

Three-Dimensional Pharmacophore Modeling of Liver-X Receptor Agonists

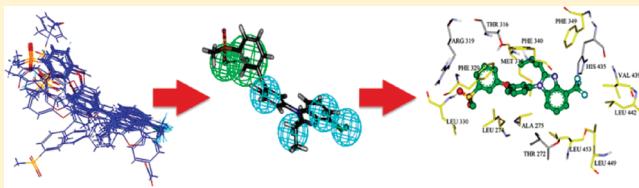
Wenxia Zhao,[†] Qiong Gu,[†] Ling Wang,[†] Hu Ge,[†] Jiabo, Li,^{*,‡} and Jun Xu^{*,†}

[†]Research Center for Drug Discovery and Institute of Human Virology, School of Pharmaceutical Sciences, Sun Yat-Sen University, 132 East Circle at University City, Guangzhou, China 510006

[‡]Accelrys, Inc., 10188 Telesis Ct # 100, San Diego, California 92121-4779, United States

 Supporting Information

ABSTRACT: High cholesterol levels contribute to hyperlipidemia. Liver X receptors (LXRs) are the drug targets. LXRs regulate the cholesterol absorption, biosynthesis, transportation, and metabolism. Novel agonists of LXR, especially LXR β , are attractive solutions for treating hyperlipidemia. In order to discover novel LXR β agonists, a three-dimensional pharmacophore model was built based upon known LXR β agonists. The model was validated with a test set, a virtual screening experiment, and the FlexX docking approach. Results show that the model is capable of predicting a LXR β agonist activity. Ligand-based virtual screening results can be refined by cross-linking by structure-based approaches. This is because two ligands that are mapped in the same way to the same pharmacophore model may have significantly different binding behaviors in the receptor's binding pocket. This paper reports our approach to identify reliable pharmacophore models through combining both ligand- and structure-based approaches.



1. INTRODUCTION

Cardiovascular diseases are the leading causes of death in the world. More than 17 million people die from them each year.¹ Hyperlipidemia is the major risk factor for cardiovascular diseases; high levels of cholesterol usually induce coronary heart diseases and atherosclerosis.² In the recent years, liver X receptors (LXRs) have gained great attention as promising targets for the treatment of hyperlipidemia. Numerous studies have demonstrated that LXRs are key sensors in the maintainence of cholesterol homeostasis.^{3–6}

LXRs have two subtypes: LXR α and LXR β .⁷ The LXR α expression predominates in the liver, adipose, kidney, adrenal tissues, and macrophages. LXR β is expressed ubiquitously.⁸ Human LXR α and LXR β share almost 78% homology in their ligand-binding domain and appear to respond to the same endogenous ligands.⁹ These subtypes can form functionally active heterodimers with retinoid X receptors (RXRs). Both RXRs and LXRs agonists can activate a LXR response element (LXRE). This activation increases the expression of several genes including the ATP-binding cassette (ABCs) transporters (ABCA1, ABCG5, and ABCG8),¹⁰ apolipoprotein E (apoE), and sterol regulatory element binding protein 1c (SREBP1c).¹¹ ABCA1 is an important transporter which can promote reverse cholesterol transport (RCT). RCT is a pathway which transports accumulated cholesterol, by plasma lipoproteins, from peripheral tissues to the liver for excretion as bile salts.¹² SREBP1c regulates multiple genes involved in cholesterol biosynthesis. The up-regulation of SREBP-1c can increase liver triglyceride (TG) synthesis.¹³ The best way to

reduce cholesterol is selectively activate RCT without impacting the synthesis of TG. Recent studies have supported that selective LXR β agonists affect the RCT while have less of an impact on TG synthesis.¹⁴

The activation of LXRs plays an important role in controlling cholesterol absorption, biosynthesis, transportation, and metabolism.⁸ For the hyperlipidemia therapy, the activation of LXRs (especially LXR β) will be an important mechanism of action. To date, only a few structural classes of LXR agonists are known. Pfizer (Wyeth Pharmaceuticals) scientists have reported a series of nanomolar level LXR agonists.^{15,16} The goal of this study is to generate three-dimensional (3D) QSAR pharmacophore models based upon the known LXR β agonists in order to seek a new generation of LXR β agonists.

2. MATERIALS AND METHODS

2.1. Selecting LXR β Agonists. Sixty LXR β agonists were selected for this study from the literature.^{15–21} Their experimental binding affinities were represented in IC₅₀. The binding affinities were determined by using recombinant human ligand binding domains (LBDs) of LXR β receptor to measure the displacement of reference ligand from the LBD. Lower IC₅₀ values indicate higher binding affinity. All 60 LXR β agonists were divided into a training set and a test set. For the training set, 19

