

Article

Quantum Chemical GA-MLR, Cluster Model, and Conceptual DFT Descriptors Studies on the Binding Interaction of Estrogen Receptor Alpha with Endocrine Disrupting Chemicals

Shu-Chun Chi ¹, Hsing-Cheng Hsi ¹  and Chia-Ming Chang ^{2,*} 

¹ Air Pollution and Environmental Material Lab, Graduate Institute of Environmental Engineering, National Taiwan University, Taipei 10617, Taiwan

² Environmental Molecular and Electromagnetic Physics (EMEP) Laboratory, Department of Soil and Environmental Sciences, National Chung Hsing University, Taichung 40227, Taiwan

* Correspondence: abinitio@dragon.nchu.edu.tw

Abstract: In the present study, the predication of the binding affinity (log RBA) of estrogen receptor alpha with three categories of environmental endocrine disrupting chemicals (EDCs), namely, PCB, phenol, and DDT, is performed by the quantum chemical genetic algorithm multiple linear regression (GA-MLR) method. The result of the optimal model indicates that log RBA increases with increasing the electrophilicity and hydrophobicity of EDCs. However, by using the quantum chemical cluster model approach, the modeling results reveal that electrostatic interaction and hydrogen bonding play a significant role. The chemical reactivity descriptors calculated based on the conceptual density functional theory also indicate that the binding mechanism of charge-controlled interaction is superior to that of frontier-controlled interaction.

Keywords: estrogen receptor alpha; endocrine disrupting chemicals; quantum chemical genetic algorithm multiple linear regression; quantum chemical cluster model; conceptual density functional theory



Citation: Chi, S.-C.; Hsi, H.-C.; Chang, C.-M. Quantum Chemical GA-MLR, Cluster Model, and Conceptual DFT Descriptors Studies on the Binding Interaction of Estrogen Receptor Alpha with Endocrine Disrupting Chemicals. *Crystals* **2023**, *13*, 228. <https://doi.org/10.3390/cryst13020228>

Academic Editor: Thomas M. Klapötke

Received: 31 December 2022

Revised: 24 January 2023

Accepted: 25 January 2023

Published: 27 January 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Endocrine disrupting chemicals (EDCs) are a class of chemicals that come from the outside world, which can interfere with the production, release, transport, metabolism, binding, and elimination processes of hormones in the body [1]. Many household and industrial products are EDCs, however, their disposal results in the release of many chemicals into the ecosystem, adversely affecting environmental and human health. One of the most extensively studied nuclear receptor targets associated with endocrine disrupting effects is the estrogen receptor alpha (ER α) [2]. The binding mechanisms of the stable ER α -EDC complexes are thought to be specific hydrogen bonds and hydrophobic interactions in the ligand binding pocket (LBP) [3]. Furthermore, the EDCs with larger sizes have higher binding affinity from the viewpoint of hydrophobic interactions [4].

Using quantum chemical descriptors, several predicting models have been applied to account for molecular and electronic properties that influence estrogen potency of EDCs. For example, the higher values of dipole moment indicate that higher polarity will result in larger intermolecular interactions. [5]. It has been found that Bis AF (4.762 Debye) and Bis S (5.571 Debye) have higher values of dipole moment, which proves the ligands bind to receptors with higher binding affinity [6]. The previous literature has shown that DDT and its metabolite (DDE) and hydroxychloride (HPTE) can bind strongly to hER α LBD [7,8]. Metabolic hydroxylation of aromatic compounds was found to enhance the binding affinity of PCB and DDT. The higher polar surface area and atomic partial charges supported that the derivatives exhibit stronger electrostatic and hydrogen bonding interactions [9]. In addition, the hydrophobicity also has a significant effect on the binding affinity [10,11].

The objective of this study is to investigate the interaction mechanism of PCB, phenol, and DDT binding to ER α . By using the quantum chemical GA-MLR method, a model is developed to predict the binding affinity (log RBA). In addition, the quantum chemical cluster model approach is used to advance the understanding of the binding interaction of PCB, phenol, and DDT by providing structures and electronic properties in detail. From the conceptual density functional theory (CDFT) perspective, the global and local reactivity descriptors are also calculated to clarify the binding mechanism.

2. Computational Details

2.1. Dataset

The dataset used to construct and validate the GA-MLR model is from the EDKB database, which has been divided into 14 categories according to chemical structure, namely, benzene, DDT, DES, flutamide, NoRing, PAH, PCB, pesticide, phenol, phthalate, phytoestrogen, siloxane, steroid, and other [12]. We chose three categories, PCB, phenol, and DDT, as ligands for this study. In addition, observed estrogen activities are expressed by relative binding affinity (RBA = (E2 IC50/Competitor IC50) \times 100) as the experimental dataset of endocrine disrupting chemicals.

2.2. Quantum Chemical Descriptors

In this study, the following quantum chemical descriptors are used to construct the GA-MLR model: the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO) energies can be used to indicate the electron-donating and electron-accepting ability, respectively, which are two important descriptors affecting the biological activity of compounds [13]. Ionization potential (IP) and electron affinity (EA) have similar implications to the frontier molecular orbital energies HOMO and LUMO [14–16]. The chemical potential (μ) drives the charge transfer reaction between estrogen receptor and endocrine disruptor, which can be divided into charge-acceptance (μ^+) and charge-donation (μ^-) parts [17]. Chemical hardness (η) is the resistance to electron redistribution, while softness (S) is the inverse of chemical hardness and is correlated with the molecular polarizability [18]. Dipole moments can be used to indicate molecular polarity [19]. Polar surface area (PSA) and apolar surface area (APSA) are related to hydrophilic and hydrophobic interactions, respectively.

2.3. GA-MLR Method

Based on the fingerprint calculation and diversity selection method, the DTC Lab software package (<https://dtclab.webs.com/software-tools>) (accessed on 4 June 2022) divided the dataset into a training set and a test set (the ratio of the two sets of data in this study was set to 75%:25%). The genetic algorithm multiple linear regression (GA-MLR) method [20] was used to perform the feature selection out of the test set. The criteria for optimal model construction have been discussed in the previous literature [21]. By using acceptable thresholds for internal and external validation metrics, it is ensured that the predicting model is robust and reliable. The thresholds for validation metrics are: $R^2 > 0.6$, $R^2_{\text{adj}} > 0.6$, $Q^2_{\text{F1}} \geq 0.5$, $Q^2_{\text{F2}} \geq 0.5$, $\overline{r^2_m} > 0.5$, $\Delta r^2_m < 0.2$, and $\text{CCC} > 0.85$ [22–26]. According to both the leverage and the standardization approaches, the applicability domain (AD) of the predicting model has been defined as the physicochemical, structural or biological space, knowledge, or information that the training set of the developed model has and is suitable for conducting predictions for new compounds [27].

