JOURNAL OF CHEMICAL INFORMATION AND MODELING

ARTICLE

pubs.acs.org/jcim

Virtual Screening for Ligands of the Insect Molting Hormone Receptor

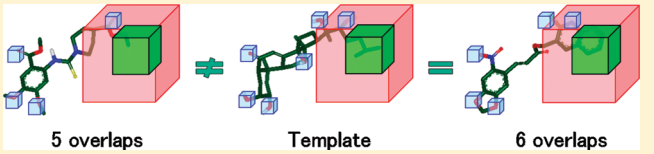
Toshiyuki Harada,[†] Yoshiaki Nakagawa,[*,†] Takehiko Ogura,[†] Yutaka Yamada,[‡] Takehiro Ohe,[§] and Hisashi Miyagawa[†]

[†]Graduate School of Agriculture, Division of Applied Life Sciences, Kyoto University, Sakyo-ku, Kyoto, 606-8502, Japan

[‡]Cybernet Systems Co., Ltd., 3 Kanda-Neribeicho, Chiyoda-ku, Tokyo, 101-0022, Japan

[§]OpenEye Japan, Inc., AIOS Toranomon 904, 1-6-12, Nishishinbashi, Minato-ku, Tokyo 105-0003, Japan

ABSTRACT: Insect growth is regulated by the orchestrated event of ecdysteroids and their receptor proteins. Agonists/antagonists of ecdysteroid receptor are predicted to disrupt normal growth, providing good candidates of new insecticides. A database of over 2 million compounds was subjected to a shape-based virtual screening cascade to identify novel nonsteroidal hits similar to the known EcR ligand ponasterone A. Testing revealed micromolar hits against two strains of insect cells. Docking experiments against EcR were used to support the predicted binding mode of these ligands based on their overlay to ponasterone A.



1. INTRODUCTION

Insects grow by repeated processes of molting and metamorphosis, which are orchestrated by two peripheral nonpeptide hormones, molting and juvenile hormones. The molting hormone of most insects is 20-hydroxyecdysone (20E; I in Figure 1). The action of molting hormones is mediated by a heterodimeric protein complex composed of the ecdysone receptor (EcR) and the ultraspiracle (USP; a homologue of retinoid X receptor, RXR). This complex activates the translation of the molting associated genes after binding of a ligand molecule to EcRs.[1] Even though vertebrates have similar gene transcription mechanisms regulated by steroid hormones, ecdysteroids do not perturb these systems. Therefore, this property has made ecdysteroids attractive candidates for insecticides with a good safety profile in mammals including human beings. However, since steroidal compounds have a complicated core structure and poor hydrophobicity, they are usually deemed unfavorable to be used as insecticides, and the nonsteroidal compounds with molting hormone activity have been sought after.

About two decades ago, a series of *N*-*t*-butyl-*N*,*N*′-dibenzoylhydrazine (DBH; II in Figure 1) were serendipitously discovered as a nonsteroidal molting hormone agonist from synthetic byproducts and attracted attention as insecticidal small molecules with new mode-of-action.[2] Later, several other nonsteroidal types of compounds such as 3,5-di-*t*-butyl-4-hydroxy-benzamide,[3] tetrahydroquinolines,[4] and γ-methylene γ-lactam type compounds[5] were randomly screened from a number of chemicals. Of these, DBHs had an excellent controlling activity toward lepidopteran insect pests, and four compounds of this type (tebufenozide, methoxyfenozide, chromafenozide, and halofenozide) have been commercialized for agricultural use to date, especially targeting for armyworm and leaf roller. Compared to the high toxicity of DBHs toward these target insects, however, their toxicity is weak or even inactive against other taxonomic orders of insects such as Diptera and Hemiptera. This has been deemed curious, given the molting processes of all the insects are commonly regulated by a steroidal hormone, 20E. Interestingly, similar species-selective activity has also been observed for other nonsteroidal agonists like tetrahydroquinoline derivatives, and thus, much effort has been directed toward understanding the molecular basis for such a selective toxicity of synthetic molting hormone agonists between insect species.[6]