Received: December 31, 2010

Published: March 24, 2011

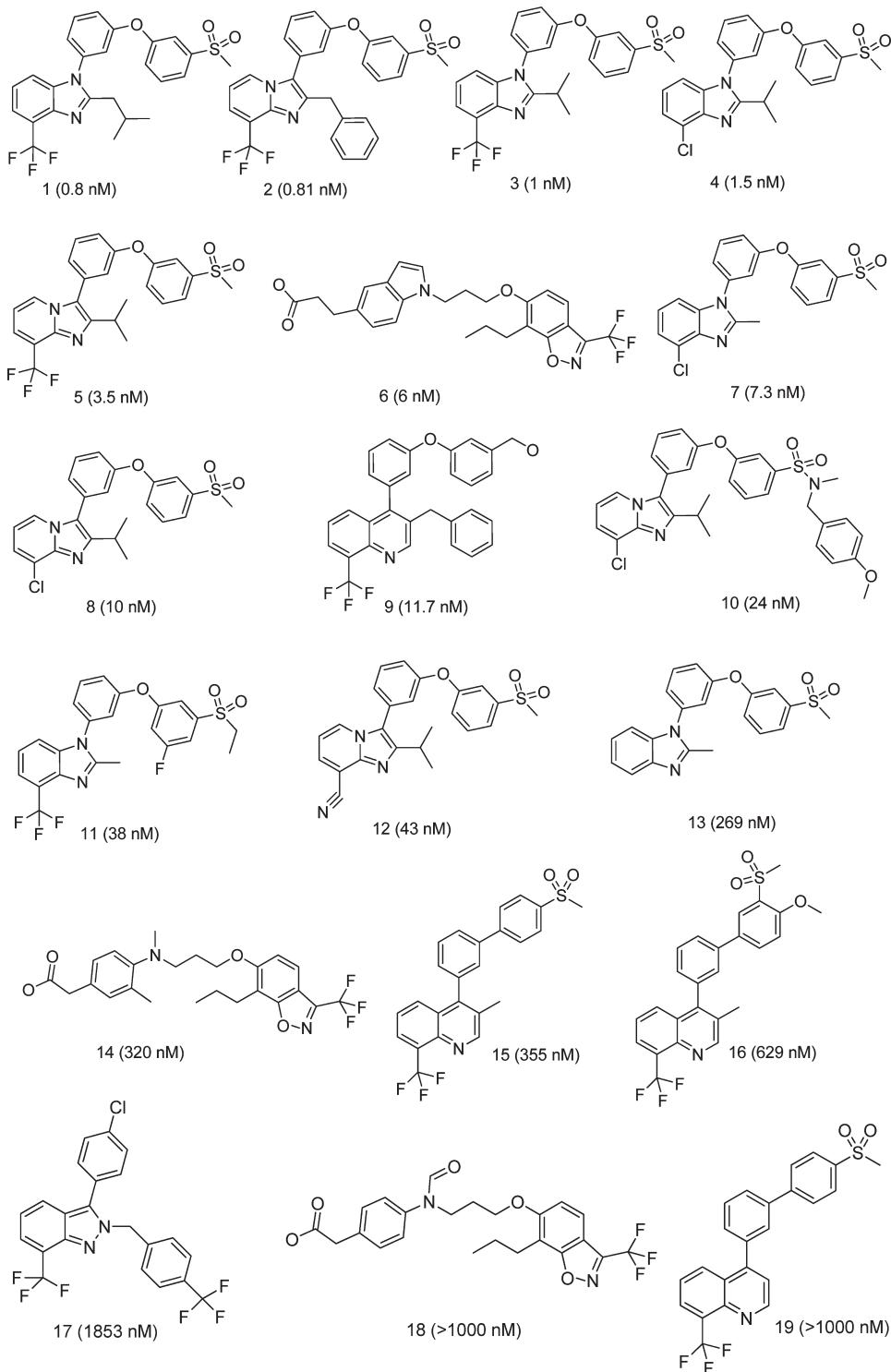


Figure 1. Structures of all 19 training set molecules used for HypoGen pharmacophore generation. IC_{50} values were listed in the parentheses.

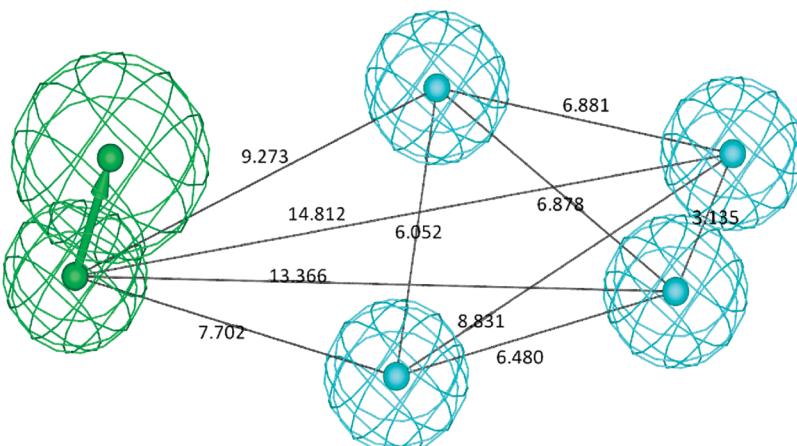
LXR β agonists were selected based upon both chemical scaffold diversity and a wide value range of activities. There are three scaffolds in the training set. The biological activities of the LXR β agonists in the training set span a range of four orders of magnitude: from 0.80 nM to greater than 1000 nM. The active distributions of 19 compounds were listed as follows: 2 compounds with $IC_{50} < 1$ nM, 5 compounds with $1 \text{ nM} \leq IC_{50} < 10$ nM, 5 compounds with $10 \text{ nM} \leq IC_{50} < 100$ nM, 4 compounds

with $100 \text{ nM} \leq IC_{50} < 1000$ nM, and 3 compounds with $IC_{50} > 1000$ nM. The remaining 41 compounds were considered as a test set, which was used for the pharmacophore validation. Figure 1 shows the chemical structures and IC_{50} values of compounds in the training set. The activity value of all the compounds were classified into three categories: high active (≤ 100 nM, +++), moderate active (100–1000 nM, ++), and less active (≥ 1000 nM, +).

Table 1. Ten Pharmacophore Hypotheses Generated by HypoGen for LXR β Agonists

hypo no.	total cost	cost difference ^a	rmsd ^b	correlation	max fit ^c	features ^d	error
1	81.24	111.31	0.79	0.98	10.12	HBA,H,H,H,H	61.08
2	86.43	106.12	1.11	0.96	9.77	HBA,H,H,H,H	66.78
3	89.84	102.71	1.22	0.95	10.28	HBA,H,H,H,RA	69.26
4	91.42	101.13	1.32	0.94	9.82	HBA,H,H,H,RA	71.72
5	91.85	100.70	1.34	0.94	9.48	HBA,H,H,H,RA	72.22
6	92.24	100.31	1.35	0.94	9.28	HBA,H,H,H,RA	72.53
7	92.53	100.02	1.36	0.93	8.71	HBA,H,H,H,RA	72.84
8	94.50	98.05	1.44	0.93	8.85	HBA,HBA,H,H,H	74.88
9	94.53	98.02	1.43	0.93	8.42	HBA,H,H,H, RA	74.53
10	94.57	97.98	1.44	0.93	8.63	HBA,HBA,H,H,H	74.74

^a Cost difference means the difference between null cost and total cost; null cost = 192.55, fixed cost = 74.77, and configuration = 17.19. All cost units are in bits. ^b Rmsd is root-mean-square deviation. ^c Max. fit is the maximum fit value for each hypothesis. The fit function does not merely check if ligand features are mapped generated pharmacophore features or not, it also contains a distance term, which measures the distance that separates the feature on the molecule from the centroid of the hypothesis feature. Both terms are used to calculate the geometric fit value. ^d Abbreviations used for features: HBA, hydrogen-bond acceptor; H, hydrophobic; and RA, ring aromatic.

**Figure 2.** The best pharmacophore model Hypo 1 with distance constraints. Features are color coded as follows: green: HBA; and cyan: HY.

