

Article

Identification of Potent LXR β -Selective Agonists without LXR α Activation by In Silico Approaches

Meimei Chen ^{1,2,*}, Fafu Yang ^{2,*}, Jie Kang ¹, Huijuan Gan ¹, Xuemei Yang ¹, Xinmei Lai ¹ and Yuxing Gao ³

¹ College of Traditional Chinese Medicine, Fujian University of Traditional Chinese Medicine, Fuzhou 350122, China; meinuohua2007@163.com (J.K.); ghjzyz@163.com (H.G.); mei_tcm@163.com (X.Y.); keerer1990@163.com (X.L.)

² College of Chemistry and Materials Science, Fujian Normal University, Fuzhou 350007, China

³ College of Chemistry and Chemical Engineering, Xiamen University, Xiamen 361005, China; gaoxingchem@xmu.edu.cn

* Correspondence: chenmeimei1984@163.com (M.C.); yangfafu@fjnu.edu.cn (F.Y.); Tel.: +86-0591-2286-1513 (M.C. & F.Y.)

Received: 19 May 2018; Accepted: 31 May 2018; Published: 4 June 2018



Abstract: Activating Liver X receptors (LXRs) represents a promising therapeutic option for dyslipidemia. However, activating LXR α may cause undesired lipogenic effects. Discovery of highly LXR β -selective agonists without LXR α activation were indispensable for dyslipidemia. In this study, in silico approaches were applied to develop highly potent LXR β -selective agonists based on a series of newly reported 3-(4-(2-propylphenoxy)butyl)imidazolidine-2,4-dione-based LXR α / β dual agonists. Initially, Kohonen and stepwise multiple linear regression SW-MLR were performed to construct models for LXR β agonists and LXR α agonists based on the structural characteristics of LXR α / β dual agonists, respectively. The obtained LXR β agonist model gave a good predictive ability ($R^2_{\text{train}} = 0.837$, $R^2_{\text{test}} = 0.843$, $Q^2_{\text{LOO}} = 0.715$), and the LXR α agonist model produced even better predictive ability ($R^2_{\text{train}} = 0.968$, $R^2_{\text{test}} = 0.914$, $Q^2_{\text{LOO}} = 0.895$). Also, the two QSAR models were independent and can well distinguish LXR β and LXR α activity. Then, compounds in the ZINC database met the lower limit of structural similarity of 0.7, compared to the 3-(4-(2-propylphenoxy)butyl)imidazolidine-2,4-dione scaffold subjected to our QSAR models, which resulted in the discovery of ZINC55084484 with an LXR β prediction value of pEC₅₀ equal to 7.343 and LXR α prediction value of pEC₅₀ equal to -1.901. Consequently, nine newly designed compounds were proposed as highly LXR β -selective agonists based on ZINC55084484 and molecular docking, of which LXR β prediction values almost exceeded 8 and LXR α prediction values were below 0.

Keywords: LXR β -selective agonists; QSAR modeling; molecular docking; Kohonen

1. Introduction

Numerous studies have demonstrated that high levels of plasma cholesterol induce dyslipidemia, atherosclerosis, and coronary heart diseases [1]. Liver X receptors (LXRs) are cholesterol sensors that protect cells from cholesterol overload [2]. Activating LXRs can stimulate reverse cholesterol transport from cells and inhibit its absorption and synthesis and promote HDL formation [3]. Currently, LXRs were identified as promising therapeutic targets for dyslipidemia, atherosclerosis, and cardiovascular diseases [4,5]. LXRs have two subtypes, including LXR α and LXR β . LXR α is mainly expressed in liver, adipose tissue, intestine, and macrophages, and LXR β is widely expressed in tissues. Both subtypes share approximately 78% homology in their ligand binding domains [6]. Activating LXR α results in undesired lipogenic effects such as increased hepatic lipogenesis, hypertriglyceridemia

and liver steatosis [7], but growing evidence suggests that LXR β -selective agonists can reduce these side effects [8,9]. Therefore, strategies to overcome the side effects related to LXR α activation in the treatment of dyslipidemia are to develop LXR β -selective agonists, avoiding hepatic lipogenesis and the development of steatosis. However, given the little differences in their ligand binding domains, it can be difficult to obtain LXR β -selective agonists without activation of LXR α .

To date, several computational approaches were also tried to predict LXR agonists. For instance, Susanne von Grafenstein, et al. identified novel LXR activators by structure-based modeling [10]. Also, Yali Li, et al. established QSAR classification models to distinguish between selective and non-selective LXR β agonists by use of classification methods [11]. He peng, et al. identified the privileged chemical fragments of LXR β agonists by application of a de novo substructure generation algorithm [12]. Veronika Temml, et al. discovered new LXR agonists by pharmacophore modeling and shape-based virtual screening [13]. However, these models did not make full use of the dual features of LXRs agonists, and did not judge whether these compounds were together with LXR α activity. Due to a high degree of similarity in the ligand binding domains of LXR β and LXR α , it is very important to consider the LXR α activation when designing LXR β agonists.

Therefore, the aim of this study was to apply a virtual screening workflow to simultaneously establish models for LXR β and LXR α agonists as a fast filter to find highly potent LXR β -selective agonists without LXR α activation, based on a series of new reported 3-(4-(2-propylphenoxy)butyl)imidazolidine-2,4-dione-based LXR α / β dual agonists, and also mined the structural features responsible for their selective activity of LXR α / β . Firstly, a Kohonen's self-organizing map and multiple linear regression combined with a stepwise technology were performed to construct models for LXR β agonists and LXR α agonists based on the same scaffolds of LXR α / β dual agonists, respectively. Secondly, the compounds in the ZINC database that fulfilled the requirement of a structural similarity of 0.7 compared to known reported LXRs agonists were subjected to the two QSAR models to screen for new LXR β -selective agonists. Then, to discover highly potent LXR β -selective agonists without LXR α activation, two QSAR models were further applied to design new compounds based on above screened compounds. Finally, molecular docking was applied to understand their binding interactions in the LXR β binding site.