2.4. Molecular Docking

The X-ray crystal structure of the ER α receptor (pdb code: 3ERT) was retrieved from the RCSB protein database (www.rcsb.org) (accessed on 6 June 2022), and all non-bonded water molecules and ligands were removed. Then, the endocrine disrupting chemicals were docked with ER α via the AutoDock Vina, a protein-ligand docking program packaged by the AutoDock Vina Extended SAMSON Extension [28]. The overall size of ER α was

set to the search domain and the center of the grid, and molecular docking simulations were performed using default parameters. Finally, among the top 200 docking poses, the best conformation was selected as the bioactive conformation of the ligand according to a standard scoring function.

2.5. Quantum Chemical Cluster Model Approach

Geometry optimization calculations were performed using AM1 Hamiltonian by MOPAC 2016 Quantum Chemistry software [29]. All amino acids in the cluster model of the ER α receptor (pdb code: 3ERT) were truncated at the α -carbon, and hydrogen atoms were added manually. During geometry optimization, the truncated α -carbon remains fixed in its input position. The protonation states of residues are derived from experimental evidence. The solvent effect is implicitly treated by the conductor-like screening model (COSMO), which uses a dielectric constant of 78.4 for water [29].

2.6. Density Functional Theory Calculations

All DFT calculations were performed by the Gaussian 16 software package [30]. The electronic properties were obtained using the M06-2X density functional method and the 6-31G(d,p) basis set. The solvent effect of water is modeled using the Self-Consistent Reaction Field (SCRF) method. For the local reactivity descriptors, the maximum partial charge of the hydrogen atom ($\rho^+_{\max(\text{H})}$), the maximum partial charge (ρ^+_{\max}), the maximum nucleophilic Fukui function (f^+_{\max}), the maximum electrophilic Fukui function (f^-_{\max}), the maximum nucleophilic condensed local softness (s^+_{\max}), and the maximum electrophilic condensed local softness (s^-_{\max}), have been obtained at the same level of the above DFT method. Polar surface area (PSA) and apolar surface area (APSA) were calculated using VEGAZZ software (PSA, probe radius = 1.4, density = 10) [31].

3. Results and Discussion

3.1. Quantum Chemical GA-MLR Model

Table 1 summarizes the M06-2X/6-31G(d,p)/SCRF calculated descriptors for various endocrine disrupting chemicals. Quantum chemical descriptors have explicit physical meaning and help to elucidate many aspects of chemical–biological interactions. Equation (1) represents the four quantum chemical descriptors used to construct the optimal GA-MLR model, namely, LUMO, μ^- , dipole moment, and APSA, which can be used to indicate electrophilicity, polarity, and hydrophobicity. The observed, predicted, and residual values of NCTR log RBA are listed in Table 2. The correlation plot of the observed and predicted values is shown in Figure 1. Statistical analysis of the model for NCTR log RBA obtained by the quantum chemical GA-MLR method yields R^2 and R^2_{adj} of 0.9101 and 0.8911, respectively. The values for Q^2_{F1} and Q^2_{F2} are 0.7820 and 0.7813, respectively, indicating that the training and test sets are close to the mean. The concordance correlation coefficient (CCC) is 0.8968, representing that the predictive model is reliable. The above values are all in line with the validation criteria of the model; therefore, the developed model has good robustness and predictive power.

Table 1. M06-2X/6-31G(d,p)/SCRF calculated descriptors for the quantum chemical GA-MLR method.

Category	Compound	HOMO (eV)	LUMO (eV)	μ^+ (a.u.)	μ^- (a.u.)	Dipole Moment (D)	APSA (\AA^2)
PCB	2,3,4,5-Tetrachloro-4'-biphenylol	−7.502	−0.430	−0.070	−0.182	4.753	399.6
	2,5-Dichloro-4'-biphenylol	−7.407	−0.083	−0.088	−0.186	2.142	353.9
	2-Chloro-4-biphenylol	−7.693	0.384	−0.077	−0.188	1.065	342.0
	4-Chloro-4'-biphenylol	−7.141	0.012	−0.082	−0.179	3.506	383.3
	4-Hydroxybiphenyl	−7.487	0.550	−0.069	−0.180	1.777	322.3
	3-Phenylphenol	−7.556	0.628	−0.066	−0.180	1.772	325.9
	2,4'-Dichlorobiphenyl	−8.203	0.275	−0.083	−0.203	3.964	404.9

Table 1. Cont.

Category	Compound	HOMO (eV)	LUMO (eV)	μ^+ (a.u.)	μ^- (a.u.)	Dipole Moment (D)	APSA (\AA^2)	
Phenol	4-n-Octylphenol	-7.270	0.860	-0.061	-0.173	1.811	452.9	
	2-sec-Butylphenol	-7.369	0.894	-0.061	-0.174	1.513	320.0	
	4-sec-Butylphenol	-7.285	0.824	-0.062	-0.173	1.837	305.3	
	4-tert-Butylphenol	-7.292	0.847	-0.062	-0.173	1.850	293.8	
	4-Chloro-3-methylphenol	-7.454	0.657	-0.069	-0.180	3.190	249.0	
	4-Phenethylphenol	-7.288	0.831	-0.060	-0.173	1.768	389.2	
	3-Ethylphenol	-7.446	0.897	-0.061	-0.176	1.434	264.1	
	a,a-Dimethyl-b-ethylallenolicacid	-6.976	-0.207	-0.087	-0.176	1.487	370.8	
	4-Chloro-2-methylphenol	-7.439	0.638	-0.069	-0.180	3.670	258.8	
	2-Cholor-4-methylphenol	-7.476	0.576	-0.071	-0.181	4.155	258.3	
	Heptylp-hydroxybenzoate	-7.793	-0.079	-0.092	-0.195	4.521	447.5	
	2-Ethylhexyl-4-hydroxybenzoate	-7.793	-0.082	-0.092	-0.195	1.723	430.2	
	Benzyl4-hydroxybenzoate	-7.805	-0.113	-0.093	-0.196	1.833	366.5	
	DDT	o,p'-DDT	-8.134	-0.246	-0.099	-0.211	3.389	491.9
		2,4-Dihydroxybenzophenone	-7.558	-0.726	-0.107	-0.195	7.199	291.3
Phenolphthalein		-7.497	-0.520	-0.101	-0.192	9.135	363.6	
Phenol red		-7.631	-0.332	-0.097	-0.194	9.753	355.4	
4,4'-Sulfonyldiphenol		-7.766	-0.261	-0.082	-0.196	8.662	258.9	
4,4'-Dihydroxy-benzophenone		-7.708	-0.662	-0.108	-0.204	4.990	272.9	
2,2'-Methylenebis(4-chlorophenol)		-7.468	0.165	-0.082	-0.189	3.955	374.5	
Bis(4-hydroxyphenyl)methane		-7.145	0.758	-0.061	-0.174	1.332	303.5	
Monohydroxy methoxychlor		-7.335	0.049	-0.087	-0.188	2.570	443.8	
Monohydroxy methoxychlor olefin		-7.057	-0.150	-0.067	-0.173	4.977	429.8	
p-Cumylphenol		-7.259	0.712	-0.061	-0.173	1.612	376.0	

Table 2. The binding affinity (log RBA) of the endocrine disrupting chemicals (experimental, predicted and residual values).

