($n = 8$).

Four compounds without hydroxyl groups were included in the test set for the LIE method. KEP and 2DBB show moderate and low binding affinity for the ER, respectively, and for PRG and 4DBB no affinity could be determined. When the theoretical value for β (0.5) was used, good estimates were made for the binding affinities for KEP and 2DBB (Table 3). Nevertheless, for ligands without a hydroxyl group, it would be ideal to determine an optimal value for β from a larger data set and to include these in our model. However, compounds without a hydroxyl group that possess affinity for the ER are scarce.

Different Binding Orientations. In the currently described LIE method, for each ligand, orientations are included that represent all possibilities. However, molecules will be present in the LBD of a receptor most of the time in the orientation with the best interaction energy. Therefore, the fact that the orientation with the best interaction energy, predicted by our model for DES, E2, and THC, matches those found in the crystal structures of ERα LBD complexed with these compounds indicates the good quality of our model. Nevertheless, other orientations of DES and E2 may also significantly contribute to the total binding affinity.

Another observation from protein X-ray studies is the binding mode of the antagonists ICI 182,780 and ICI 164,384. These ER antagonists are 7α-substituted derivatives of E2 and bind with very high affinity to both subtypes of the ER.[8,38] Their side chains interact with specific residues, preventing helix 12 from associating with the rest of the LBD. To do so, the side chain must protrude from the binding cavity via the 11β position of the E2 moiety.[6,39] Consequently, these molecules must bind in an "upside down" orientation similar to orientation 2 of E2 in our model, and apparently, the ER hosts this kind of binding. This also supports our hypothesis that E2, and also other compounds, may bind in several orientations.

Remarkably, the most favorable binding orientation of GEN predicted by our model (orientation 3) is distinct from its binding orientation observed in the crystal structure of the ERβ, which is equivalent to orientation 1 of E2.[10] During the MD simulation, the perpendicular (T-shaped) aromatic interaction between the Phe404 and the phenolic ring of GEN was distorted, while for orientations 3 and 4 this interaction was stable, and this led to higher interaction energies. In our model, with orientations 3 and 4, the flavone portion of GEN occupies the narrow part of the binding cavity, similar to the "A" ring of E2 in orientation 1. The O4 moiety interacts with the Glu353 and Arg394 in the ER, and the phenolic ring interacts with His524. In contrast, in the crystal structure of GEN complexed to ERβ, the phenolic ring occupies the same position as the "A" ring of E2 in orientation 1. These differences in orientation of GEN in ERα and ERβ may be related to the fact that GEN is a full agonist for ERα and a partial agonist for ERβ, inducing an antagonized receptor conformation.[10]

Although the architecture of the binding cavity of ERβ shows great similarity to the binding cavity of ERα, it has been shown that ERα and ERβ respond differently to agonists and antagonists.[7,8] GEN binds with moderate affinity to both ERs; however, it shows significant higher affinity for the ERβ ($\sim$30-fold).[7] The two amino acid differences between the residues of the binding cavity of ERα and ERβ may have a direct effect on the observed differences in ligand binding preferences.[8] Because of the "inward" orientation of its side chain, the methionine at position 336 in ERβ occupies a larger volume in the binding cavity than the corresponding leucine at position 384 in ERα, possibly favoring another binding orientation of GEN. Also, the side chain of isoleucine 373 in ERβ occupies more binding cavity space than the corresponding methionine 421 in ERα. Both substitutions are lining the binding cavity at the histidine 524 side of E2 and may have significant effects on the preferred binding modes of GEN and other ligands.