2.2. Generating Conformations. In order to figure out 3D pharmacophore features from the training set, all proper conformations have to be sampled from the conformation space of the training set. Most of pharmacophore modeling studies start with simple energy minimized conformations.^{22,23} However, for those compounds with larger structures and more rotatable bonds, their energy minimized conformations can be different from the conformations for activities. Induced fit theory considers that a molecule should rearrange itself to fit into the protein's active site; the energy spent for the rearrangement is compensated by protein–ligand binding.^{24–26} Because the structures in our training set are larger and the LXR β crystal structures are known, the initial conformations of the structures in the training set are generated from ligand–receptor docking procedure in order to easily reach optimized binding modes. 3KFC was used for generating the conformations by docking the compounds in the training set to the active site of LXR β .²⁷ The docking algorithm is FlexX.²⁸ For each compound in the training set, multiple conformations were generated from the docking process; the best scored conformation was kept as the initial conformation for diversified conformation space generation. The diversified conformations were generated through Molecular Operating Environment (MOE). For each compound in the

training set, up to 250 conformations can be generated under MMFF-94 force field. The conformation spaces were used for the 3D pharmacophore modeling.

2.3. Modeling 3D Pharmacophores. The Catalyst module of the Discovery Studio 2.1 (DS 2.1) was used to model the 3D pharmacophores.²⁹ The Catalyst module requires us to select pharmacophore features from a list: hydrogen-bond acceptor (HBA), hydrogen-bond acceptor_lipid, hydrogen-bond donor (HBD), hydrophobic (HY), hydrophobic_aromatic, hydrophobic_aliphatic, pos_ionizable, neg_ionizable, pos_charge, neg_charge, and ring_aromatic (RA). After analyzing the structure features in our training set, the charge-related pharmacophore features (pos_ionizable, neg_ionizable, pos_charge, neg_charge) were excluded because there were no charge groups in the training set. Considering the binding pocket of LXR β is hydrophobic, we select the hydrophobic feature to construct the pharmacophore model. Finally, HBA, HBD, HY, and RA features were selected for building 3D pharmacophore models.

Based upon the selected pharmacophore features, the Catalyst generated 10 pharmacophore hypotheses by combining pharmacophore feature consensus and the related activity values.

2.4. Validating Pharmacophores. The purpose of the pharmacophore validation is to evaluate the quality of a pharmacophore

Table 2. Experimental and Predicted IC₅₀ Values of the 19 Training Set Compounds Based on the Pharmacophore Model Hypo 1

compound no.	experimental IC ₅₀ , nM	predicted IC ₅₀ , nM	error ^a	fit value	experimental scale ^b	predicted scale ^b
1	0.8	0.77	-1.04	10.12	+++	+++
2	0.81	1.54	1.91	9.82	+++	+++
3	1	1.67	1.67	9.78	+++	+++
4	1.5	1.00	-1.50	10.00	+++	+++
5	3.5	4.45	1.27	9.36	+++	+++
6	6	11.54	1.92	8.94	+++	+++
7	7.3	6.31	-1.16	9.20	+++	+++
8	10	2.70	-3.70	9.57	+++	+++
9	11.7	34.71	2.97	8.46	+++	+++
10	24	36.99	1.54	8.44	+++	+++
11	38	5.79	-6.57	9.24	+++	+++
12	43	59.34	1.38	8.23	+++	+++
13	269	352.46	1.31	7.46	++	++
14	320	145.87	-2.19	7.84	++	++
15	355	325.44	-1.09	7.49	++	++
16	629	298.11	-2.11	7.53	++	++
17	1853	1568.44	-1.18	6.81	+	+
18	>1000	1226.28	-1.22	6.92	+	+
19	>1000	1954.52	1.30	6.71	+	+

^aA + indicates that the predicted IC₅₀ value is higher than the experiment IC₅₀ value, and a - indicates that the predicted IC₅₀ value is lower than the experiment IC₅₀ value. ^bActivity scale: high active, +++ (IC₅₀ < 100 nM); moderate active, ++ (100 nM ≤ IC₅₀ ≤ 1000 nM); low active, + (IC₅₀ ≥ 1000 nM).

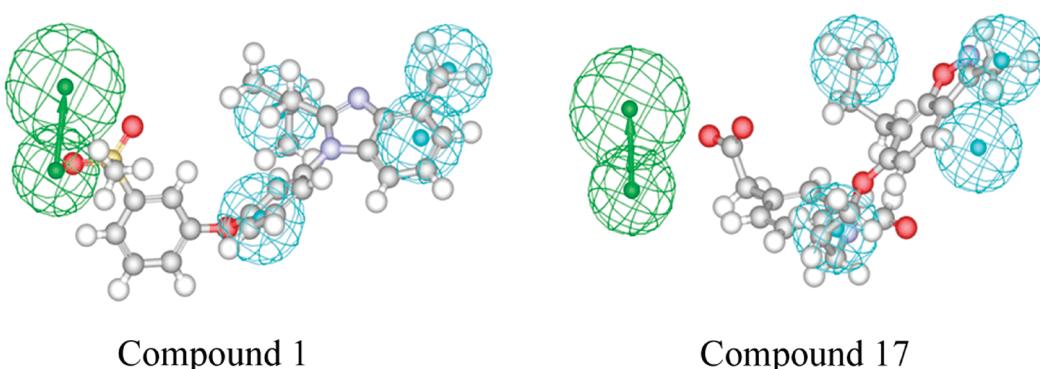


Figure 3. In the training set, Compound 1 is the most active compound because it fits the Hypo 1 very well. Compound 17 has the lowest activity because it fits the three hydrophobic features, but it is off from the hydrogen acceptor feature in Hypo 1. Features are color coded as follows: green: HBA; and cyan: HY.

hypothesis. The significance of a hypothesis was evaluated by a number of parameters: correlation coefficient, total cost, null cost, root-mean-square deviation (rmsd), configuration cost, and fit value. Apart from the above calculated parameters, 41 diversified compounds were used to test each hypothesis. A cross validation approach was also performed by using the data retrieved from Fischer's randomization test. A confidence level of 95% was set (significance = $[1 - (1 + 0)/(19 + 1)] \times 100\% = 95\%$), and 19 random spreadsheets (random hypotheses) were generated for significance tests. The multiple orthogonal testing methods ensure the selection of the best pharmacophore hypothesis.

2.5. Virtual Screening Using Pharmacophore Hypotheses. The best generated pharmacophore hypothesis was used as a 3D query to search new LXR β agonists from the NCI database, which has 260 071 compounds. Lipinski rules were applied to exclude nondrug-like compounds with following criterion: molecular weight <500, $-4 < \log p < 5$, numbers of donor/acceptor <10,

the numbers of rotatable bond <10. The diversified conformations of the remaining structures (180 871 structures) were generated. For each compound, up to 250 conformations were kept. The best pharmacophore was used to screen those qualified compounds by flexibly mapping pharmacophores to the query (the best generated pharmacophore hypothesis) in DS. A number of hits were found based upon given criterion. Docking studies then examined the quality of the virtual screening campaign.