2. Results and Discussion

2.1. Results of Dataset Division by Kohonen Map

After the descriptors reduction, totally 183 molecular descriptors were used as variables to build Kohonen maps (5 \times 5 neurons, 500 epochs). On the basis of the trained network, the compounds fell into different neuron of the Kohonen map (see Figure 1). Similar chemicals were within the same cell of the Kohonen map. The selection of the training set chemicals was performed by the minimal distance from the centroid of each cell in the Kohonen map [14]. The remaining objects, close to the training set chemicals, were used as the test set. As a result, there were 40 compounds in the training set and 12 compounds in the test set used for building models for LXR β agonists, additionally, 30 compounds in the training set and 11 compounds in the test set used for constructing models for LXR α agonists. Figure 1 showed the distribution of the training set and test set, marked by circle and triangle symbols, respectively.

2.2. MLR Model Results of LXR β Activity

After SW-MLR was performed, the best QSAR model for LXR β agonists was generated with nine molecular descriptors. The obtained QSAR model was shown as follows:

$$\text{pEC}_{50} = -0.777 \times \text{vsurf_IW2} + 0.007 \times \text{SMR_VSA6} - 1.236 \times \text{glob} + 4.560 \times \text{GCUT_SLOGP_2} - 60.185 \times \text{E_strain} - 0.189 \times \text{dipoleX} - 0.247 \times \text{AM1_LUMO} + 0.154 \times \text{vsurf_IW5} + 0.016 \times \text{vsurf_DD13} + 5.087$$

$N_{\text{train}} = 40$, $R^2_{\text{train}} = 0.837$, $F_{\text{train}} = 17.235 > F_{0.01}(9, 30) = 3.06$ (the cut off value of F distribution), $\text{RMSE}_{\text{train}} = 0.118$, $Q^2_{\text{LOO}} = 0.715$, $\text{RMSE}_{\text{LOO}} = 0.156$, $N_{\text{test}} = 12$, $R^2_{\text{test}} = 0.843$, $\text{RMSE}_{\text{test}} = 0.534$.

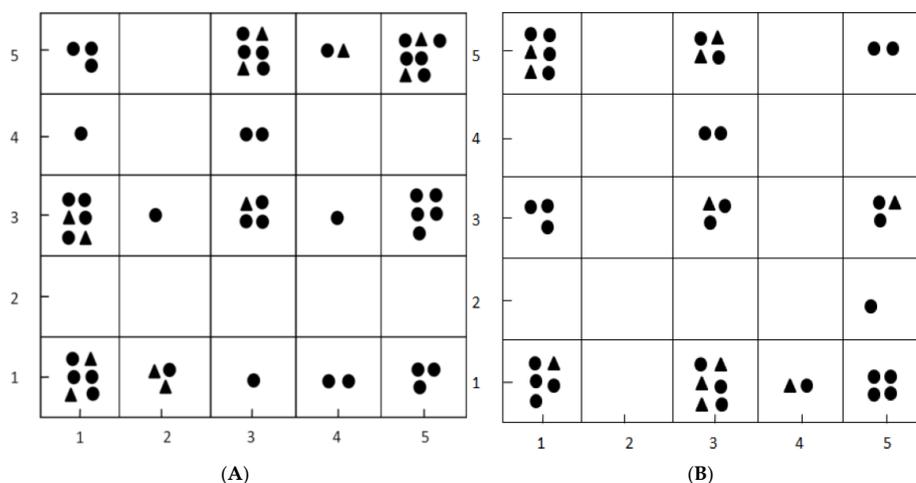


Figure 1. The Kohonen maps for the training set and test set for LXR β modeling (A) and LXR α modeling (B): black dots represent compounds of the training set and black triangles represent compounds of the test set.

The selected variables and their chemical meanings, standard coefficients, significance and variable inflation factor (VIF) were presented in Table 1. It can be seen that all selected descriptors had a statistical significance of less than 0.05, indicating that they were obvious features in defining the activity of LXR β -selective agonists. As shown in Table 1, the VIF of all descriptors was smaller than 5, indicating no multicollinearity existed among the descriptors in models [15]. Table S2 (see Supplementary Materials) lists the correlation matrix of the selected descriptors in the QSAR model. All linear correlation coefficient values for each pair of descriptors were smaller than 0.85, showing that they were independent [16]. The predicted results of QSAR model are given in Table S1 (see Supplementary Materials) and shown in Figure 2. As described in Table 2, obviously, the QSAR model was very successfully built with statistical significance and good prediction ability. The R^2_{train} value of this model reveals that it can explain 83.7% of variance in activity. The Q^2_{LOO} value of 0.715 was bigger than 0.5, indicating that the developed model had very good stability and predictive ability. In addition, the value of R^2_{test} for the external prediction was 0.843, showing good prediction and generalization ability.

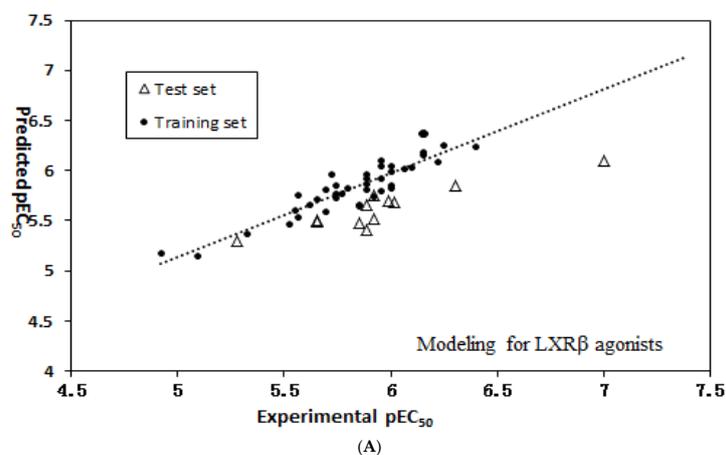


Figure 2. Cont.

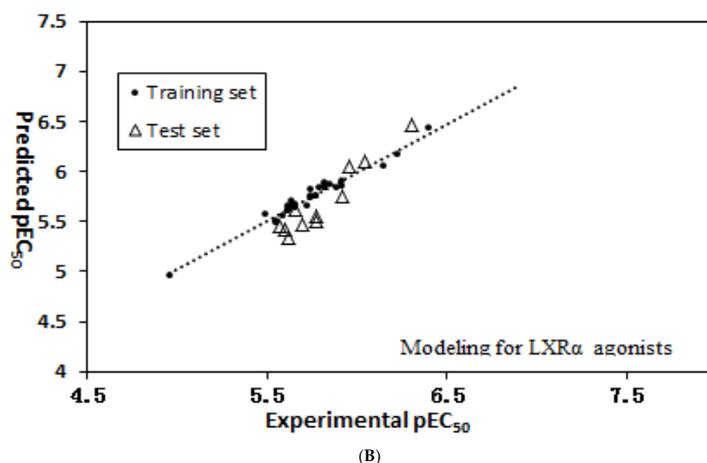


Figure 2. Plots of experimental pEC_{50} values against predicted pEC_{50} values by QSAR models for LXR β agonists (A) and LXR α agonists (B).