Conclusions

A computational model was developed for prediction of binding affinities of ligands to the ERα using MD simulations in combination with the linear interaction energy (LIE) approach. We obtained an excellent linear correlation ($r^2 = 0.94$, $n = 19$) between experimental and calculated ΔG values for compounds that bind to the ERα with K_d values ranging from 0.15 mM to 30 pM, a 5 000 000-fold difference in binding affinity. The excellent correlation for the compounds of the training set, which are structurally very diverse, is remarkable. The predictive value of our model is shown by predicting binding affinities of a collection of structurally diverse estrogenic compounds that are in very good agreement with the experimental values; the predicted ΔG values for these compounds showed an average deviation of only 0.61 kcal/mol for absolute binding affinities.

An important advantage of the LIE method is the fact that it includes a solvent model and takes into account all possible binding orientations. The LIE approach provides a very good method for prediction of absolute ligand binding affinities, as well as binding orientation of ligands. The presented LIE model is of great value for the prediction of estrogenic activity of compounds such as xenoestrogens and potential metabolites of estrogens, which are hard to generate or isolate and therefore difficult to test in vitro but may be applied to other systems as well.

Experimental Section

Chemicals. ^{3}[H]-E2 was obtained from Amersham (Buckinghamshire, U.K.). Dextran (MW 60000−90000) was obtained from Duchefa, Haarlem, The Netherlands. 8-Prenylnaringenin was a kind gift from Dr. Fred Stevens, Department of Bioorganic Chemistry, Vrije Universiteit, Amsterdam, The Netherlands.[40] Pure hydroxybenzo[*a*]pyrenes (purity, $\geq$98%) were obtained from the Chemical Carcinogen Reference Standard Repository of the National Cancer Institute (KS). Pure hydroxychrysenes were kindly provided by Dr. Albrecht Seidel, Biochemical Institute for Environmental Carcinogens, Hamburg, Germany. All other chemicals were purchased from Sigma-Aldrich Co. (St. Louis, MO).

Competitive Ligand Binding Assay. Competitive binding assays were performed as described before.[41] Briefly, compounds were incubated with sheep uterus cytosol and [^{3}H]-E2 for 3 h at 4 °C. Receptor-bound [^{3}H]-E2 was separated from unbound [^{3}H]-E2 by charcoal−dextran absorption and centrifugation and measured in a 1900 Tricarb scintillation counter (Packard Instruments, Perkin-Elmer Wellesley) after the addition of scintillation cocktail (HiSafe 3, Perkin-Elmer). The ligand binding affinities (K_d) were calculated with Graphpad Prism 3.0 (GraphPad Software Inc., 2000, San Diego, CA). All compounds were tested for binding affinity in our laboratory except for the DMBA metabolites and 11β-chloroethyl-E2 (E2-Cl), for which K_d values were taken from the literature.[42,43]

Computational Details. Ligand Parametrization. For the prediction of ligand binding affinities, docking studies were performed with the various ligands of the training set ($n = 19$) and the test set ($n = 12$). Selection of the investigated ligands was based on structural diversity and a wide range of binding affinities for the ER. Initial structures of the ligands were generated with the xLEaP module of Amber 6.0. Coordinates of crystal structures of ligands were used for E2 (Protein Data Bank (PDB) code 1ERE), DES (PDB code 3ERD), GEN (PDB code 1QKM), ZEA (Cambridge Structural Database, CSD), and THC (PDB code 1L2J). Subsequently, conformation analysis and subsequent hierarchical clustering were performed using Sybyl 6.7 (Tripos Inc., St Louis, MO) to obtain four energy-optimized geometries covering the broadest range of the conformational space of the ligand. Electrostatic potentials were calculated for all conformers using GAMESS at the 6-31G* level of theory.