2.6. Evaluating Hits with Docking Studies. The hits that overlapped with the pharmacophore model were docked back to the binding pocket of LXR β for checking the fitnesses (by means of FlexX algorithm). The poses were sorted based on scoring functions. A lower scoring value indicates a more favorable binding. A crystal structure of LXR β -ligand complex (PDB code: 3KFC) was used as the reference structure for hit validations. For each hit, top 10 poses were kept for the further studies.

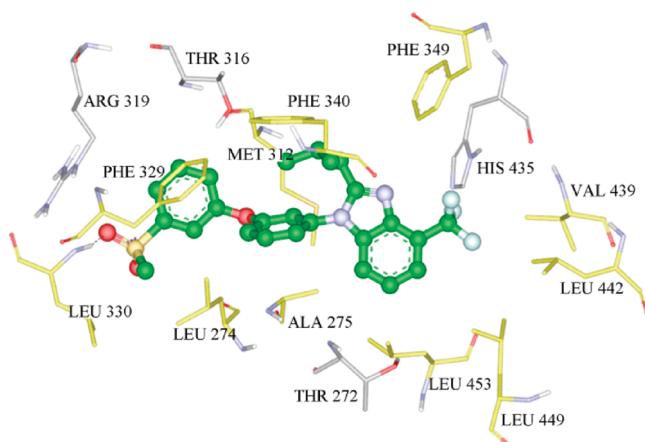


Figure 4. Binding orientation of Compound 1 in the training set. The binding orientation was taken from FlexX pose. Ligands are shown in ball and stick, while active sites are shown in stick. The hydrophobic residues are shown in yellow, and the hydrogen bonds are shown in purple.

3. RESULTS AND DISCUSSION

3.1. Pharmacophore Model Generation. Catalyst HypoGen requires at least 16 diversified compounds in a training set; the activities of these compounds span over four orders of magnitude.³⁰ In this study, the training set consists of 19 compounds, and the IC₅₀ values range from 0.8 to greater than 1000 nM. Top 10 pharmacophore hypotheses were generated. Each pharmacophore model consists of five pharmacophore features. All the models have the features of hydrogen acceptor and hydrophobic, which indicate that hydrogen acceptor and hydrophobic play important roles for the activities of LXR β agonists.

As shown in Table 1, the correlation values of all pharmacophore models were greater than 0.92, which means that all the models are capable of predicting LXR β agonist IC₅₀ if they share the similar scaffolds to the ones in the training set. Besides the high correlations, all of 10 pharmacophore models have a high cost difference (>97), low rmsd (<1.5), high fit value (>8.6), and low error (<75). All these data support the predictive ability of the generated 10 pharmacophore models. Among the 10 generated pharmacophores, the top ranked hypothesis (Hypo 1) contains one hydrogen acceptor and four hydrophobic features. Hypo 1 has the highest correlation coefficient (0.98), the best cost difference (111.31), the lowest errors (61.09), the lowest rmsd (0.79), and a high fit value (10.12). Therefore, Hypo 1 was picked as the final 3D pharmacophore model for further analyses.

Hypotheses were ranked by cost. Three cost values, namely fixed, null, and total costs, were used to assess the quality of generated hypotheses. Every created pharmacophore model represents its own total cost. Fixed cost is the lowest possible cost, it represents ideal hypothetical model. Fixed cost predicts all training set compounds well with lowest deviation. While null cost represents the largest cost of hypothesis with no features, it estimates every activity of the training set compound to be the average activity. The cost difference means the difference between the total and null costs.

The fixed and null cost values were 74.77 and 192.55, respectively. The configuration cost was 17.19 (a little higher than maximum limit of 17). Configuration cost depends on the complexity of the hypothesis space. This may be because the

Table 3. Experimental and Predicted IC₅₀Values of the 41 Test Set Compounds Based on the Pharmacophore Model Hypo 1

compound no.	experimental IC ₅₀ , nM	predicted IC ₅₀ , nM	error ^a	fit value	experimental scale ^b	predicted scale ^b
1	1.2	8.29	6.91	9.09	+++	+++
2	1.3	0.81	-1.60	10.09	+++	+++
3	1.7	1.49	-1.14	9.83	+++	+++
4	3.5	13.85	3.96	8.86	+++	+++
5	3.6	13.85	3.85	8.86	+++	+++
6	3.7	13.85	3.74	8.86	+++	+++
7	4.2	0.65	-6.43	10.19	+++	+++
8	4.7	13.85	2.95	8.86	+++	+++
9	5.1	6.31	1.24	9.20	+++	+++
10	5.4	5.79	1.07	9.24	+++	+++
11	5.6	6.31	1.13	9.20	+++	+++
12	5.7	5.79	1.02	9.24	+++	+++
13	6.5	1.62	-4.01	9.79	+++	+++
14	7	5.79	-1.21	9.24	+++	+++
15	7.5	2.93	-2.56	9.54	+++	+++
16	10	30.12	3.01	8.52	+++	+++
17	13	5.79	-2.25	9.24	+++	+++
18	13	5.79	-2.24	9.24	+++	+++
19	17	13.85	-1.23	8.86	+++	+++
20	20	6.31	-3.17	9.20	+++	+++
21	24	5.79	-4.15	9.24	+++	+++
22	28	20.38	-1.37	8.69	+++	+++
23	30	5.327	-5.63	9.28	+++	+++
24	30	76.23	2.54	8.12	+++	+++
25	33	76.23	2.31	8.12	+++	+++
26	44	35.87	-1.23	8.45	+++	+++
27	49	76.23	1.56	8.12	+++	+++
28	65	38.32	-1.70	8.42	+++	+++
29	69	60.71	-1.14	8.22	+++	+++
30	98	232.57	2.37	7.64	+++	++
31	100	215.49	2.16	7.67	++	++
32	100	148.04	1.48	7.83	++	++
33	140	123.2	-1.14	7.91	++	++
34	170	127.05	-1.34	7.90	++	++
35	188	19.99	-9.40	8.70	++	+++
36	199	70.23	-2.83	8.16	++	+++
37	230	150.93	-1.52	7.83	++	++
38	410	851.37	2.08	7.07	++	++
39	700	349.16	-2.01	7.46	++	++
40	>1000	969.43	1.03	7.02	+	++
41	>1000	3432.46	3.43	6.47	+	+

^a A + indicates that the predicted IC₅₀ value is higher than the experiment IC₅₀ value, a - indicates that the predicted IC₅₀ value is lower than the experiment IC₅₀ value. ^b Activity scale: high active, +++ (IC₅₀ < 100 nM); moderate active, ++ (100 nM ≤ IC₅₀ ≤ 1000 nM); low active, + (IC₅₀ ≥ 1000 nM).

training set compounds increase the entropy of the hypothesis.^{29,31} Other good figures should surpass the drawbacks related to the configuration cost. In Hypo 1, the fixed cost value (74.77) was close to the total cost value (81.24), and the cost difference between total and null costs was found to be 111.31,

which shows the over 90% chance of representing a true correlation data. Furthermore, compounds in the test set and in Fischer's randomization test were also used to validate the reliability of the pharmacophore models. Figure 2 shows the chemical features and the geometric parameters of Hypo 1.