Table 1. Selected descriptors of MLR model for LXR β agonists.

Descriptor	Chemical Meaning	Coefficient	Standard Coefficient	VIF	<i>p</i> -Value
vsurf_IW2	Hydrophilic integrity moment	−0.777	−0.592	2.246	0.000
SMR_VSA6	Sum of v_i such that R_i is in (0.485,0.56]	0.007	0.325	1.232	0.000
glob	Globularity, or inverse condition number (smallest eigenvalue divided by the largest eigenvalue) of the covariance matrix of atomic coordinates.	−1.236	−0.668	1.779	0.000
GCUT_SLOGP_2	The GCUT descriptors using atomic contribution to logP	4.560	0.687	2.785	0.000
E_strain	Local strain energy	−60.185	−0.315	1.196	0.000
dipoleX	The x component of the dipole moment	−0.189	−0.358	1.240	0.000
AM1_LUMO	The energy (eV) of the Lowest Unoccupied Molecular Orbital calculated using the AM1 Hamiltonian	−0.247	−0.397	2.712	0.003
vsurf_IW5	Hydrophilic integrity moment	0.154	0.284	1.790	0.007
vsurf_DD13	Contact distances of vsurf_DDmin	0.016	0.212	1.192	0.013
Constant		5.087			

Table 2. Statistical parameters of two QSAR models for LXR β and LXR α agonists.

QSARModel	Training Set					Test Set	
	R^2_{train}	$RMSE_{train}$	F	Q^2_{LOO}	$RMSE_{LOO}$	R^2_{test}	$RMSE_{test}$
LXR beta	0.837	0.118	17.235	0.715	0.156	0.843	0.232
LXR alpha	0.968	0.045	44.068	0.895	0.081	0.914	0.155

2.3. QSAR Model Results of LXR α Activity

After the SW-MLR was performed, the best QSAR model for LXR α activators was generated with twelve molecular descriptors. The obtained QSAR model was generated as follows:

The selected variables and their chemical meanings, standard coefficients, and variable inflation factors are shown in Table 3. The values of VIF and significance showed that these 12 descriptors were

obvious features in defining LXR α activity. All linear correlation coefficient values for each pair of descriptors were smaller than 0.85, showing that they were independent [16]. The predicted results of the QSAR model were listed in Table S3 (see Supplementary Materials) and shown in Figure 2. Table 2 listed the statistical results of the proposed model. As described in Table 2, the obtained QSAR model was very successful and of good predictive ability. The QSAR model can give 96.8% variances in LXR α activity in the training set. The Q^2_{LOO} value of leave-one-out (LOO) cross-validation was 0.895 (much bigger than 0.5), showing that the developed QSAR model had good stability and predictive ability. Additionally, the R^2_{test} for the external prediction also reached 0.914, indicating good prediction and generalization ability of the LXR α QSAR model [17].

Table 3. Selected descriptors of MLR model for LXR α agonists.

Descriptor	Chemical Meaning	Coefficient	Standard Coefficient	VIF	<i>p</i> -Value
GCUT_SLOGP_2	The GCUT descriptors using atomic contribution to logP	8.952	1.539	4.862	0.000
vsurf_DD12	Contact distances of vsurf_DDmin	0.024	0.366	1.348	0.000
Q_VSA_POS	Total positive van der Waals surface area	0.005	0.673	2.161	0.000
SlogP_VSA2	Sum of v_i such that Li is in $(-0.2,0]$	0.023	0.813	4.763	0.000
E_ang	Angle bend potential energy	-0.019	-0.325	2.126	0.000
pmiY	y component of the principal moment of inertia	4.808×10^{-5}	0.209	1.468	0.001
dipoleY	The y component of the dipole moment	0.164	0.281	1.408	0.000
vsurf_DW12	Contact distances of vsurf_EWmin	-0.019	-0.203	1.467	0.001
BCUT_SMR_0	The BCUT descriptors using atomic contribution to molar refractivity	34.772	0.351	2.546	0.000
SlogP_VSA3	Sum of v_i such that Li is in $(0,0.1]$	0.005	0.238	1.590	0.000
vsurf_CW6	Capacity factor	-4.271	-0.273	3.370	0.003
Q_VSA_FPPOS	Fractional positive polar van der Waals surface area	-1.339	-0.144	1.810	0.024
Constant		83.858			

Finally, the Y-randomization tests were performed to confirm the robustness of two QSAR models [18]. Table 4 listed the results of ten Y-randomization tests for these two LXR α and LXR β QSAR models. It can be observed that all new R^2_{train} and Q^2_{LOO} values of the Y-randomization tests were much smaller than those of the original models. Thereby, the two QSAR models with good predictive abilities were not due to a chance correlation or structural dependency of the training set. Overall, these two QSAR models for LXR α and LXR β agonists were quite satisfied, exhibiting the significantly high predictive ability, reliability, and robustness, which can be used to predict LXR β and LXR α activity.

Table 4. R^2_{train} and Q^2_{LOO} values of QSAR models after ten Y-randomization tests.

No. of Test		1	2	3	4	5	6	7	8	9	10
LXR β model	R^2_{train}	0.185	0.182	0.155	0.257	0.282	0.208	0.126	0.244	0.155	0.241
	Q^2_{LOO}	0.063	0.006	0.018	0.001	0.037	0.015	0.081	0.002	0.047	0.003
LXR α model	R^2_{train}	0.259	0.174	0.288	0.287	0.334	0.298	0.25	0.264	0.258	0.287
	Q^2_{LOO}	0.033	0.097	0.086	0.009	0.015	0.039	0.023	0.016	0.032	0.027

2.4. Interpretation of the Descriptors

It is possible to gain some vital structural features to govern the LXR β -selective activity by interpreting the molecular descriptors in the QSAR models. In the QSAR models of LXR β agonists

and LXR α agonists, nine and twelve descriptors were uncovered, respectively. Additionally, only one descriptor (GCUT_SLOGP_2) is the same for two QSAR models. In order to investigate whether there was some correlation between two QSAR models, the correlation coefficients of their descriptors were calculated as listed in Table 5. Obviously, all linear correlation coefficient values for each pair of descriptors between two QSAR models were smaller than 0.6, indicating that these two QSAR models were independent and can well distinguish LXR β and LXR α activity.