In the cases of E2, EST, and ZEA, which consist of a flexible, closed multiple ring system, conformation analyses were performed with Macromodel (Schrödinger, Inc.). In the cases of GEN and biphenyls, conformation analysis resulted in one optimal conformation. In these cases, geometry optimizations with several fixed torsions over the rotatable bond between two aromatic rings were performed with GAMESS at the STO-3G level of theory. Because the hydroxylated metabolites of BaP and CHN are planar, no conformational analysis was performed on the structures of their metabolites; however, two conformations with opposite orientations of the hydroxyl hydrogen in the plane were used for geometry optimization. For all ligands, the generated optimized structures were used as different conformers in the calculation of electrostatic potentials at the 6-31G* level of theory. Calculated electrostatic potentials were subsequently applied in the multiconformational restraint electrostatic potential (RESP) fitting procedure to obtain atomic charges for all compounds.[44]

Extra Amber 6.0 Ligand Parameters. The Amber 6.0 package comprised libraries for all atom types for the amino acids in the LBD of the ER. Most of the atom types used for the ligands were also already parametrized to our satisfaction except for torsions between nonsubstituted aromatic and nonaromatic sp^2-hybridizedcarbon atoms in DES, GEN, ZEA, and two aromatic rings in biphenyls. Their torsional force constants were optimized using an ab initio approach. The potential energy surface of rotation over the dihedral angle as calculated in GAMESS using the STO-3G basis set for geometry optimizations and the 6-31G* basis set for energy calculations, compared to the potential energy surface in Amber 6.0 (data not shown), showed that the ideal force constant for the single bond between the two sp^2-hybridized carbons in Amber should be 4 kcal/mol for GEN, DES, and ZEA and 3 kcal/mol for the biphenyls. Subsequently, these force constants were used in all MD simulations. Optimal bond lengths were based on those found in the crystal structures of GEN and DES, and biphenyls, and set to 1.40 and 1.51 Å respectively.

ERα LBD Structure. The crystal structure of the LBD (residues Ser 305 to Leu 549) of the ERα complexed with DES (PDB code 3ERD)[3] was used to build the protein model, and all molecular dynamics simulations were based on this protein structure. Missing side chains, from surface loops, were modeled using the homology module of InsightII (Biosym, San Diego, CA). The terminal amino acids were treated as charged.

The charges of the ionizable groups were set to correspond to a pH of 7.4, resulting in a net charge of -5 e$^-$. Coordinates of six crystallographic water molecules, directly interacting with either the ER residues GLU353 and ARG394 or the 3-OH group of DES, were also obtained from the published crystal structure.

Docking of the Ligands. Positions of E2 (PDB code 1ERE) and DES in the ER LBD from the published crystal structures[3,6] were used as templates for the docking of the ligands. Ligands were placed in the binding site of the ER by manual docking, using InsightII. If available, ligand crystal structure coordinates were used; otherwise, low-energy conformations obtained by GAMESS were applied. In the direct surroundings of the "A" ring of E2, the ER binding site is rather narrow and flat, allowing little variation in the position of the aromatic ring. The remainder of the cavity is tolerant for a variety of hydrophobic groups. To calculate the preferred orientation of ligands, E2 was docked in four orientations as depicted in Figure 1B. Subsequently, all ligands were docked in four different orientations, thereby fitting aromatic moieties on the "A" ring of E2 in its four orientations. Special care was taken in minimizing overlap between protein and ligand and, when appropriate, in adjusting the orientation of OH groups in the direction of the interacting amino acid residues, in particular Glu353, Arg394, or His524.

The 12-methyl group in 12-methylbenz[*a*]anthracenediol (12-MBA-diol) and 7,12-dimethylbenz[*a*]anthracenediol (BA-diol) causes these structures to be nonplanar, with a dihedral angle involving the methyl group of 22°. This dihedral angle was observed in several dimethylbenz[*a*]anthracene (DMBA) crystal structures (CSD) and was found after geometry minimizations in GAMESS (STO-3G). Since the optimal dihedral may be 22° or $-22°$ and the transition between these states is unlikely to occur in MD simulations because of the rather high energy barrier (17.0 kcal/mol), we docked both isomers in four orientations, resulting in eight final orientations for each of the two benz[*a*]anthracenes. Similarly, the large bromo substituents in the biphenyl structures prohibit complete rotation over the single bond connecting the phenyl rings. Therefore, these compounds were also docked in eight orientations. Because of the large bromo substituents, some of these orientations, however, experienced severe steric hindrance, and these were not included in further studies.