Category	Compound	Expt.	Pred.	Δ (Expt. - Pred.)	
PCB	2,3,4,5-Tetrachloro-4'-biphenylol	-0.64	-0.60	-0.04	
	2,5-Dichloro-4'-biphenylol	-1.44	-1.59	0.15	
	2-Chloro-4-biphenylol	-2.77	-2.74	-0.03	
	4-Chloro-4'-biphenylol	-2.18	-1.33	-0.85	
	4-Hydroxybiphenyl	-3.04	-2.81	-0.23	
	3-Phenylphenol	-3.44	-3.03	-0.41	
	2,4'-Dichlorobiphenyl	-3.61	-3.65	0.04	
Phenol	4-n-Octylphenol	-2.31	-2.20	-0.11	
	2-sec-Butylphenol	-3.54	-3.26	-0.28	
	4-sec-Butylphenol	-3.37	-3.15	-0.22	
	4-tert-Butylphenol	-3.61	-3.27	-0.34	
	4-Chloro-3-methylphenol	-3.38	-3.83	0.45	
	4-Phenethylphenol	-2.69	-2.57	-0.12	
	3-Ethylphenol	-3.87	-3.75	-0.12	
	a,a-Dimethyl-b-ethylallenolicacid	-0.02	-0.34	0.32	
	4-Chloro-2-methylphenol	-3.67	-3.79	0.12	
	2-Cholor-4-methylphenol	-3.66	-3.80	0.14	
	Heptylp-hydroxybenzoate	-2.09	-2.05	-0.04	
	2-Ethylhexyl-4-hydroxybenzoate	-1.74	-1.70	-0.04	
	Benzyl4-hydroxybenzoate	-2.54	-2.11	-0.43	
	DDT	o,p'-DDT	-2.85	-2.31	-0.54
		2,4-Dihydroxybenzophenone	-2.61	-1.95	-0.66
Phenolphthalein		-1.87	-2.07	0.20	
Phenol red		-3.25	-2.78	-0.47	
4,4'-Sulfonyldiphenol		-3.07	-3.62	0.55	
4,4'-Dihydroxy-benzophenone		-2.46	-2.48	0.02	
2,2'-Methylenebis(4-chlorophenol)		-2.45	-2.60	0.15	
Bis(4-hydroxyphenyl)methane		-3.02	-2.96	-0.06	
Monohydroxy methoxychlor		-0.89	-1.51	0.62	
Monohydroxy methoxychlor olefin		-0.63	-0.45	-0.18	
p-Cumylphenol		-2.30	-2.35	0.05	

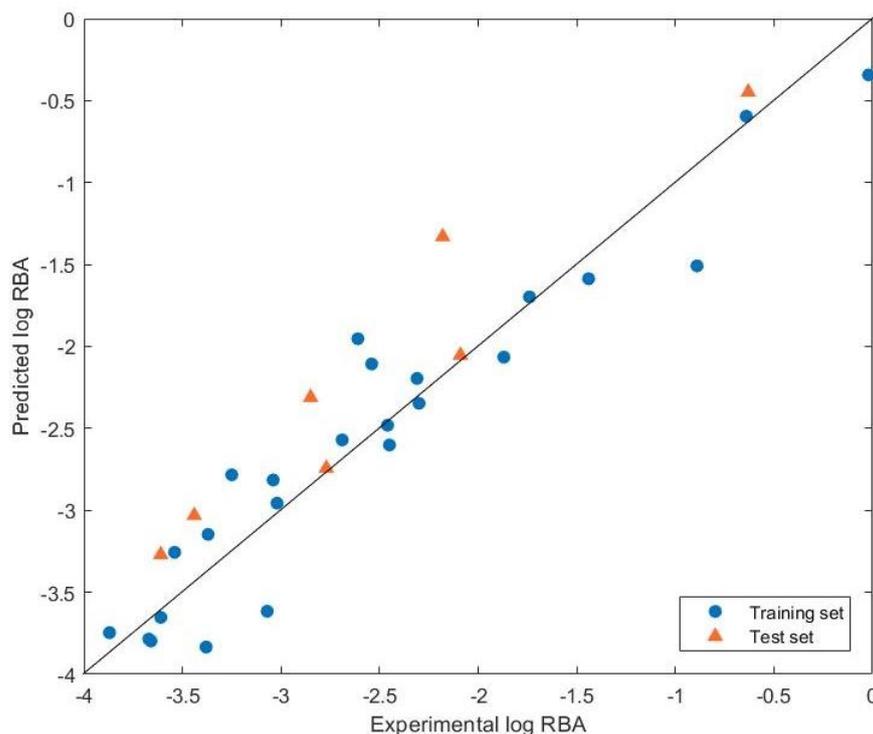


Figure 1. The scatter plot of experimental and quantum chemical GA-MLR predicted log RBA.

$$\text{Log RBA} = 9.8648 (\pm 2.1265) - 2.4609 (\pm 0.2291) \text{ LUMO} + 73.3633 (\pm 11.1238) \mu^- - 0.1667 (\pm 0.039) \text{ Dipole Moment} + 0.0067 (\pm 0.0012) \text{ APSA} \quad (1)$$

Internal validation metrics:

$$\begin{aligned} R^2 &= 0.9101 \\ R^2_{\text{adj}} &= 0.8911 \\ \text{Standard Error of Estimation (SEE)} &= 0.3386 \\ Q^2_{\text{LOO}} &= 0.8484 \\ \text{SDEP}_{\text{LOO}} &= 0.3912 \\ \text{Scaled } r^2_m \text{ (LOO)} &= 0.7956 \\ \text{Scaled } \Delta r^2_m \text{ (LOO)} &= 0.0637 \\ \text{Mean Absolute Error (MAE)} &= 0.2989 \end{aligned}$$

External validation metrics using a test set:

$$\begin{aligned} Q^2_{\text{F1}} &= 0.7820 \\ Q^2_{\text{F2}} &= 0.7813 \\ \text{Scaled } r^2_m \text{ (Test)} &= 0.7125 \\ \text{Scaled } \Delta r^2_m \text{ (Test)} &= 0.0998 \\ \text{CCC (Test)} &= 0.8968 \\ \text{Mean Absolute Error (MAE, Test)} &= 0.3403 \end{aligned}$$