3.2. Predicting IC₅₀ Values for the Compounds from the Training Set Using Hypo 1. Table 2 lists the IC₅₀ values for the compounds in the training set estimated by Hypo 1. The compounds were divided into three categories based on their IC₅₀ values: high active (IC₅₀ < 100 nM, +++)+, moderate active (100 nM ≤ IC₅₀ < 1000 nM), and low active (IC₅₀ ≥ 1000 nM) compounds. All the compounds were correctly estimated and assigned with proper activity levels. Values in column "error" are less than 10, which means the differences between the experimental IC₅₀ and predicted IC₅₀ values are less than one order magnitude. Figure 3 explains why Compound 1 is most active, and Compound 17 is lowest active based upon Hypo 1 pharmacophore model.

Compound 1 (Figure 1) is the most active compound, FlexX algorithm assigned the best score for it, based upon the Hypo 1

model (the FlexX score = -32.8, this score is related to free energy changes). Figure 4 shows the binding mode of compound 1 with the LXR β receptor. Binding analysis showed that Compound 1 formed a hydrogen bond with the backbone N atom of Leu330 and also had hydrophobic interactions with "Phe pocket" (PHE329, PHE340, PHE349). This explained why Compound 1 is the most active.

3.3. Validating Hypo 1. In order to validate the Hypo 1 pharmacophore model, we use it to predict the activities for 41 compounds in the test set. These compounds are structurally diverse (see Supporting Information, Figure S1).

Table 3 lists Hypo 1 validation results from the test set. The activities for the 41 compounds have been predicted. In order to depict the data points that have values of >1000 nM, they were set to 1500 nM. These data points appear as outliers in the scatter plot, as shown in Figure 5. For the test set, the correlation coefficient of estimated and observed data is 0.91. These data show that the predictions are highly correlated with the experimental data; the errors of the predictions are less than 10. Therefore, the quality of the model is acceptable.

In order to further validate the Hypo 1 pharmacophore model, the Fischer's randomization modeling tests were carried out in DS 2.1. This cross-validation approach is to confirm that Hypo 1 pharmacophore model was not generated by chance, and there are strong correlations between the chemical structures and bioactivities. Nineteen random pharmacophore spreadsheets were generated using exactly the same parameters as used in generating the 10 pharmacophore hypotheses, as described in Section 3.1. Figure 6 shows that Hypo 1 is of a superior quality than all the randomly generated hypotheses.

Table 4. Statistical Parameter from Screening Test Set Molecules

parameter	values
total molecules in database (D)	181 051
total number of actives in database (A)	180
total Hits (Ht)	2447
active Hits (Ha)	167
% ratio of actives [(Ha/A) × 100]	92.78
enrichment factor (EF)	68.64

Figure 5. The correlations between estimated and observed IC₅₀ values.

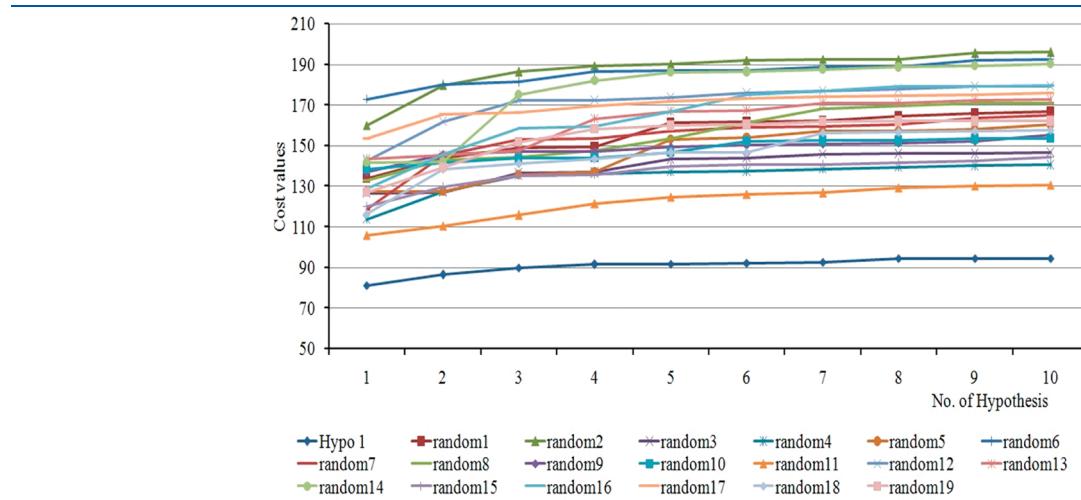


Figure 6. Comparison of Hypo 1 and randomly generated hypotheses. All random models have significantly high cost values than the one for Hypo 1.

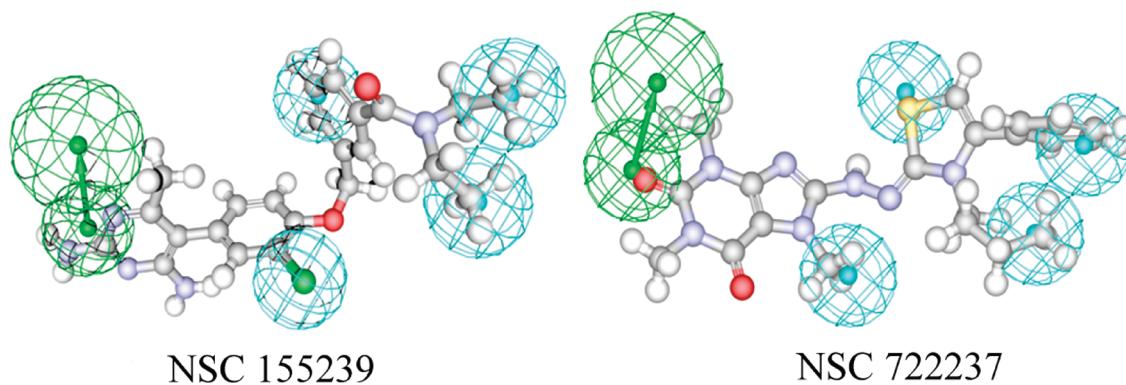


Figure 7. The pharmacophore mapping of NSC155239 and NSC722237 for Hypo 1.