MD Simulations of Ligands Bound to the ER and Free Ligands in Water. The starting structure of each receptor−ligand was solvated in water using the tLEaP module of Amber 6.0. The complex was surrounded by a rectangular box of equilibrated water molecules, based on the TIP3P water model.[45] The minimum distance between the complex and the boundaries was 6 Å, resulting in typical box volume sizes of about 3.0×10^6 Å^3, and contained ca. 20 000 atoms including ER and ligand. All calculations were carried out in the Sander module of Amber 6.0 under periodic boundary conditions with a cutoff distance of 10.0 Å.

Subsequently, MD simulations were performed to equilibrate the total volume of the water box for 20 ps. MD runs were performed at constant pressure (1 atm) and constant temperature (300 K). During equilibration, positional restraints were applied on the backbone C_α's of the ER and the crystal waters in the ER to fix their Cartesian coordinates. After the system had adapted suitable volume and pressure, an energy minimization was carried out for 10 ps, using steepest descent for 0.01 ps, followed by the full conjugate gradient minimization method. The positional restraints remained applied to the backbone C_α's and the crystal waters during minimization.

Finally, the minimized structures of the solvated receptor−ligand complexes were used for MD simulations. Binding of the ligands to the ERα was simulated for 100 ps at constant pressure. No restraints were applied during these MD runs. During the first 15 ps, the temperature was raised from 0 to 300 K. During the subsequent 35 ps, velocities and coordinates were saved every picosecond. These trajectory files were used for analysis in the Carnal/Anal module of Amber 6.0 for structural stability and hydrogen bonding and for calculation of the energy terms for the linear interaction energy method.

Simulations of free ligands in water were performed in a similar protocol. Ligands were solvated by surrounding the ligands with a box of water molecules (TIP3P) with a minimum distance between the ligand and the boundaries of 20 Å, resulting in a box volume of about 1.5×10^6 Å^3 with ca. 10 000 atoms. All calculations were carried out under periodic boundary conditions with a cutoff distance of 10.0 Å. Equilibration of the water box (20 ps), minimization (10 ps), and subsequent MD simulations (100 ps) were carried out as described for the receptor−ligand complex. No restraints were applied. For each ligand, 50 trajectory files were obtained and the last 35 (constant temperature) were used for analysis.

References

(1) McLachlan, J. A.; Korach, K. S.; Newbold, R. R.; Degen, G. H. Diethylstilbestrol and other estrogens in the environment. *Fundam. Appl. Toxicol.* **1984**, *4*, 686−691.

(2) Åqvist, J.; Medina, C.; Samuelsson, J. E. A new method for predicting binding affinity in computer-aided drug design. *Protein Eng.* **1994**, *7*, 385−391.

(3) Shiau, A. K.; Barstad, D.; Loria, P. M.; Cheng, L.; Kushner, P. J.; Agard, D. A.; Greene, G. L. The structural basis of estrogen receptor/coactivator recognition and the antagonism of this interaction by tamoxifen. *Cell* **1998**, *95*, 927−937.

(4) Case, D. A.; Pearlman, D. A.; Caldwell, J. W.; Cheatham, T. E., III; Wang, J.; Ross, W. S.; Simmerling, C.; Darden, T.; Merz, K. M.; Stanton, R. V.; Cheng, A.; Vincent, J. J.; Crowley, M.; Tsui, V.; Gohlke, H.; Radmer, R.; Duan, Y.; Pitera, J.; Massova, I.; Seibel, G. L.; Singh, U. C.; Weiner, P.; Kollman, P. A. Amber, 1999. http://amber.scripps.edu/. Amber is the collective name for a suite of programs that allow users to carry out molecular dynamics simulations, particularly on biomolecules.