2.5. Screening New Highly LXR β -Selective Agonists

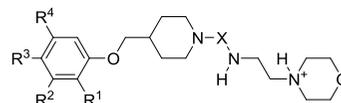
In general, compounds with high structural similarity (bigger than 0.7) to the basic scaffold of the training set will be given more accurate predictions than compounds without similarity [19], and will also have similar biological activity [20]. Thus, to ensure maximum accuracy of these predictions, the structural similarity between compounds in the ZINC database and 3-(4-phenoxybutyl)imidazolidine-2,4-dione skeleton were calculated. Compounds with structural similarity below 0.7 were removed from the ZINC database, which resulted in the retrieval of 637 compounds. The compounds with predictive activity values ($pEC_{50} > 6.0$ for LXR β and $pEC_{50} < 1.0$ for LXR α) were identified as potential LXR β -selective agonists. Thus, 11 compounds were discovered from these molecules based on our QSAR models. Among them, ZINC55084484 had the best LXR β prediction values ($pEC_{50} = 7.343$) and LXR α prediction values ($pEC_{50} = -1.901$), much better in LXR β -selective activity than the best reported compounds ($pEC_{50} = 7.0$ for LXR β and $pEC_{50} = 6.095$ for LXR α) in Table S1 (see Supplementary Materials). Thus, to find highly potent LXR β -selective agonists, we designed new compounds based on ZINC55084484. As listed in Table 6, we found that the presence of absorbent groups such as propionyloxy, propionamido and 2,2,2-trifluoroethylamino at R1 and R3 of the benzene ring can significantly enhance LXR β agonist activity, among which the addition of 2,2,2-trifluoroethylamino at R1 (Table 6) performed best. It was also observed that the addition of a chlorine atom at the ortho or para position of 2,2,2-trifluoroethylamino can lead to better LXR β agonist activity, such as compounds N1 and N3 with predicted pEC_{50} values of 8.497 and 8.429, respectively.

2.6. Molecular Docking Study

Molecular docking embedded in Molecular Operating Environment software (MOE2008.10, Chemical Computing Group Inc., Montreal, QC, Canada) was applied to better understand the binding modes and important interactions of new designed LXR β -selective agonists. In this docking study, the root-mean-square distance (RMSD) parameter of the ligand between the three-dimensional crystal structure of the LXR β complex (PDB: 5JY3) and in the redocked structure was 0.807 Å, showing that these docking parameters were suitable, and the docking results were reliable [21]. Docking results of the newly designed LXR β -selective agonists were listed in Table 6. Obviously, these newly designed compounds had better docking scores for LXR β than the template compound (ZINC55084484), which were in agreement with the QSAR results. It was also observed that the presence of absorbent groups such as propionyloxy and 2,2,2-trifluoroethylamino at R¹ and R³ of benzene ring significantly enhanced the LXR β agonist activity, which almost corresponded with the QSAR results. The best docked conformation of the most active compound N1, as shown in Figure 3, revealed that the presence of 2,2,2-trifluoroethylamino and chlorine at R¹ and R³ of benzene ring allowed for potentiation of strong hydrophilic interactions with Phe340, Ile345, Phe268, Ala343, Phe268, Leu449, Thr272, Leu453, Phe271 and Trp457 in the active site of LXR β . Comparative to molecular docking between compound N1 and the template compound ZINC55084484, shown in Figure 3, the former had a better binding score than the latter. This revealed that carbonyl was not conducive to the activity compared with methylene in the X place, limiting the molecular flexibility, and the presence of 2,2,2-trifluoroethylamino and chlorine at R¹ and R³ of benzene ring allowed for potentiation of the strong hydrophilic interactions in the active site of LXR β . It can be concluded that more H-bonds and hydrophobic interactions between substituent groups at benzene ring with above amino acids were beneficial to the activity.

Table 5. Correlation coefficients of descriptors between two QSAR models.

	AM1_LUMO	GCUT_SLOGP_2	E_strain	dipoleX	SMR_VSA6	vsurf_DD13	vsurf_IW2	vsurf_IW5	Glob
BCUT_SMR_0	0.031	-0.255	0.254	0.199	-0.243	-0.172	0.075	-0.069	-0.063
GCUT_SLOGP_2	-0.055	0.014	0.066	-0.194	-0.371	0.090	0.226	-0.133	0.110
Q_VSA_FPPOS	-0.242	-0.315	0.072	-0.008	-0.041	0.258	-0.178	-0.184	-0.117
Q_VSA_POS	-0.017	0.279	-0.182	-0.164	0.240	0.113	-0.202	-0.012	0.013
E_ang	-0.069	-0.431	0.007	-0.030	0.173	0.531	0.148	-0.261	-0.361
dipoleY	-0.216	0.045	-0.168	-0.189	0.051	-0.001	-0.062	-0.086	0.254
pmiY	0.204	-0.159	-0.195	-0.098	-0.020	0.110	0.216	0.067	-0.353
SlogP_VSA2	0.029	-0.166	-0.012	0.083	0.210	0.054	-0.076	0.051	-0.347
SlogP_VSA3	-0.155	0.246	0.095	-0.032	-0.236	0.015	0.026	0.082	0.085
vsurf_CW6	-0.056	0.002	0.352	-0.231	-0.116	-0.127	-0.137	-0.156	0.201
vsurf_DD12	-0.564	0.005	0.161	0.067	-0.400	0.226	0.243	0.060	-0.211
vsurf_DW12	0.116	-0.075	-0.046	-0.324	0.099	-0.099	-0.092	-0.304	-0.091

Table 6. Chemical structures of newly designed LXR β -selective agonists based on two QSAR Models.