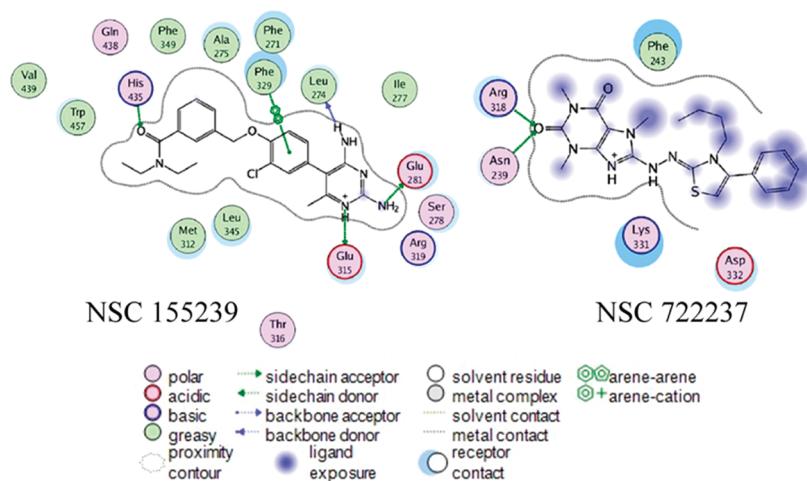


Figure 8. The ligand–receptor interaction map for NSC155239 and NSC722237. The fuzzy blue blob indicates ligand exposure in solvent.

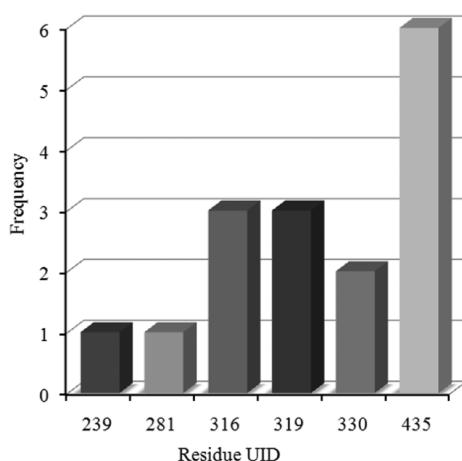


Figure 9. The “hot residues” of LXR β and their numbers of the interactions with the ligands (frequencies) (PDB codes: 3KFC, 1UPV, 1UPW, 1PQ6, 1PQ9, 1PQC, and 1P8W).

3.4. Virtual Screening Against NCI Compounds Using Hypo 1. In order to discover novel LXR β agonists, Hypo 1 was used to screen against the 260 071 compounds in NCI library. Based upon the Lipinski’s rules, we selected 180 871 drug-like compounds from NCI library for virtual screening. The

pharmacophore model Hypo 1 was used to do virtual screening over the NCI library (180 871 compounds) and 180 known LXR β agonists. There were 2447 hits, from which 167 hits were from the 180 known compounds. Hence, the enrichment factor = $[167 \times (180\,871 + 180)] / (2447 \times 180) = 69$, which means that it is 69 times more probable to pick active compounds from the database than expected by chance. Table 4 indicates that the Hypo 1 was not generated by chance.

The 2280 compounds from the NCI library were mapped to Hypo 1 with all feature mappings, and all the IC₅₀ values were predicted. From the 2280 compounds (from NCI library), 147 compounds were selected based upon the fact that their estimated IC₅₀ values were less than 10 nM. In order to increase the confidence of the 147 hits, we docked them back to the LXR β binding pocket in an LXR β crystal structure (PDB code: 3KFC) with the FlexX algorithm. In order to ensure that the FlexX algorithm was able to properly dock ligands to the binding pocket, the ligand derived from the cocrystal structure 3KFC was redocked to the cocrystal structure. The rmsd value between the docked and experimental poses for the ligand was 0.91. This proved that the FlexX was capable of docking compounds to 3KFC properly, and 145 compounds could be successfully docked into the binding pocket.

We randomly selected 140 compounds that were not picked by the Hypo 1 and docked them to 3KFC. For the 145 compounds picked by the Hypo 1, the best compound showed the dock score of

–32.20 (lower value is better), and 14% of them scored lower than –20. For the 140 arbitrary compounds, the best compound showed the dock score of –23.67, and 3.6% of them scored lower than –20. These data show that the Hypo 1 was not generated by chance.

In order to objectively rank the 145 compounds by both the predicted IC_{50} (X) from HypoGen and the docking score (Y) from FlexX, a consensus score (s) was computed from the normalized X and the normalized Y . The normalized X and Y were computed by the following formulas:

$$x'_i = \left(1 - \frac{x_i - \min(X)}{\max(X) - \min(X)}\right) * 100 \quad (1)$$

$$y'_i = \left(1 - \frac{y_i - \min(Y)}{\max(Y) - \min(Y)}\right) * 100 \quad (2)$$

Where, x' and y' stand for the normalized HypoGen predicted IC_{50} and the FlexX docking score, respectively. The consensus score (s_i) was computed in formula 3:

$$s_i = \sqrt{x'_i y'_i} \quad (3)$$

The consensus scores for the 145 compounds are listed in Table S1 (see Supporting Information). The compounds are ranked in the descending order of the consensus score. In order to validate the consensus score, the best and worst compounds in Table S1, Supporting Information were examined with MOE.

The best consensus scored compound is NSC155239, and the worst consensus scored compound is NSC722237. Figure 7 depicts the pharmacophore mappings of NSC155239 and NSC722237 for Hypo 1. The ligand–receptor interaction maps for NSC155239 and NSC722237 are depicted in Figure 8.

Both NSC155239 and NSC722237 matched Hypo 1 pharmacophore model well. No significant pharmacophore difference between NSC155239 and NSC722237 could be recognized (Figure 7). At this time, the pharmacophore model could not explain why NSC155239 and NSC722237 have different activities. On the other hand, it was clearly recognized that NSC155239 and NSC722237 interacted with LXR β binding pocket differently. NSC155239 interacted with many “hot” residues in the pocket (Figure 8). However, NSC722237 only interacted with five residues, and most of it was exposed to the solvent (Figure 8).