In addition, the four descriptors in Equation (1) are tested with the VIF and showed that all values are less than 3.152 (Table 3), confirming the absence of multicollinearity in the modeling results. The t-values of the descriptors represent the individual contribution of one descriptor relative to other descriptors in the model; the +/− signs indicate whether the descriptor contributes positively or negatively to molecular potency. Furthermore, the larger the absolute t-value, the larger the contribution of the descriptor to the molecular potency. As can be seen in Table 3, the contribution of the descriptors is LUMO > μ^- > APSA > dipole moment, implying that electrophilicity plays an important role in the interaction of endocrine disruptors with estrogen receptor α . The negative contribution of LUMO indicates that the value of log RBA increases as the electron-accepting ability of EDCs decreases. From the LUMO values shown in Table 1, it can be seen that DDT

has the largest negative value and is a soft electrophile, which mainly reacts with proteins belonging to soft nucleophiles, resulting in a higher log RBA value. Soft electrophilic descriptors have been shown to be positively correlated with binding affinity [32,33]. The positive contribution of μ^- suggests that lowering the charge-donating chemical potential of EDCs reduces log RBA. μ^- is expressed as the chemical potential that controls the charge donation process, however, phenol has the smallest value due to fewer benzene rings (nucleophilic groups), resulting in fewer electrons flowing to ER α .

Table 3. The variance inflation factors (VIFs) and t-values of four descriptors in the quantum chemical GA-MLR model (Equation (1)).

Descriptor	VIF	t-Value
LUMO	3.152	−10.740
μ^-	2.561	6.595
Dipole Moment	2.076	−4.273
APSA	1.162	5.496

The positive contribution of APSA indicates that log RBA increases with increasing hydrophobicity. This is consistent with the positive correlation between aromatic chemicals and hydrophobicity proposed in the previous literature [32,33]. It can be found that the APSA value increases with the number of halides (chlorides) and benzene rings, making the DDT value the largest and the phenol value the smallest. The negative contribution of the dipole moment indicates that the log RBA decreases with increasing polarity of the endocrine disrupting chemical, since as the molecular dipole moment increases, so does the hydrophilicity of EDCs [34].

3.2. Quantum Chemical Cluster Model

The quantum chemical cluster method with the combined computational technique of AM1/DFT is used to clarify the binding interactions of ER α with EDCs. It is well known that the semiempirical AM1 method for the geometry optimization of molecules is faster than the DFT method, and there is an outstanding relationship between the optimized geometries of AM1 and DFT [35–37]. In addition, the semiempirical AM1 method confirms that stable structures appear in ligand orientations as seen by X-ray crystallography [38]. Therefore, the AM1 Hamiltonian is used in this study to obtain the equilibrium structures of nine compounds in complex with the ER α receptor (pdb code: 3ERT). However, although semiempirical methods can be used to obtain acceptable equilibrium geometries, they are not reliable enough to accurately calculate electronic properties [39]. Therefore, this study uses M06-2X/6-31G(d,p)/SCRF calculations to obtain binding energies. Nine compounds with different binding energies (ΔE) are selected from three categories of PCB, phenol, and DDT for the following discussion (Table 4). The nine compounds are 2-Chloro-4-biphenylol, 4-Chloro-4'-biphenylol, 2,4'-Dichlorobiphenyl, 4-Phenethylphenol, 4-Chloro-2-methylphenol, 4-Chloro-3-methylphenol, 2,4-Dihydroxybenzophenone, p-Cumylphenol, and 2,2'-Methylenebis(4-chlorophenol).

2-Chloro-4-biphenylol (−16.93 kcal/mol) and 4-Chloro-4'-biphenylol (−24.33 kcal/mol) are monohydroxylated PCBs (Figure 2a,b), the hydroxyl groups of which are linked to the side chains of Leu387, Glu353, and Arg394. 2,4'-Dichlorobiphenyl (−9.84 kcal/mol) is an unhydroxylated PCB (c), which does not form any hydrogen bonds with amino acid residues. Compared to monohydroxylated PCBs, unhydroxylated PCBs poorly bind to ER α receptors [40]. Among them, due to the key amino acid residue Phe404 near 4-Chloro-4'-biphenylol, it provides hydrophobic group binding with the benzene ring of the ligand, forming a π - π interaction to generate a larger binding energy.

Table 4. The conceptual density functional theory descriptors and binding energies of nine endocrine disrupting chemicals.

Category	Compound	$\rho^+_{\text{Max(H)}} \text{ (a.u.)}$	$\rho^+_{\text{Max}} \text{ (a.u.)}$	f^+_{Max}	Sites for Nucleophilic Attack	$s^+_{\text{Max}} \text{ (a.u.)}$	f^-_{Max}	Site for Electrophilic Attack	$s^-_{\text{Max}} \text{ (a.u.)}$	$\Delta E \text{ (kcal/mol)}$
PCB	2-Chloro-4-biphenylol	0.369	0.369	0.129	C11	0.582	0.125	O1	0.567	-16.93
	4-Chloro-4'-biphenylol	0.366	0.366	0.107	C11	0.553	0.095	O2	0.492	-24.33
	2,4'-Dichlorobiphenyl	0.167	0.167	0.097	C5	0.407	0.241	C12	1.010	-9.84
Phenol	4-Chloro-3-methylphenol	0.367	0.367	0.115	C15	0.519	0.191	O1	0.860	-16.94
	4-Phenethylphenol	0.363	0.363	0.118	C5	0.522	0.124	C11	0.548	-22.19
	4-Chloro-2-methylphenol	0.366	0.366	0.111	C4	0.504	0.185	C11	0.837	-10.19
DDT	2,4-Dihydroxybenzophenone	0.371	0.371	0.151	O2	0.860	0.132	C10	0.748	-24.37
	p-Cumylphenol	0.364	0.364	0.070	C14	0.312	0.120	O1	0.537	-26.62
	2,2'-Methylenebis(4-chlorophenol)	0.371	0.371	0.099	C13	0.461	0.118	C12	0.548	-21.21