Name	R ¹	R ²	R ³	R ⁴	X	Predicted pEC ₅₀ Values		
						LXR β	LXR α	Docking Scores
ZINC55084484	H	H	H	H	CO	7.343	-1.901	-7.713
N1	2,2,2-trifluoroethylamino	H	H	H	Cl	8.497	-1.911	-11.205
N2	2,2,2-trifluoroethylamino	H	H	H	CH ₂	8.390	-2.076	-9.394
N3	2,2,2-trifluoroethylamino	Cl	H	H	CH ₂	8.429	-1.730	-9.757
N4	2,2,2-trifluoroethylamino	F	H	H	CO	8.328	0.215	-10.236
N5	propionyloxy	H	H	H	CO	8.148	-0.753	-9.528
N6	2,2,2-trifluoroethylamino	H	propionyloxy	H	CH ₂	7.932	-0.9524	-9.323
N7	propionyloxy	H	propionyloxy	H	CO	7.923	-0.905	-9.177
N8	2,2,2-trifluoroethylamino	F	H	H	CH ₂	8.178	-1.760	-10.068
N9	2,2,2-trifluoroethylamino	H	propionyloxy	H	CO	8.111	-1.211	-10.321

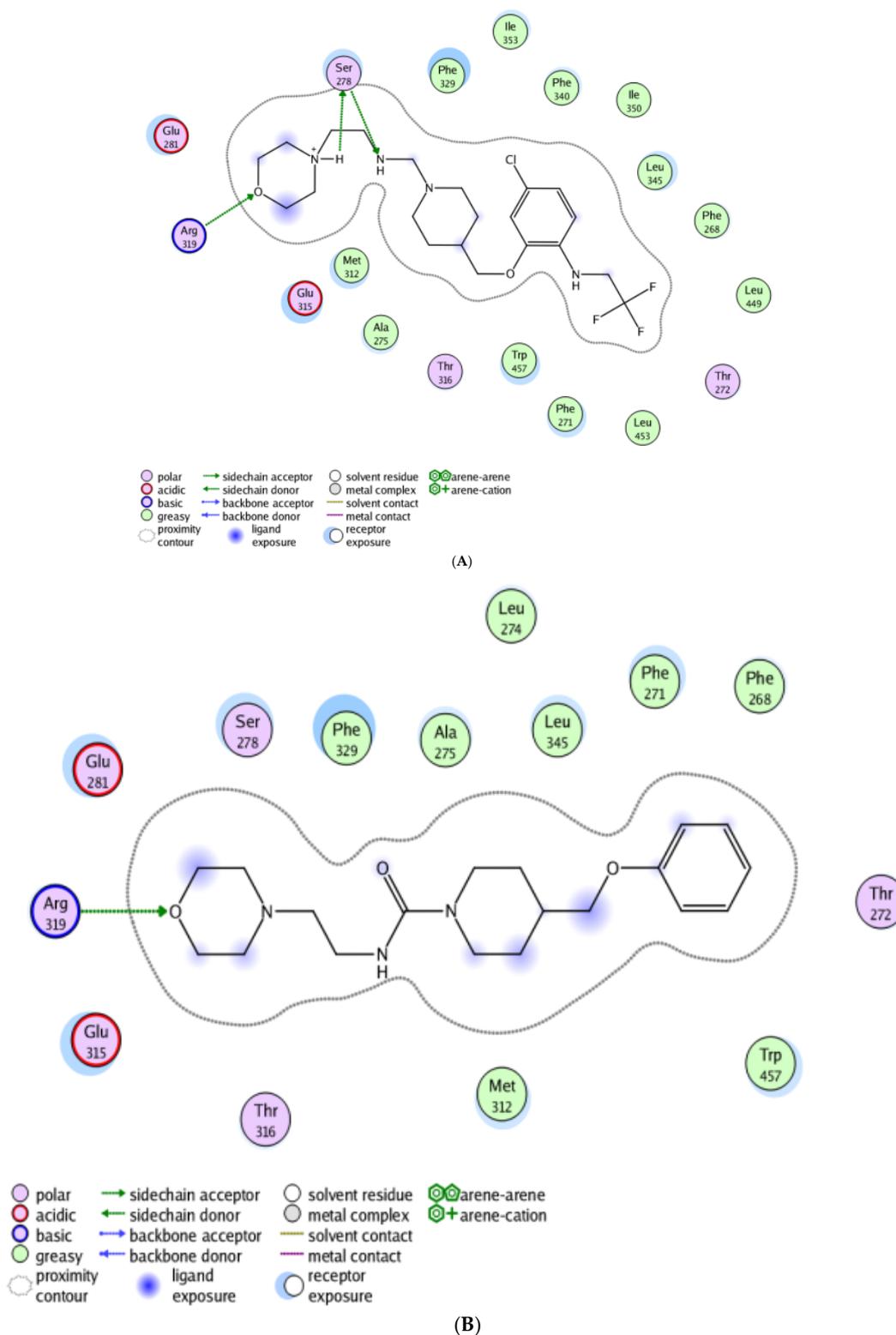


Figure 3. The binding modes of new designed compound N1 (A) and the template compound (ZINC55084484) (B) in the LXR β active site.

Additionally, a drug-likeness analysis was also performed to evaluate the oral drug-like property of these new designed LXR β -selective agonists using the Lipinski rule of five, which predicts that great absorption or permeation is more likely when molecular weight is no more than 500 Da, the number of

H-bond donors (a_don) are less than five, the number of H-bond acceptors (a_acc) are less than 10 and the octanol-water partition coefficient ($\log P(o/w)$) is lower than five [19,22]. The drug-like property descriptors of these compounds were listed in Table 7. All values of five such descriptors were largely coincided with the five rules of oral medications. Thereby, these nine newly designed compounds were suggested to be highly LXR β -selective agonists.

Table 7. Drug-like property descriptors of new designed LXR β -selective agonists.

Name	Predicted pEC ₅₀ Values		Weight	a_acc	a_don	logP(o/w)
	LXR β	LXR α				
ZINC55084484	7.343	−1.901	347.459	4	1	1.218
N1	8.497	−1.911	465.968	4	2	2.757
N2	8.390	−2.076	431.523	4	2	2.128
N3	8.429	−1.730	465.968	4	2	2.718
N4	8.328	0.215	463.496	3	2	1.777
N5	8.148	−0.753	420.53	4	1	1.376
N6	7.932	−0.9524	503.586	5	2	2.573
N7	7.923	−0.905	492.593	5	1	1.821
N8	8.178	−1.760	449.513	4	2	2.279
N9	8.111	−1.211	517.569	4	2	2.071