In order to ensure NSC155239 did interact with the “hot” residues, seven cocrystal structures of LXR β (reported in RCSB Protein Data Bank) were superimposed. By summarizing the interactions between the receptor residues and the ligand in the cocrystal structures, Asn239, Glu281, Thr316, Arg319, Leu330, and His435 were found to be the most frequently bound with the ligands. Most of them could form hydrogen bonds with the ligands. His435 was particularly crucial for LXR β binding.³² Figure 9 depicts the numbers of the interactions for all the “hot” residues in LXR β . The LXR β binding pocket consists of hydrogen binding area and hydrophobic area (“Phe pocket”), which is formed by Phe239, Phe340, and Phe349.

NSC155239 formed hydrogen bonds with four residues: His435, Leu274, Glu281, and Glu315 in LXR β . NSC155239 also formed π – π interaction with Phe329, in the “Phe pocket”. NSC 722237 had only weak interactions with five residues in the LXR β pocket. These differences explain the reason why NSC155239 was ranked as the most potent compound and why NSC 722237 was the least potent compound.

4. CONCLUSIONS

In this work, 3D QSAR pharmacophore models were generated from 19 LXR β agonists in the training set. The best pharmacophore model Hypo 1 was validated with the test set, a virtual screening experiment, and the FlexX docking approach. The pharmacophore model is capable of predicting the IC_{50} values of the agonists with similar scaffolds. The ligand-based pharmacophore model can be used for virtual screening. However, it would be better to validate the hits with the structure-based docking approach. By combining ligand- and structure-based approaches, we expect better results for novel lead identification campaigns.

■ ASSOCIATED CONTENT

S Supporting Information. The structures of 41 test set compounds (Figure S1), and the consensus scores of the 145 compounds from NCI library (Table S1). This material is available free of charge via the Internet at <http://pubs.acs.org>.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: (J.X.) xujun9@mail.sysu.edu.cn; (J.L.) jiaboli@yahoo.com.

■ ACKNOWLEDGMENT

This work has funded in part of the Major scientific and technological special project by the ministry of Science and Technology of China (2010ZX09102-305), Guangdong Recruitment Program of Creative Research Groups, and the Fundamental Research Funds for the Central Universities.

■ REFERENCES

- (1) Governments Ignore the World's Leading Cause of Death; World Heart Federation: Geneva, Switzerland http://www.world-heart-federation.org/fileadmin/user_upload/documents/press-releases-esc-4sept.pdf. Accessed September 4, 2005.
- (2) Spreafico, M.; Smiesko, M.; Peristera, O.; Rossato, G.; Vedani, A. Probing Small-Molecule Binding to the Liver-X Receptor: A Mixed-Model QSAR Study. *Mol. Inf.* **2010**, *29*, 27–36.
- (3) Yang, C. D.; McDonald, J. G.; Patel, A.; Zhang, Y.; Umetani, M.; Xu, F.; Westover, E. J.; Covey, D. F.; Mangelsdorf, D. J.; Cohen, J. C.; Hobbs, H. H. Sterol Intermediates from Cholesterol Biosynthetic Pathway as Liver X Receptor Ligands. *J. Biol. Chem.* **2006**, *281*, 27816–27826.
- (4) Whitney, K. D.; Watson, M. A.; Collins, J. L.; Benson, W. G.; Stone, T. M.; Numerick, M. J.; Tippin, T. K.; Wilson, J. G.; Winegar, D. A.; Kliewer, S. A. Regulation of Cholesterol Homeostasis by the Liver X Receptors in the Central Nervous System. *Mol. Endocrinol.* **2002**, *16*, 1378–1385.
- (5) Yang, C. D.; McDonald, J. G.; Cohen, J. C.; Hobbs, H. H. Sterol Intermediates from Cholesterol Biosynthetic Pathway as Liver X Receptor Ligands. *FASEB J.* **2007**, *21*, A979–A979.
- (6) Zhao, C. Y.; Dahlman-Wright, K. Liver X Receptor in Cholesterol Metabolism. *J. Endocrinol.* **2010**, *204*, 233–240.
- (7) Janowski, B. A.; Willy, P. J.; Devi, T. R.; Falck, J. R.; Mangelsdorf, D. J. An Oxysterol Signalling Pathway Mediated by the Nuclear Receptor LXR Alpha. *Nature* **1996**, *383*, 728–731.
- (8) Fan, X. K.; H., J.; Bouton, D.; Warner, M.; Gustafsson, J. A. Expression of Liver X Receptor Beta is Essential for Formation of Superficial Cortical Layers and Migration of Later-Born Neurons. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 13445–13450.