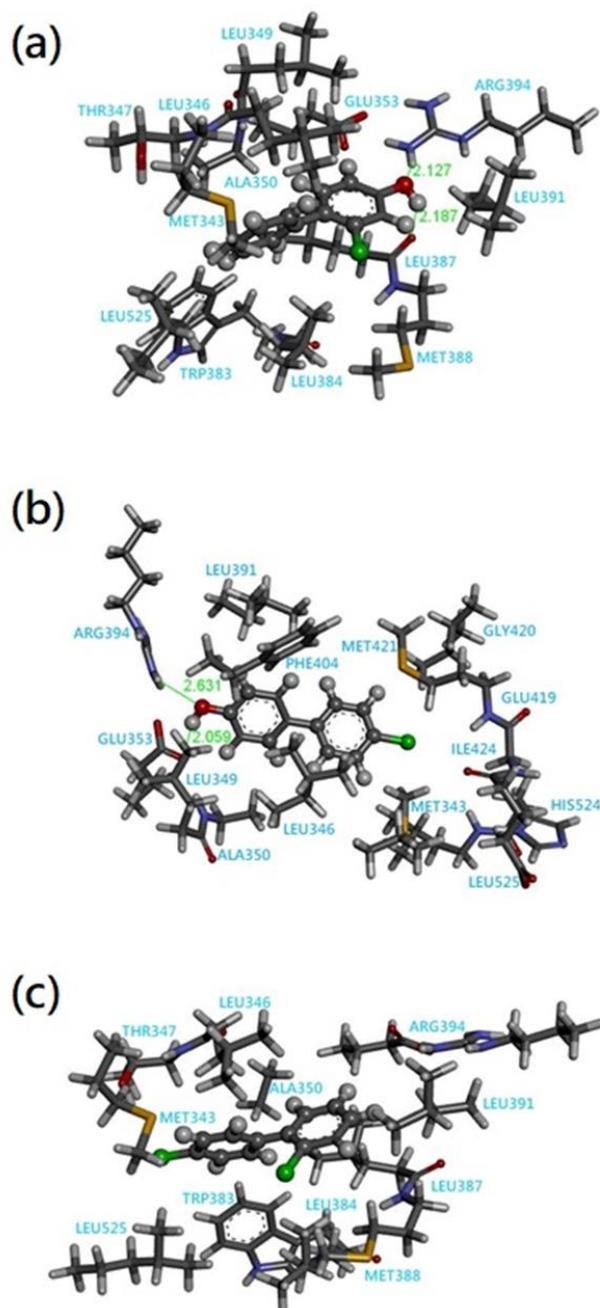
**Figure 2.** The quantum chemical cluster models for the ER α receptor (PDB 3ERT) with PCBs: (a) 2-Chloro-4-biphenylol, (b) 4-Chloro-4'-biphenylol, and (c) 2,4'-Dichlorobiphenyl.

Figure 3a shows 4-Chloro-3-methylphenol (-22.19 kcal/mol), whose hydroxyl groups form hydrogen bonds with Glu353 and Arg394 with bond distances of 1.968 Å and 2.183 Å. Figure 3b shows 4-Phenethylphenol (-10.19 kcal/mol), whose hydroxyl groups form hydrogen bonds with Leu387 and Lys449 with bond distances of 2.117 Å and 2.643 Å. Figure 3c shows 4-Chloro-2-methylphenol (-16.94 kcal/mol), whose hydroxyl group forms a hydrogen bond with Glu353 with a bond distance of 1.974 Å. It can be found that not only the length of the hydrogen bond affects the strength of the binding energy, but also the amino acid residues (Glu353 and Arg394) that form hydrogen bonds with the ligand are also important. The biological activity of ER α depends on the specific binding of ligands to the ligand-binding cleft (LBC) and activation function 2 (AF-2) [41]. Furthermore, due to the presence of Phe404 around 4-Chloro-3-methylphenol and 4-Chloro-2-methylphenol, a π - π interaction is formed, generating a larger binding energy. Among them, the bond distance between 4-Chloro-3-methylphenol and Phe404 is significantly shorter than that of 4-Chloro-2-methylphenol. Therefore, 4-Chloro-3-methylphenol forms a strong π - π interaction to result in a larger binding energy.

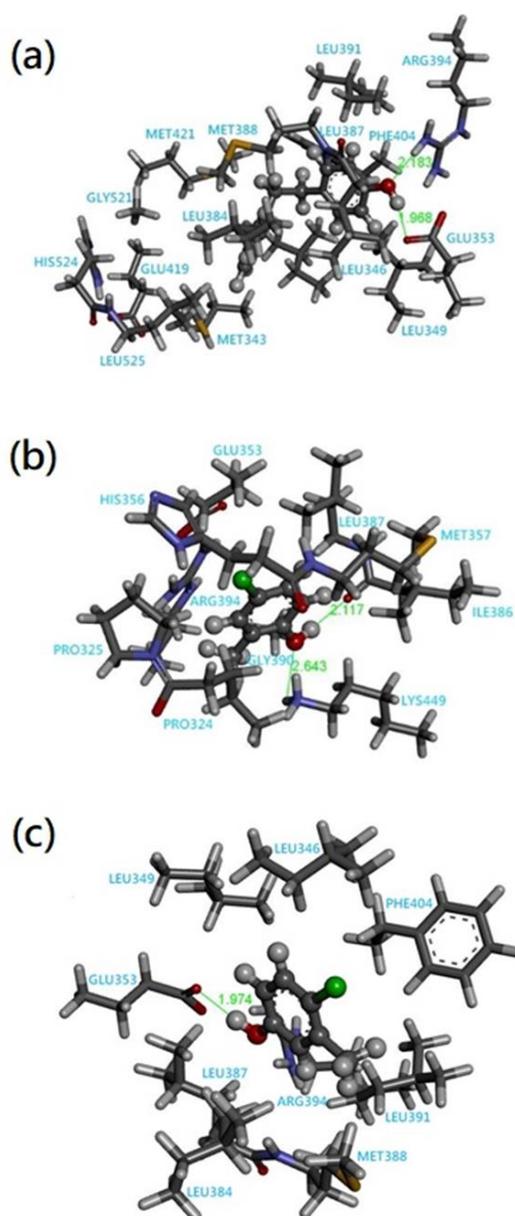


Figure 3. The quantum chemical cluster models for the ER α receptor (PDB 3ERT) with phenols: (a) 4-Chloro-3-methylphenol, (b) 4-Phenethylphenol, and (c) 4-Chloro-2-methylphenol.

The results of structural analysis show that hydrogen bondings play a significant role in ER α -ligand interactions, and the key residues are Glu353, Arg394, His524, Leu525, and Leu387. Glu353 and Arg394, the regions near these two amino acids, are hydrophilic [44], being easily combined with ligands to form hydrogen bonds. In particular, Glu353 can form a strong hydrogen bond with a phenolic group [3], and a protein–ligand complex structure can be observed that forms hydrogen bonds with Glu353; its binding energy is large. Next, the side chains of His524 and Leu525 form polar and/or non-polar interactions with the benzene ring in the ligand, making them form part of the binding site of helix 12 or the previous ring [45] that stabilizes the conformation of the complex. Moreover, residues of the binding pocket form a hydrophobic network, and the formation of hydrophobic interactions with ligands is also important in ER α -ligand interactions. From Figures 2–4, not only it can be observed that the key residues are Leu346, Leu387, Leu391, Phe404 and Ala350, of which leucine is the most abundant, it is also found that the ligand will be closer to these residues to provide van der Waals interactions. The key residues mentioned in the previous literature to use MM-GBSA to identify ER α -ligand interactions are Leu346, Leu387, Leu391, Phe404, and Ala350 [44], which is consistent with the results of this study. Among the three categories of PCB, phenol, and DDT, DDT has the largest binding energy, which is not only related to hydrogen bonding but also to the hydrophobic interaction. This means that efficiently filling the LBC cavity with larger molecules can promote binding between hydrophobic amino acid residues and ligands [41].