3. Materials and Methods

3.1. Dataset Division

In this QSAR analysis, a series of fifty-three 3-(4-(2-propylphenoxy)butyl)imidazolidine-2,4-dione-based LXR α / β dual agonists were taken from K. Shibuya [23,24] and the basic scaffold was presented in Figure 4. The activated activity (EC₅₀) values were covered to logarithmic scale pEC₅₀ values, which were used as the dependent parameters in the QSAR study. The molecular structures and activity data of LXRs agonists were presented in Table S1 (see Supplementary Materials). All 2D structures of compounds in Table S1 were sketched using ChemDraw software and were converted into 3D structures using energy minimization module embedded in Molecular Operating Environment software (MOE2008.10, Chemical Computing Group Inc., Montreal, QC, Canada). Then, their conformer structures were optimized by stochastic conformational search and followed to generate 327 diverse descriptors by utilizing the QSAR module of MOE [25]. The redundant information among descriptors was conducted by deleting constant or almost constant values for all molecules and removing one of inter-correlated descriptors (a pairwise correlation coefficient greater than 0.95) [26]. Finally, a total set of 183 descriptors remained and were used for QSAR modeling. To obtain reliable QSAR models, the studied chemicals were firstly separated into a training set and a test set using a Kohonen's self-organizing map (5 × 5 neurons, 500 epochs), which ensured the training set spanned the whole descriptor space and kept a balanced distribution of the chemicals in two data sets [27].

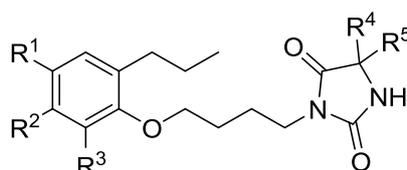


Figure 4. 3-(4-(2-propylphenoxy)butyl)imidazolidine-2,4-dione based LXR α / β dual agonists.

3.2. Stepwise Multiple Linear Regression (SW-MLR)

Feature selection is considered as one of the key steps in development of 2D-QSAR models. In this study, a stepwise technology combined with MLR (SW-MLR) was applied to select a suitable set of

descriptors that could be used as input values for model construction. The stepwise regression is a variation on forward selection. At each stage in the process, after adding a new variable, an F-test was performed to check if some variables could be removed without significantly increasing the residual sum of squares [28]. So, different MLR models were developed in this procedure. The statistical parameters such as squared correlation coefficient (R^2), root mean standard error (RMSE), and Fisher statistic were calculated to assess the performance of derived QSAR models [29].

3.3. Model Validation

Model validation is a critical step in assessing the predictive ability and reliability of QSAR models. It includes internal and external validations. Generally, the leave-one-out (LOO) cross-validation technology is often considered as the most economical and popular internal validation to evaluate the predictive ability of the model [30]. LOO cross-validation involves using one object from the dataset as the validation set, and the remaining dataset serves as the training data. This is repeated so that every object in the dataset is used once as the validation data, which employs all the information available. Usually, the model is acceptable when the value of LOO cross-validation squared correlation coefficient (Q^2_{LOO}) is bigger than 0.5 [13]. Moreover, external validation is significant and essential to evaluate the generalization performance of the proposed model. The statistical parameters, such as the squared correlation coefficient (R^2_{test}) and root mean square errors ($\text{RMSE}_{\text{test}}$) of the test set were calculated to evaluate the performance of the model [29].

All algorithms were written in MATLAB 8.0 and run on a computer [Intel(R) Pentium(R), 2.00-GB RA].

3.4. Screening New LXR β -Selective Agonists

The ZINC database that contained over 35 million diverse purchasable compounds was subjected to our QSAR model prediction for discovering new highly-potent LXR β -selective agonists [31]. Given that our QSAR models were constructed based on a 3-(4-phenoxybutyl)imidazolidine-2,4-dione skeleton, only compounds with these skeletons can be well-predicted by our QSAR models. Thereby, the molecular structural similarity between compounds in ZINC database and 3-(4-phenoxybutyl)imidazolidine-2,4-dione skeleton was first calculated using the Tanimoto coefficient in Open Babel 2.3.1 [32]. Generally, a good cutoff for the Tanimoto coefficient for biologically similar molecules is 0.7 or 0.8 [20]. Here, compounds with a structural similarity bigger than 0.7 were selected out from the ZINC database and imported into MOE for further analysis. Hydrogen atoms and partial charges were assigned, and then they were energy minimized using the molecular mechanics force field method with a convergence criterion of 0.01 kcal/mol. Then, the two above obtained QSAR models were applied to screen new LXR β -selective agonists from these compounds. Subsequently, to discover highly potent LXR β -selective agonists that do not activate LXR α , the models were further applied to design new compounds based on above screened compounds from ZINC database.

3.5. Molecular Docking Study

Molecular docking was further performed to study the binding modes and important interactions of new designed LXR β -selective agonists. The docking simulation was carried out as follows [33]. First, the three-dimensional crystal structure of the LXR β -GW3965 complex from the RSCB protein databank (PDB: 5JY3) was protonated using AMBER99 force field and minimized with a RMSD gradient of 0.05 kcal/mol Å. In addition, the binding site and docking placement were using the ligand atom mode and trianglematcher algorithm, respectively. Finally, two rescoring methods including London dG and Affinity dG, along with the force field method, were adopted to compute the interactions.

4. Conclusions

In this paper, modeling techniques such as Kohonen and SW-MLR, structural similarity analysis, and molecular docking were successfully applied to establish models to develop highly potent

LXR β -selective agonists without activation of LXR α based on a series of newly reported LXR α / β dual agonists. The best obtained QSAR model for LXR β can explain 83.7% of the variance in activity with a low RMSE of 0.118, and the best derived QSAR model for LXR α can give better predictive ability with R^2_{train} of 0.968 and RMSE of 0.045. Also, the two QSAR models uncovered approximately different important features in defining LXR α and LXR β activity. They were independent and could well distinguish LXR β and LXR α activity. A total of 11 compounds from the ZINC database that fulfilled the requirement of structural similarity of 0.7 compared to known dual LXR α / β agonists were predicted with activity values of $\text{pEC}_{50} > 6.0$ for LXR β and $\text{pEC}_{50} < 1.0$ for LXR α . Among them, ZINC55084484 had the best LXR β prediction values ($\text{pEC}_{50} = 7.343$) and LXR α prediction value ($\text{pEC}_{50} = -1.901$), much better in LXR β -selective activity than the best reported compounds in Table S1 ($\text{pEC}_{50} = 7.0$ for LXR β and $\text{pEC}_{50} = 6.095$ for LXR α). Thereupon, nine new compounds were designed as highly potent LXR β -selective agonists based on ZINC55084484, of which LXR β prediction values almost surpassed 8 and LXR α prediction values were below 0. Additionally, the docking results of the newly designed LXR β -selective agonists corresponded with the QSAR results well. The best docked conformation of the most active compound N1 revealed that carbonyl was not conducive to the activity compared with methylene in the X place, limiting the molecular flexibility, and the presence of 2,2,2-trifluoroethylamino and chlorine at R¹ and R³ of the benzene ring allowed for potentiation of strong hydrophilic interactions in the active site of LXR β . Overall, this study could provide valuable guidance for the future design of highly potent LXR β -selective agonists in the drug discovery process.