(5) Evans, R. M. The steroid and thyroid hormone receptor superfamily. *Science* **1988**, *240*, 889−895.

(6) Brzozowski, A. M.; Pike, A. C.; Dauter, Z.; Hubbard, R. E.; Bonn, T.; Engstrom, O.; Ohman, L.; Greene, G. L.; Gustafsson, J. A.; Carlquist, M. Molecular basis of agonism and antagonism in the oestrogen receptor. *Nature* **1997**, *389*, 753−758.

(7) Kuiper, G. G.; Carlsson, B.; Grandien, K.; Enmark, E.; Haggblad, J.; Nilsson, S.; Gustafsson, J. A. Comparison of the ligand binding specificity and transcript tissue distribution of estrogen receptors alpha and beta. *Endocrinology* **1997**, *138*, 863−870.

(8) Barkhem, T.; Carlsson, B.; Nilsson, Y.; Enmark, E.; Gustafsson, J.; Nilsson, S. Differential response of estrogen receptor alpha and estrogen receptor beta to partial estrogen agonists/antagonists. *Mol. Pharmacol.* **1998**, *54*, 105−112.

(9) Torchia, J.; Rose, D. W.; Inostroza, J.; Kamei, Y.; Westin, S.; Glass, C. K.; Rosenfeld, M. G. The transcriptional co-activator p/CIP binds CBP and mediates nuclear-receptor function. *Nature* **1997**, *387*, 677−684.

(10) Pike, A. C.; Brzozowski, A. M.; Hubbard, R. E.; Bonn, T.; Thorsell, A. G.; Engstrom, O.; Ljunggren, J.; Gustafsson, J. A.; Carlquist, M. Structure of the ligand-binding domain of oestrogen receptor beta in the presence of a partial agonist and a full antagonist. *EMBO J.* **1999**, *18*, 4608−4618.

(11) IARC Benzo[*a*]pyrene. *IARC Monograms on the Evaluation of the Carcinogenic Risk of Chemicals to Man*; International Agency for Research on Cancer: Lyon, France, 1973; pp 91−136.

(12) Oost, R. v. d.; Beyer, J.; Vermeulen, N. P. E. Fish biomarkers and environmental risk assessment: a review. *Environ. Toxicol. Appl. Pharmacol.*, in press.

(13) Alexandrov, K.; Cascorbi, I.; Rojas, M.; Bouvier, G.; Kriek, E.; Bartsch, H. CYP1A1 and GSTM1 genotypes affect benzo[*a*]-pyrene DNA adducts in smokers' lung: comparison with aromatic/hydrophobic adduct formation. *Carcinogenesis* **2002**, *23*, 1969−1977.

(14) Melendez-Colon, V. J.; Luch, A.; Seidel, A.; Baird, W. M. Cancer initiation by polycyclic aromatic hydrocarbons results from formation of stable DNA adducts rather than apurinic sites. *Carcinogenesis* **1999**, *20*, 1885−1891.

(15) Weinstein, I. B.; Jeffrey, A. M.; Jennette, K. W.; Blobstein, S. H.; Harvey, R. G.; Harris, C.; Autrup, H.; Kasai, H.; Nakanishi, K. Benzo[*a*]pyrene diol epoxides as intermediates in nucleic acid binding in vitro and in vivo. *Science* **1976**, *193*, 592−595.

(16) Charles, G. D.; Bartels, M. J.; Zacharewski, T. R.; Gollapudi, B. B.; Freshour, N. L.; Carney, E. W. Activity of benzo[*a*]pyrene and its hydroxylated metabolites in an estrogen receptor-alpha reporter gene assay. *Toxicol. Sci.* **2000**, *55*, 320−326.

(17) van Lipzig, M. M. H.; Vermeulen, N. P. E.; Meerman, J. H. N. Unpublished results.