3.3. Conceptual Density Functional Theory

Pearson's hard-soft acid-base (HSAB) principle and conceptual density functional theory have provided important insights into the nature of chemical reactions and the stability of molecular systems. Global HSAB reactivity descriptors are often used to indicate the stability of compounds, while local HSAB reactivity descriptors are often used to discuss site selectivity issues [46–49]. The HSAB reactivity descriptors can also help to clarify the mechanism of target–toxicant interactions and have been applied to explain chemically induced toxicity [50].

The maximum partial charge of the hydrogen atom ($\rho^+_{\max(\text{H})}$) and the maximum partial charge (ρ^+_{\max}) can be considered as the indicator of electrostatic interactions. The results shown in Table 4 demonstrated the order of binding energies is DDT > phenol ~ PCB. In particular, 2,2'-Methylenebis(4-chlorophenol) has the largest $\rho^+_{\max(\text{H})}$. On the contrary, the $\rho^+_{\max(\text{H})}$ of 2,4'-Dichlorobiphenyl is the smallest, which leads to the weak binding interaction. It can be seen in Figure 4c that the chlorine atom of the ligand is inclined towards Phe404, and the electrophilic region at the top of the chlorine atom tends to the π -system of the aromatic groups in the side chains of tyrosine, phenylalanine, histidine, and tryptophan. It not only forms interactions between receptor and ligand, but also for the regulation and stabilization of intramolecular short peptides and proteins [7]. Among the nine complexes, p-Cumylphenol has the highest binding energy, and its f^+_{\max} , s^+_{\max} , f^-_{\max} , and s^-_{\max} are the smallest. Furthermore, it not only has large $\rho^+_{\max(\text{H})}$ and ρ^+_{\max} , but also has two hydrogen bonds. Therefore, it is improved for hard–hard (charge-controlled) interaction.

Fukui functions f^+ and f^- are reactivity indices which can govern nucleophilic and electrophilic attacks, respectively. The larger value of Fukui function for a specific site supports the reactivity of that site [51]. Besides, the local softness should be interpreted as the concentration of the corresponding global softness, related to the Fukui function, which has been shown to be useful for ligand docking, active site detection, and protein folding prediction [52,53]. Pearson's HSAB principle states that soft acids or bases tend to react with soft bases or acids, whereas hard acids or bases preferentially react with hard bases or acids. Soft–soft interactions are essentially covalent front-controlled, while hard–hard interactions are essentially ionic charge-controlled [52].

From the local reactivity descriptors and binding energies summarized in Table 4, 2,4-Dihydroxybenzophenone has the largest f^+_{\max} and s^+_{\max} , which can be considered as soft–soft (frontier-controlled interaction), indicating that there are strong nucleophilicity.

Due to its smaller f_{\max}^+ and s_{\max}^+ values for p-Cumylphenol, it tends to have hard–hard (charge-controlled) interactions, and has two hydrogen bonds with ER α . Combined with its large nonpolar surface area (APSA) and strong hydrophobic interactions, the binding affinity of p-Cumylphenol is maximized. 4-Chloro-4'-biphenylol has the smallest f_{\max}^- and s_{\max}^- , and it has two hydrogen bonds with ER α , thus proving to be a hard–hard (charge-controlled) interaction. The lowest binding energy is 2,4'-Dichlorobiphenyl due to the non-H-bond with ER α and the lowest value of quantum chemical descriptors. Furthermore, as the number of hydroxyl groups increases, the values of f_{\max}^- and s_{\max}^- will decrease.

4. Conclusions

The quantum chemical GA-MLR and cluster model approaches are used to investigate the binding interactions of PCB, phenol, and DDT to ER α . The results of hydrogen bonding analysis reveal that the key residues are Glu353, Arg394, His524, Leu525, and Leu387. In particular, the hydrogen bonding with the amino acid Glu353 has a stronger binding tendency. The mechanism of binding interaction of endocrine disrupting chemicals is also proved from chemical reactivity theory according to the conceptual density functional theory descriptors. The binding mechanism of the electrostatic interaction is superior to that of the electrophilic interaction.

Author Contributions: Conceptualization, S.-C.C., H.-C.H. and C.-M.C.; methodology, S.-C.C., H.-C.H. and C.-M.C.; software, S.-C.C.; validation, C.-M.C.; formal analysis, S.-C.C., H.-C.H. and C.-M.C.; investigation, S.-C.C.; resources, C.-M.C.; data curation, S.-C.C.; writing—original draft preparation, S.-C.C.; writing—review and editing, H.-C.H. and C.-M.C.; visualization, S.-C.C., H.-C.H. and C.-M.C.; supervision, C.-M.C.; project administration, C.-M.C.; funding acquisition, C.-M.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Science Council of Taiwan, Republic of China and grant number [MOST 111-2321-B-005-004].

Data Availability Statement: Not applicable.

Acknowledgments: The authors thank the National Science Council of Taiwan, Republic of China, MOST 111-2321-B-005-004 for providing financial support. Computer time was provided by the National Center for High-Performance Computing.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. EFSA Scientific Committee. Scientific Opinion on the hazard assessment of endocrine disruptors: Scientific criteria for identification of endocrine disruptors and appropriateness of existing test methods for assessing effects mediated by these substances on human health and the environment. *EFSA J.* **2013**, *11*, 3132. [[CrossRef](#)]
2. Sellami, A.; Montes, M.; Lagarde, N. Predicting Potential Endocrine Disrupting Chemicals Binding to Estrogen Receptor α (ER α) Using a Pipeline Combining Structure-Based and Ligand-Based in Silico Methods. *Int. J. Mol. Sci.* **2021**, *22*, 2846. [[CrossRef](#)]
3. Lee, S.; Barron, M.G. Structure-Based Understanding of Binding Affinity and Mode of Estrogen Receptor alpha Agonists and Antagonists. *PLoS ONE* **2017**, *12*, e0169607. [[CrossRef](#)]
4. Toporova, L.; Balaguer, P. Nuclear receptors are the major targets of endocrine disrupting chemicals. *Mol. Cell. Endocrinol.* **2020**, *502*, 110665. [[CrossRef](#)]
5. Uzzaman, M.; Mahmud, T. Structural modification of aspirin to design a new potential cyclooxygenase (COX-2) inhibitors. *Silico Pharmacol.* **2020**, *8*, 1. [[CrossRef](#)]
6. Uzzaman, M.; Hasan, M.K.; Mahmud, S.; Yousuf, A.; Islam, S.; Uddin, M.N.; Barua, A. Physicochemical, spectral, molecular docking and ADMET studies of Bisphenol analogues; A computational approach. *Inform. Med. Unlocked* **2021**, *25*, 100706. [[CrossRef](#)]
7. Matsushima, A. A Novel Action of Endocrine-Disrupting Chemicals on Wildlife; DDT and Its Derivatives Have Remained in the Environment. *Int. J. Mol. Sci.* **2018**, *19*, 1377. [[CrossRef](#)]
8. Zhuang, S.; Zhang, J.; Wen, Y.; Zhang, C.; Liu, W. Distinct mechanisms of endocrine disruption of DDT-related pesticides toward estrogen receptor α and estrogen-related receptor γ . *Environ. Toxicol. Chem.* **2012**, *31*, 2597–2605. [[CrossRef](#)]
9. Landeros-Martinez, L.L.; Glossman-Mitnik, D.; Flores-Holguin, N. Interaction of Tamoxifen Analogs With the Pocket Site of Some Hormone Receptors. A Molecular Docking and Density Functional Theory Study. *Front. Chem.* **2018**, *6*, 293. [[CrossRef](#)]