Supplementary Materials: The following are available online, Table S1: Molecular structures and corresponding pEC_{50} values of experimental and predicted of the 3-(4-(2-propylphenoxy)butyl)imidazolidine-2,4-dione based LXR α / β dual agonists, Table S2: The correlation matrix of descriptors of LXR β QSAR model, Table S3: The correlation matrix of descriptors of LXR α QSAR model.

Author Contributions: M.C. contributed to the analysis of the study and manuscript writing; J.K., H.G., X.Y., X.L. and Y.G. did data analysis; F.Y. designed the experiments.

Funding: This work is funded by National Natural Science Foundation program of China (81503497, 21406036 and 8167151245), Fujian Provincial Natural Science fund subject of China (2015J01340 and 2017J01571), and Fujian Education Department of China: Fujian Provincial universities' incubation project for prominent young scientific researchers, and Fujian 2011 Chinese Medicine Health Management Collaboration Center of China (JG2017001).

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

References

1. Moriyama, Y.; Okamura, T.; Inazu, A.; Doi, M.; Iso, H.; Mouri, Y.; Ishikawa, Y.; Suzuki, H.; Iida, M.; Koizumi, J.; et al. A low prevalence of coronary heart disease among subjects with increased high-density lipoprotein cholesterol levels, including those with plasma cholesteryl ester transfer protein deficiency. *Prev. Med.* **2016**, *27*, 659–667. [[CrossRef](#)] [[PubMed](#)]
2. Ma, Z.; Deng, C.; Hu, W.; Zhou, J.; Fan, C.; Di, S.; Liu, D.; Yang, Y.; Wang, D. Liver X receptors and their agonists: Targeting for cholesterol homeostasis and cardiovascular diseases. *Curr. Issues Mol. Biol.* **2017**, *22*, 41–64. [[CrossRef](#)] [[PubMed](#)]
3. Breevoort, S.R.; Angdisen, J.; Schulman, I.G. Macrophage-independent regulation of reverse cholesterol transport by liver x receptors. *Arterioscler. Thromb. Vasc. Biol.* **2014**, *34*, 1650–1660. [[CrossRef](#)] [[PubMed](#)]
4. Hansen, M.K.; Connolly, T.M. Nuclear receptors as drug targets in obesity, dyslipidemia and atherosclerosis. *Curr. Opin. Investig. Drugs* **2008**, *9*, 247–255. [[PubMed](#)]
5. Fitz, N.F.; Cronican, A.; Pham, T.; Fogg, A.; Fauq, A.H.; Chapman, R.; Lefterov, I.; Koldamova, R. Liver X receptor agonist treatment ameliorates amyloid pathology and memory deficits caused by high-fat diet in APP23 mice. *J. Neurosci.* **2010**, *30*, 6862–6872. [[CrossRef](#)] [[PubMed](#)]
6. Gabbi, C.; Warner, M.; Gustafsson, J.Å. Action mechanisms of liver x receptors. *Biochem. Biophys. Res. Commun.* **2014**, *446*, 647–650. [[CrossRef](#)] [[PubMed](#)]
7. Jin, S.H.; Yang, J.H.; Shin, B.Y.; Seo, K.; Shin, S.M.; Cho, I.J.; Ki, S.H. Resveratrol inhibits lxr α -dependent hepatic lipogenesis through novel antioxidant sestrin2 gene induction. *Toxicol. Appl. Pharm.* **2013**, *271*, 95–105. [[CrossRef](#)] [[PubMed](#)]