(18) Kramer, V. J.; Giesy, J. P. Specific binding of hydroxylated polychlorinated biphenyl metabolites and other substances to bovine calf uterine estrogen receptor: structure-binding relationships. *Sci. Total Environ.* **1999**, *233*, 141−161.
(19) Garner, C. E.; Jefferson, W. N.; Burka, L. T.; Matthews, H. B.; Newbold, R. R. In vitro estrogenicity of the catechol metabolites of selected polychlorinated biphenyls. *Toxicol. Appl. Pharmacol.* **1999**, *154*, 188−197.
(20) You, L.; Casanova, M.; Bartolucci, E. J.; Fryczynski, M. W.; Dorman, D. C.; Everitt, J. I.; Gaido, K. W.; Ross, S. M.; Heck, Hd. H. Combined effects of dietary phytoestrogen and synthetic endocrine-active compound on reproductive development in Sprague−Dawley rats: genistein and methoxychlor. *Toxicol. Sci.* **2002**, *66*, 91−104.
(21) Tong, W.; Perkins, R.; Xing, L.; Welsh, W. J.; Sheehan, D. M. QSAR models for binding of estrogenic compounds to estrogen receptor alpha and beta subtypes. *Endocrinology* **1997**, *138*, 4022−4025.
(22) Bissantz, C.; Folkers, G.; Rognan, D. Protein-based virtual screening of chemical databases. 1. Evaluation of different docking/scoring combinations. *J. Med. Chem.* **2000**, *43*, 4759−4767.
(23) Oostenbrink, B. C.; Pitera, J. W.; van Lipzig, M. M.; Meerman, J. H.; van Gunsteren, W. F. Simulations of the estrogen receptor ligand-binding domain: affinity of natural ligands and xenoestrogens. *J. Med. Chem.* **2000**, *43*, 4594−4605.
(24) Åqvist, J.; Hansson, T. On the validity of electrostatic linear response in polar solvents. *J. Phys. Chem.* **1996**, *100*, 9512−9521.
(25) Hansson, T.; Marelius, J.; Åqvist, J. Ligand binding affinity prediction by linear interaction energy methods. *J. Comput.-Aided Mol. Des.* **1998**, *12*, 27−35.
(26) Marelius, J.; Hansson, T.; Åqvist, J. Calculation of ligand binding free energies from molecular dynamics simulations. *Int. J. Quantum Chem.* **1998**, *69*, 77−88.
(27) Shiau, A. K.; Barstad, D.; Radek, J. T.; Meyers, M. J.; Nettles, K. W.; Katzenellenbogen, B. S.; Katzenellenbogen, J. A.; Agard, D. A.; Greene, G. L. Structural characterization of a subtype-selective ligand reveals a novel mode of estrogen receptor antagonism. *Nat. Struct. Biol.* **2002**, *9*, 359−364.
(28) Waller, C. L.; Oprea, T. I.; Chae, K.; Park, H. K.; Korach, K. S.; Laws, S. C.; Wiese, T. E.; Kelce, W. R.; Gray, L. E., Jr. Ligand-based identification of environmental estrogens. *Chem. Res. Toxicol.* **1996**, *9*, 1240−1248.
(29) Anstead, G. M.; Carlson, K. E.; Katzenellenbogen, J. A. The estradiol pharmacophore: ligand structure−estrogen receptor binding affinity relationships and a model for the receptor binding site. *Steroids* **1997**, *62*, 268−303.
(30) Madigou, T.; Tiffoche, C.; Lazennec, G.; Pelletier, J.; Thieulant, M. L. The sheep estrogen receptor: cloning and regulation of expression in the hypothalamo-pituitary axis. *Mol. Cell. Endocrinol.* **1996**, *121*, 153−163.