10. Żwierzeło, W.; Maruszczyńska, A.; Skórka-Majewicz, M.; Goschorska, M.; Baranowska-Bosiacka, I.; Dec, K.; Styburski, D.; Nowakowska, A.; Gutowska, I. The influence of polyphenols on metabolic disorders caused by compounds released from plastics—Review. *Chemosphere* **2020**, *240*, 124901. [[CrossRef](#)]
11. Manzetti, S.; van der Spoel, E.R.; van der Spoel, D. Chemical Properties, Environmental Fate, and Degradation of Seven Classes of Pollutants. *Chem. Res. Toxicol.* **2014**, *27*, 713–737. [[CrossRef](#)] [[PubMed](#)]
12. Ding, D.; Xu, L.; Fang, H.; Hong, H.; Perkins, R.; Harris, S.; Bearden, E.D.; Shi, L.; Tong, W. The EDKB: An established knowledge base for endocrine disrupting chemicals. *BMC Bioinform.* **2010**, *11* (Suppl. 6), S5. [[CrossRef](#)] [[PubMed](#)]
13. Akkoç, S.; Tüzün, B.; Özalp, A.; Kökbudak, Z. Investigation of structural, electronical and in vitro cytotoxic activity properties of some heterocyclic compounds. *J. Mol. Struct.* **2021**, *1246*, 131127. [[CrossRef](#)]
14. Tsuneda, T.; Song, J.W.; Suzuki, S.; Hirao, K. On Koopmans' theorem in density functional theory. *J. Chem. Phys.* **2010**, *133*, 174101. [[CrossRef](#)] [[PubMed](#)]
15. Perdew, J.P.; Levy, M. Comment on “Significance of the highest occupied Kohn-Sham eigenvalue”. *Phys. Rev. B* **1997**, *56*, 16021. [[CrossRef](#)]
16. Casida, M.E. Correlated optimized effective-potential treatment of the derivative discontinuity and of the highest occupied Kohn-Sham eigenvalue: A Janak-type theorem for the optimized effective-potential model. *Phys. Rev. B* **1999**, *59*, 4694–4698. [[CrossRef](#)]
17. Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **2010**, *132*, 154104. [[CrossRef](#)] [[PubMed](#)]
18. Demircioğlu, Z.; Büyükgüngör, O.; Ersanlı, C.C. Chemical Activity Studies with Density Functional Theory. *SETSCI Conf. Index. Syst.* **2018**, *3*, 695–697.
19. Pathak, S.K.; Srivastava, R.; Sachan, A.K.; Prasad, O.; Sinha, L.; Asiri, A.M.; Karabacak, M. Experimental (FT-IR, FT-Raman, UV and NMR) and quantum chemical studies on molecular structure, spectroscopic analysis, NLO, NBO and reactivity descriptors of 3,5-Difluoroaniline. *Spectrochim. Acta A Mol. Biomol. Spectrosc.* **2015**, *135*, 283–295. [[CrossRef](#)]
20. Aiken, L.S.; West, S.G.; Pitts, S.C. Multiple Linear Regression. In *Handbook of Psychology*; Weiner, I.B., Ed.; Wiley: Hoboken, NJ, USA, 2003; pp. 481–507. [[CrossRef](#)]
21. Ambure, P.; Gajewicz-Skretna, A.; Cordeiro, M.; Roy, K. New Workflow for QSAR Model Development from Small Data Sets: Small Dataset Curator and Small Dataset Modeler. Integration of Data Curation, Exhaustive Double Cross-Validation, and a Set of Optimal Model Selection Techniques. *J. Chem. Inf. Model.* **2019**, *59*, 4070–4076. [[CrossRef](#)] [[PubMed](#)]
22. Golbraikh, A.; Tropsha, A. Beware of q²! *J. Mol. Graph. Model.* **2002**, *20*, 269–272. [[CrossRef](#)] [[PubMed](#)]
23. Tropsha, A. Best Practices for QSAR Model Development, Validation, and Exploitation. *Mol. Inform.* **2010**, *29*, 476–488. [[CrossRef](#)]
24. Alam, S.; Khan, F. QSAR, docking, ADMET, and system pharmacology studies on tormentic acid derivatives for anticancer activity. *J. Biomol. Struct. Dyn.* **2018**, *36*, 2373–2390. [[CrossRef](#)] [[PubMed](#)]
25. Roy, K.; Ambure, P.; Kar, S.; Ojha, P.K. Is it possible to improve the quality of predictions from an “intelligent” use of multiple QSAR/QSPR/QSTR models? *J. Chemom.* **2018**, *32*, e2992. [[CrossRef](#)]
26. Olasupo, S.B.; Uzairu, A.; Adamu, G.S.; Uba, S. Computational Modeling and Pharmacokinetics/ADMET Study of Some Arylpiperazine Derivatives as Novel Antipsychotic Agents Targeting Depression. *Chem. Afr.* **2020**, *3*, 979–988. [[CrossRef](#)]
27. Jaworska, J.S.; Comber, M.; Auer, C.; Van Leeuwen, C.J. Summary of a workshop on regulatory acceptance of (Q)SARs for human health and environmental endpoints. *Environ. Health Perspect.* **2003**, *111*, 1358–1360. [[CrossRef](#)]
28. Trott, O.; Olson, A.J. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **2010**, *31*, 455–461. [[CrossRef](#)]
29. Stewart, J.J. Optimization of parameters for semiempirical methods VI: More modifications to the NDDO approximations and re-optimization of parameters. *J. Mol. Model.* **2013**, *19*, 1–32. [[CrossRef](#)]
30. Frisch, M.J.; Trucks, G.W.; Schlegel, H.B.; Scuseria, G.E.; Robb, M.A.; Cheeseman, J.R.; Scalmani, G.; Barone, V.; Petersson, G.A.; Nakatsuji, H.; et al. Gaussian 16 Rev. C.01. Gaussian, Inc.: Wallingford, CT, USA, 2016.
31. Pedretti, A.; Mazzolari, A.; Gervasoni, S.; Fumagalli, L.; Vistoli, G. The VEGA suite of programs: An versatile platform for cheminformatics and drug design projects. *Bioinformatics* **2021**, *37*, 1174–1175. [[CrossRef](#)]
32. Mekenyan, O.G.; Veith, G.D. Relationships between descriptors for hydrophobicity and soft electrophilicity in predicting toxicity. *SAR QSAR Environ. Res.* **1993**, *1*, 335–344. [[CrossRef](#)]
33. Veith, G.D.; Mekenyan, O.G. A QSAR Approach for Estimating the Aquatic Toxicity of Soft Electrophiles [QSAR for Soft Electrophiles]. *Quant. Struct.-Act. Relatsh.* **1993**, *12*, 349–356. [[CrossRef](#)]
34. Bary, G.; Ghani, L.; Jamil, M.I.; Arslan, M.; Ahmed, W.; Ahmad, A.; Sajid, M.; Ahmad, R.; Huang, D. Designing small organic non-fullerene acceptor molecules with diflorobenzene or quinoline core and dithiophene donor moiety through density functional theory. *Sci. Rep.* **2021**, *11*, 19683. [[CrossRef](#)] [[PubMed](#)]
35. Sun, Y.; Chen, D.; Liu, C. Evaluation of the effectiveness of AM1 geometry used in calculating O–H bond dissociation enthalpy. *J. Mol. Struct. THEOCHEM* **2002**, *618*, 181–189. [[CrossRef](#)]
36. Zheng, G.; Irle, S.; Morokuma, K. Performance of the DFTB method in comparison to DFT and semiempirical methods for geometries and energies of C₂₀–C₈₆ fullerene isomers. *Chem. Phys. Lett.* **2005**, *412*, 210–216. [[CrossRef](#)]
37. Derosa, P.A. A combined semiempirical-DFT study of oligomers within the finite-chain approximation, evolution from oligomers to polymers. *J. Comput. Chem.* **2009**, *30*, 1220–1228. [[CrossRef](#)] [[PubMed](#)]