8. Hu, Y.; Yang, Y.; Yu, Y.; Wen, G.; Shang, N.; Zhuang, W.; Lu, D.; Zhou, B.; Liang, B.; Yue, X.; et al. Synthesis and identification of new flavonoids targeting liver x receptor β involved pathway as potential facilitators of $\alpha\beta$ clearance with reduced lipid accumulation. *J. Med. Chem.* **2013**, *56*, 6033–6053. [[CrossRef](#)] [[PubMed](#)]
9. Lund, E.G.; Peterson, L.B.; Adams, A.D.; Lam, M.N.; Burton, C.A.; Chin, J.; Guo, Q.; Huang, S.; Latham, M.; Lopez, J.C.; et al. Different roles of liver X receptor alpha and beta in lipid metabolism: Effects of an alpha-selective and a dual agonist in mice deficient in each subtype. *Biochem. Pharmacol.* **2006**, *71*, 453–463. [[CrossRef](#)] [[PubMed](#)]
10. Von Grafenstein, S.; Mihaly-Bison, J.; Wolber, G.; Bochkov, V.N.; Liedl, K.R.; Schuster, D. Identification of Novel Liver X Receptor Activators by Structure-Based Modeling. *J. Chem. Inf. Model.* **2012**, *52*, 1391–1400. [[CrossRef](#)] [[PubMed](#)]
11. Li, Y.; Wang, L.; Liu, Z.; Li, C.; Xu, J.; Gu, Q.; Xu, J. Predicting selective liver X receptor β agonists using multiple machine learning methods. *Mol. Biosyst.* **2015**, *11*, 1241–1250. [[CrossRef](#)] [[PubMed](#)]
12. Peng, H.; Liu, Z.; Yan, X.; Ren, J.; Xu, J. A *de novo* substructure generation algorithm for identifying the privileged chemical fragments of liver x receptor β agonists. *Sci. Rep.* **2017**, *7*, 11121. [[CrossRef](#)] [[PubMed](#)]
13. Temml, V.; Voss, C.V.; Dirsch, V.M.; Schuster, D. Discovery of new liver x receptor agonists by pharmacophore modeling and shape-based virtual screening. *J. Chem. Inf. Model.* **2014**, *54*, 367–371. [[CrossRef](#)] [[PubMed](#)]
14. Hegymegibarakonyi, B.; Orfi, L.; Kéri, G.; Kövesdi, I. Application of Kohonen Self-Organizing feature maps in QSAR of human ADMET and kinase data sets. *Acta Pharm. Hung.* **2013**, *83*, 143–148.
15. Wang, Y.; Chen, J.; Yang, X.; Lyakurwa, F.; Li, X.; Qiao, X. In silico model for predicting soil organic carbon normalized sorption coefficient (K(OC)) of organic chemicals. *Chemosphere* **2015**, *119*, 438–444. [[CrossRef](#)] [[PubMed](#)]
16. Jalali-Heravi, M.; Asadollahi-Baboli, M.; Shahbazikhah, P. QSAR study of heparanase inhibitors activity using artificial neural networks and Levenberg-Marquardt algorithm. *Eur. J. Med. Chem.* **2008**, *43*, 548–556. [[CrossRef](#)] [[PubMed](#)]
17. You, W.; Yang, Z.; Ji, G. Feature selection for high-dimensional multi-category data using pls-based local recursive feature elimination. *Expert Syst. Appl.* **2014**, *41*, 1463–1475. [[CrossRef](#)]
18. Rastija, V.; Agiä, D.; Tomiä, S.; Nikoliä, S.; Hranjec, M.; Grace, K.Z.; Abramić, M. Synthesis, QSAR, and molecular dynamics simulation of amidino-substituted benzimidazoles as dipeptidyl peptidase iii inhibitors. *Acta Chim. Slov.* **2015**, *62*, 867–878. [[CrossRef](#)] [[PubMed](#)]
19. Lipinski, C.A.; Lombardo, F.; Dominy, B.W.; Feeney, P.J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.* **2001**, *46*, 3–26. [[CrossRef](#)]
20. Singh, K.P.; Gupta, S.; Rai, P. Predicting acute aquatic toxicity of structurally diverse chemicals in fish using artificial intelligence approaches. *Ecotoxicol. Environ. Saf.* **2013**, *95*, 221–233. [[CrossRef](#)] [[PubMed](#)]
21. Wei, Y.; Li, J.; Chen, Z.; Wang, F.; Huang, W.; Hong, Z.; Lin, J. Multistage virtual screening and identification of novel HIV-1 protease inhibitors by integrating SVM, shape, pharmacophore and docking methods. *Eur. J. Med. Chem.* **2015**, *101*, 409–418. [[CrossRef](#)] [[PubMed](#)]
22. Chen, M.; Yang, F.; Yang, X.; Lai, X.; Gao, Y. Systematic understanding of mechanisms of a chinese herbal formula in treatment of metabolic syndrome by an integrated pharmacology approach. *Int. J. Mol. Sci.* **2016**, *17*, 2114. [[CrossRef](#)] [[PubMed](#)]
23. Koura, M.; Matsuda, T.; Okuda, A.; Watanabe, Y.; Yamaguchi, Y.; Kurobuchi, S.; Matsumoto, Y.; Shibuya, K. Design, synthesis and pharmacology of 1,1-bistrifluoromethylcarbinol derivatives as liver X receptor β -selective agonists. *Bioorg. Med. Chem. Lett.* **2015**, *25*, 2668–2674. [[CrossRef](#)] [[PubMed](#)]
24. Matsuda, T.; Okuda, A.; Watanabe, Y.; Miura, T.; Ozawa, H.; Tosaka, A.; Yamazaki, K.; Yamaguchi, Y.; Kurobuchi, S.; Koura, M.; et al. Design and discovery of 2-oxochromene derivatives as liver X receptor β -selective agonists. *Bioorg. Med. Chem. Lett.* **2015**, *25*, 1274–1278. [[CrossRef](#)] [[PubMed](#)]
25. Chen, M.; Yang, X.; Lai, X.; Gao, Y. 2D and 3D QSAR models for identifying diphenylpyridylethanamine based inhibitors against cholesteryl ester transfer protein. *Bioorg. Med. Chem. Lett.* **2015**, *25*, 4487–4495. [[CrossRef](#)] [[PubMed](#)]
26. Chen, M.; Yang, X.; Lai, X.; Kang, J.; Gan, H.; Gao, Y. Structural Investigation for Optimization of Anthranilic Acid Derivatives as Partial FXR Agonists by in Silico Approaches. *Int. J. Mol. Sci.* **2016**, *17*, 536. [[CrossRef](#)] [[PubMed](#)]

27. Toropov, A.A.; Toropova, A.P. Quasi-QSAR for mutagenic potential of multi-walled carbon-nanotubes. *Chemosphere* **2015**, *124*, 40–46. [[CrossRef](#)] [[PubMed](#)]
28. Ghorbanzadeh, M.; Zhang, J.; Andersson, P.L. Binary classification model to predict developmental toxicity of industrial chemicals in zebrafish. *J. Chemometr.* **2016**, *30*, 298–307. [[CrossRef](#)]
29. Bolboacă, S.D.; Jäntschi, L. Sensitivity, specificity, and accuracy of predictive models on phenols toxicity. *J. Comput. Sci.* **2014**, *5*, 345–350. [[CrossRef](#)]
30. Zou, S.; Zhang, J.; Zhang, Z. A novel approach for predicting microbe-disease associations by bi-random walk on the heterogeneous network. *PLoS ONE* **2017**, *12*, e0184394. [[CrossRef](#)] [[PubMed](#)]
31. Irwin, J.J.; Shoichet, B.K. ZINC—A free database of commercially available compounds for virtual screening. *J. Chem. Inf. Model.* **2005**, *45*, 177–182. [[CrossRef](#)] [[PubMed](#)]
32. O’Boyle, N.M.; Banck, M.; James, C.A.; Morley, C.; Vandermeersch, T.; Hutchison, G.R. Open Babel: An open chemical toolbox. *J. Cheminform.* **2011**, *3*, 33. [[CrossRef](#)] [[PubMed](#)]
33. Hegazy, M.F.; Ibrahim, A.Y.; Mohamed, T.A.; Shahat, A.A.; El Halawany, A.M.; Abdelazim, N.; Alsaid, M.S.; Paré, P.W. Sesquiterpene lactones from cynara cornigera: Acetyl cholinesterase inhibition and in silico ligand docking. *Planta Med.* **2016**, *82*, 138–146. [[PubMed](#)]

Sample Availability: Samples of the compounds are not available from the authors.



© 2018 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 (<http://creativecommons.org/licenses/by/4.0/>).