(31) Reese, J. C.; Katzenellenbogen, B. S. Characterization of a temperature-sensitive mutation in the hormone binding domain of the human estrogen receptor. Studies in cell extracts and intact cells and their implications for hormone-dependent transcriptional activation. *J. Biol. Chem.* **1992**, *267*, 9868−9873.
(32) Pike, J. W.; Dokoh, S.; Haussler, M. R.; Liberman, U. A.; Marx, S. J.; Eil, C. Vitamin D3−resistant fibroblasts have immunoassayable 1,25-dihydroxyvitamin D3 receptors. *Science* **1984**, *224*, 879−881.
(33) Northrop, J. P.; Gametchu, B.; Harrison, R. W.; Ringold, G. M. Characterization of wild type and mutant glucocorticoid receptors from rat hepatoma and mouse lymphoma cells. *J. Biol. Chem.* **1985**, *260*, 6398−6403.
(34) Tora, L.; Mullick, A.; Metzger, D.; Ponglikitmongkol, M.; Park, I.; Chambon, P. The cloned human oestrogen receptor contains a mutation which alters its hormone binding properties. *EMBO J.* **1989**, *8*, 1981−1986.
(35) Fanchenko, N. D.; Sturchak, S. V.; Shchedrina, R. N.; Pivnitsky, K. K.; Novikov, E. A.; Ishkov, V. L. The specificity of the human uterine receptor. *Acta Endocrinol.* **1979**, *90*, 167−175.
(36) Tsuzuki, S.; Honda, K.; Azumi, R. Model chemistry calculations of thiophene dimer interactions: origin of π-stacking. *J. Am. Chem. Soc.* **2002**, *124*, 12200−12209.
(37) Jaffe, R.; Smith, G. A quantum chemistry study of benzene dimer. *J. Chem. Phys.* **1996**, *105*, 2780−2788.
(38) Kuiper, G. G.; Lemmen, J. G.; Carlsson, B.; Corton, J. C.; Safe, S. H.; van der Saag, P. T.; van der Burg, B.; Gustafsson, J. A. Interaction of estrogenic chemicals and phytoestrogens with estrogen receptor beta. *Endocrinology* **1998**, *139*, 4252−4263.
(39) Pike, A. C.; Brzozowski, A. M.; Walton, J.; Hubbard, R. E.; Thorsell, A. G.; Li, Y. L.; Gustafsson, J. A.; Carlquist, M. Structural insights into the mode of action of a pure antiestrogen. *Structure (London)* **2001**, *9*, 145−153.
(40) Milligan, S. R.; Kalita, J. C.; Pocock, V.; Van De Kauter, V.; Stevens, J. F.; Deinzer, M. L.; Rong, H.; De Keukeleire, D. The endocrine activities of 8-prenylnaringenin and related hop (*Humulus lupulus* L.) flavonoids. *J. Clin. Endocrinol. Metab.* **2000**, *85*, 4912−4915.
(41) Murk, A. J.; Legler, J.; van Lipzig, M. M. H.; Meerman, J. H. N.; Belfroid, A. C.; Spenkelink, A.; van der Burg, B.; Rijs, G. B.; Vethaak, D. Detection of estrogenic potency in wastewater and surface water with three in vitro bioassays. *Environ. Toxicol. Chem.* **2002**, *21*, 16−23.
(42) Morreal, C. E.; Sinha, D. K.; Schneider, S. L.; Bronstein, R. E.; Dawidzik, J. Antiestrogenic properties of substituted benz[*a*]-anthracene-3,9-diols. *J. Med. Chem.* **1982**, *25*, 323−326.
(43) Bindal, R. D.; Carlson, K. E.; Reiner, G. C.; Katzenellenbogen, J. A. 11 beta-chloromethyl-[3H]estradiol-17 beta: a very high affinity, reversible ligand for the estrogen receptor. *J. Steroid Biochem.* **1987**, *28*, 361−370.
(44) Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges. The Resp Model. *J. Phys. Chem.* **1993**, *97*, 10269−10280.
(45) Jorgensen, W. L.; Chandraskhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926−935.

JM0309607