38. Villar, R.; Gil, M.J.; García, J.I.; Martínez-Merino, V. Are AM1 ligand-protein binding enthalpies good enough for use in the rational design of new drugs? *J. Comput. Chem.* **2005**, *26*, 1347–1358. [[CrossRef](#)]
39. Derosa, P.; Koraboina, K.; Sanders, M. A combined model to study conductive properties of polymers with atomic resolution. In Proceedings of the AIChE Annual Meeting, Conference Proceedings, San Francisco, CA, USA, 12–17 November 2006.
40. Celik, L.; Lund, J.D.D.; Schiøtt, B. Exploring Interactions of Endocrine-Disrupting Compounds with Different Conformations of the Human Estrogen Receptor α Ligand Binding Domain: A Molecular Docking Study. *Chem. Res. Toxicol.* **2008**, *21*, 2195–2206. [[CrossRef](#)]
41. Ye, H.; Dudley, S.Z.; Shaw, I.C. Intimate estrogen receptor- α /ligand relationships signal biological activity. *Toxicology* **2018**, *408*, 80–87. [[CrossRef](#)] [[PubMed](#)]
42. Ekena, K.; Weis, K.E.; Katzenellenbogen, J.A.; Katzenellenbogen, B.S. Identification of amino acids in the hormone binding domain of the human estrogen receptor important in estrogen binding. *J. Biol. Chem.* **1996**, *271*, 20053–20059. [[CrossRef](#)] [[PubMed](#)]
43. Ekena, K.; Weis, K.E.; Katzenellenbogen, J.A.; Katzenellenbogen, B.S. Different residues of the human estrogen receptor are involved in the recognition of structurally diverse estrogens and antiestrogens. *J. Biol. Chem.* **1997**, *272*, 5069–5075. [[CrossRef](#)] [[PubMed](#)]
44. Xue, Q.; Liu, X.; Liu, X.-C.; Pan, W.-X.; Fu, J.-J.; Zhang, A.-Q. The Effect of Structural Diversity on Ligand Specificity and Resulting Signaling Differences of Estrogen Receptor α . *Chem. Res. Toxicol.* **2019**, *32*, 1002–1013. [[CrossRef](#)] [[PubMed](#)]
45. 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. [[CrossRef](#)]
46. Pearson, R.G.; Songstad, J. Application of the Principle of Hard and Soft Acids and Bases to Organic Chemistry. *J. Am. Chem. Soc.* **1967**, *89*, 1827–1836. [[CrossRef](#)]
47. Chakraborty, D.; Chattaraj, P.K. Conceptual density functional theory based electronic structure principles. *Chem. Sci.* **2021**, *12*, 6264–6279. [[CrossRef](#)] [[PubMed](#)]
48. Geerlings, P.; Chamorro, E.; Chattaraj, P.K.; De Proft, F.; Gázquez, J.L.; Liu, S.; Morell, C.; Toro-Labbé, A.; Vela, A.; Ayers, P. Conceptual density functional theory: Status, prospects, issues. *Theor. Chem. Acc.* **2020**, *139*. [[CrossRef](#)]
49. Geerlings, P.; Proft, F.D.; Langenaeker, W. Conceptual Density Functional Theory. *Chem. Rev.* **2003**, *103*, 1793–1873. [[CrossRef](#)]
50. Lopachin, R.M.; Gavin, T.; Decaprio, A.; Barber, D.S. Application of the Hard and Soft, Acids and Bases (HSAB) theory to toxicant—Target interactions. *Chem. Res. Toxicol.* **2012**, *25*, 239–251. [[CrossRef](#)]
51. Rokhina, E.V.; Suri, R.P. Application of density functional theory (DFT) to study the properties and degradation of natural estrogen hormones with chemical oxidizers. *Sci. Total. Environ.* **2012**, *417–418*, 280–290. [[CrossRef](#)]
52. Torrent-Sucarrat, M.; De Proft, F.; Ayers, P.W.; Geerlings, P. On the applicability of local softness and hardness. *Phys. Chem. Chem. Phys.* **2010**, *12*, 1072–1080. [[CrossRef](#)]
53. Faver, J.; Kenneth, M.; Merz, J. Utility of the Hard/Soft Acid-Base Principle via the Fukui Function in Biological Systems. *J. Chem. Theory Comput.* **2010**, *6*, 548–559. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.