- (9) Calkin, A. C.; T., P. Liver X Receptor Signaling Pathways and Atherosclerosis. *Arterioscler, Thromb, Vasc. Biol.* **2010**, *30*, 1513–1518.
- (10) Repa, J. J.; Turley, S. D.; Lobaccaro, J. M. A.; Medina, J.; Li, L.; Lustig, K.; Shan, B.; Heyman, R. A.; Dietschy, J. M.; Mangelsdorf, D. J. Regulation of Absorption and ABC1-Mediated Efflux of Cholesterol by RXR Heterodimers. *Science* **2000**, *289*, 1524–1529.
- (11) Repa, J. J.; Liang, G. S.; Ou, J. F.; Bashmakov, Y.; Lobaccaro, J. M. A.; Shimomura, I.; Shan, B.; Brown, M. S.; Goldstein, J. L.; Mangelsdorf, D. J. Regulation of Mouse Sterol Regulatory Element-Binding Protein-1c Gene (SREBP-1c) by Oxysterol Receptors, LXR Alpha and LXR Beta. *Genes & Dev.* **2000**, *14*, 2819–2830.
- (12) Wang, X.; Collins, H. L.; Ranalletta, M.; Fukui, I. V.; Billheimer, J. T.; Rothblat, G. H.; Tall, A. R.; Rader, D. J. Macrophage ABCA1 and ABCG1, But not SR-BI, Promote Macrophage Reverse Cholesterol Transport in Vivo. *J. Clin. Invest.* **2007**, *117*, 2216–2224.
- (13) Brown, M. S.; Goldstein, J. L. The SREBP Pathway: Regulation of Cholesterol Metabolism by Proteolysis of a Membrane-Bound Transcription Factor. *Cell* **1997**, *89*, 331–340.
- (14) Bradley, M. N.; Hong, C.; Chen, M.; Joseph, S. B.; Wilpitz, D. C.; Wang, X.; Lusis, A. J.; Collins, A.; Hseuh, W. A.; Collins, J. L.; Tangirala, R. K.; Tontonoz, P. Ligand Activation of LXR Beta Reverses Atherosclerosis and Cellular Cholesterol Overload in Mice Lacking LXR Alpha and ApoE. *J. Clin. Invest.* **2007**, *117*, 2337–2346.
- (15) Travins, J. M.; Bernotas, R. C.; Kaufman, D. H.; Quinet, E.; Nambi, P.; Feingold, I.; Huselton, C.; Wilhelmsson, A.; Goos-Nilsson, A.; Wrobel, J. 1-(3-Aryloxyaryl) Benzimidazole Sulfones are Liver X Receptor Agonists. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 526–530.
- (16) Singhaus, R. R.; Bernotas, R. C.; Steffan, R.; Matelan, E.; Quinet, E.; Nambi, P.; Feingold, I.; Huselton, C.; Wilhelmsson, A.; Goos-Nilsson, A.; Wrobel, J. 3-(3-Aryloxyaryl) Imidazo[1,2-a] Pyridine Sulfones as Liver X Receptor Agonists. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 521–525.
- (17) Szewczyk, J. W.; Huang, S.; Chin, J.; Tian, J.; Mitnal, L.; Rosa, R. L.; Peterson, L.; Sparrow, C. P.; Adams, A. D. SAR Studies: Designing Potent and Selective LXR Agonists. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 3055–3060.
- (18) Bernotas, R. C.; Kaufman, D. H.; Singhaus, R. R.; Ullrich, J.; Unwalla, R.; Quinet, E.; Nambi, P.; Wilhelmsson, A.; Goos-Nilsson, A.; Wrobel, J. 4-(3-Aryloxyaryl) Quinoline Alcohols are Liver X Receptor Agonists. *Bioorg. Med. Chem.* **2009**, *17*, 8086–8092.
- (19) Ullrich, J. W.; Morris, R.; Bernotas, R. C.; Travins, J. M.; Jetter, J.; Unwalla, R.; Quinet, E.; Nambi, P.; Feingold, I.; Huselton, C.; Enroth, C.; Wilhelmsson, A.; Goos-Nilsson, A.; Wrobel, J. Synthesis of 4-(3-Biaryl) Quinoline Sulfones as Potent Liver X Receptor Agonists. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 2903–2907.
- (20) Bernotas, R. C.; Singhaus, R. R.; Kaufman, D. H.; Ullrich, J.; Fletcher, H.; Quinet, E.; Nambi, P.; Unwalla, R.; Wilhelmsson, A.; Goos-Nilsson, A.; Farnegardh, M.; Wrobel, J. Biarylether Amide Quinolines as Liver X Receptor Agonists. *Bioorg. Med. Chem.* **2009**, *17*, 1663–1670.
- (21) Wrobel, J.; Steffan, R.; Bowen, S. M.; Magolda, R.; Matelan, E.; Unwalla, R.; Basso, M.; Clerin, V.; Gardell, S. J.; Nambi, P.; Quinet, E.; Remnick, J. I.; Vlasuk, G. P.; Wang, S.; Feingold, I.; Huselton, C.; Bonn, T.; Farnegardh, M.; Hansson, T.; Nilsson, A. G.; Wilhelmsson, A.; Zamaratski, E.; Evans, M. J. Indazole-Based Liver X Receptor (LXR) Modulators with Maintained Atherosclerotic Lesion Reduction Activity but Diminished Stimulation of Hepatic Triglyceride Synthesis. *J. Med. Chem.* **2008**, *51*, 7161–7168.
- (22) Li, H. F.; Lu, T.; Zhu, T.; Jiang, Y. J.; Rao, S. S.; Hu, L. Y.; Xin, B. T.; Chen, Y. D. Virtual Screening for Raf-1 Kinase Inhibitors Based on Pharmacophore Model of Substituted Ureas. *Eur. J. Med. Chem.* **2009**, *44*, 1240–1249.
- (23) Sakkiah, S.; Thangapandian, S.; John, S.; Kwon, Y. J.; Lee, K. W. 3D QSAR Pharmacophore Based Virtual Screening and Molecular Docking for Identification of Potential HSP90 Inhibitors. *Eur. J. Med. Chem.* **2010**, *45*, 2132–2140.
- (24) Sundarapandian, T.; Shalini, J.; Sugunadevi, S.; Woo, L. K. Docking-Enabled Pharmacophore Model for Histone Deacetylase 8 Inhibitors and Its Application in Anti-Cancer Drug Discovery. *J. Mol. Graphics Modell.* **2010**, *29*, 382–395.
- (25) Yankeelov, J. A., Jr.; Koshland, D. E., Jr. Evidence for Conformation Changes Induced by Substrates of Phosphoglucomutase. *J. Biol. Chem.* **1965**, *240*, 1593–602.
- (26) Kirchmair, J.; Laggner, C.; Wolber, G.; Langer, T. Comparative Analysis of Protein-Bound Ligand Conformations with Respect to Catalyst's Conformational Space Subsampling Algorithms. *J. Chem. Inf. Model.* **2005**, *45*, 422–30.
- (27) Bernotas, R. C.; Singhaus, R. R.; Kaufman, D. H.; Travins, J. M.; Ullrich, J. W.; Unwalla, R.; Quinet, E.; Evans, M.; Nambi, P.; Olland, A.; Kauppi, B.; Wilhelmsson, A.; Goos-Nilsson, A.; Wrobel, J. 4-(3-Aryloxyaryl) Quinoline Sulfones are Potent Liver X Receptor Agonists. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 209–212.
- (28) Rarey, M. K., B.; Lengauer, T.; Klebe, G. A Fast Flexible Docking Method Using an Incremental Construction Algorithm. *J. Mol. Biol.* **1996**, *261*, 470–489.
- (29) *Discovery Studio*, version 2.1; Accelrys Inc: San Diego, CA, 2007.
- (30) Kurogi, Y.; Guner, O. F. Pharmacophore Modeling and Three-Dimensional Database Searching for Drug Design Using Catalyst. *Curr. Med. Chem.* **2001**, *8*, 1035–1055.
- (31) Aparoy, P.; Kumar Reddy, K.; Kalangi, S. K.; Chandramohan Reddy, T.; Reddanna, P. Pharmacophore Modeling and Virtual Screening for Designing Potential 5-Lipoxygenase Inhibitors. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 1013–1018.
- (32) Farnegardh, M.; Bonn, T.; Sun, S.; Ljunggren, J.; Ahola, H.; Wilhelmsson, A.; Gustafsson, J. A.; Carlquist, M. The Three-Dimensional Structure of the Liver X Receptor Beta Reveals a Flexible Ligand-Binding Pocket that Can Accommodate Fundamentally Different Ligands. *J. Biol. Chem.* **2003**, *278*, 38821–38828.