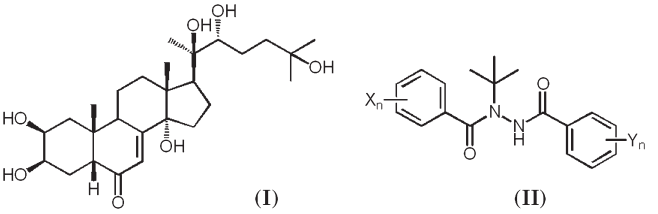


Figure 1. Chemical structures of 20-hydroxyecdysone (**I**) and dibenzoylhydrazine (**II**).

While the difference in the hormonal activity between steroidal and nonsteroidal ecdysone agonists has been implicated in the difference in their binding mode to EcR, a structural evidence was obtained by a crystal structure-based comparative studies of ligand-bound EcR complexes, using ponasterone A (PonA; one of the most potent ecdysteroids) and a DBH analogue, BYI-06830.[7] According to the elucidated structures, the binding pockets of the EcR from *Heliothis virescens* (HvEcR) for PonA and DBH were overlapped only partially, as shown in Figure 2, and the terminal *i*-propyl moiety of the side chain of PonA shared the same space as the *t*-butyl moiety of DBH. By contrast, the ring moiety of DBH away from *t*-butyl group was found to be accommodated in an extra pocket that was not occupied by PonA. Therefore, it was concluded that the structural difference of this extra pocket among the insect species can be responsible

Received: October 13, 2010
Published: January 28, 2011

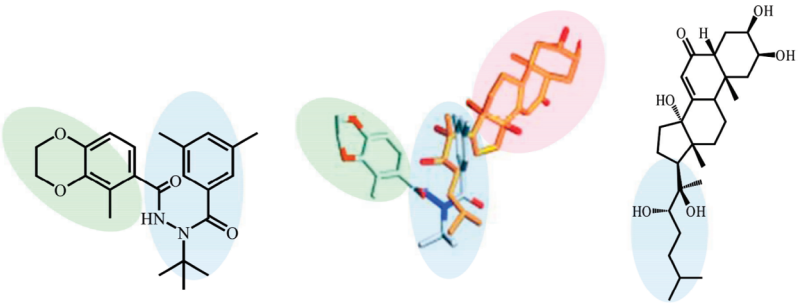


Figure 2. Overlap between structures of BYI-06830 and ponasterone A bound to HvEcR.

for the selectivity of DBHs or their restricted use as insecticides due to the relatively narrow spectrum of target insect pests.

Given the DBH-specific interaction with EcR accounting for their selective toxicity, it is reasonable to assume that a ligand that shares the same binding pocket as ecdysteroids should have a molting hormone agonist activity toward broad range of insects. The design of such a ligand is most likely achieved by mimicking the structures of ecdysteroids. In this regard, we previously synthesized a series of PonA analogues with various steroidal moieties and measured their ligand-receptor binding activities in *Drosophila* Kc cells,[8] in order to examine the factors that affect the interaction between EcR and a steroidal ligand. A 3-D QSAR study revealed that the binding activity of the steroid analogues was mainly dependent on the number of hydrogen bonds (HBs) formed between ligand and receptor,[9] and other physicochemical properties of a ligand such as logP and electrostatic nature affect the binding only slightly if any. It has been shown that the definite space arrangement of a hydrophobic substructure in the ligands is required for the effective interaction with the receptor and a set of hydrogen bond donors/acceptors.[9,10] Based on this result, it was suggested that not only mimicking the shape of steroidal skeleton but also assembling as many properly arranged HB donors/acceptors as possible in a molecule is important for designing a molting hormone agonists having a comparative activity to ecdysteroids.

In this study, we report the first application of a computer-aided drug design approach to discover a novel ecdysone agonist with nonsteroidal scaffold. Then we screened new ligands that interact with EcR from a chemical database comprising of 2.1 million compounds, using a technique of virtual screening. Virtual screening is one of emerging approaches for efficient drug discovery, in which likely active compounds are selected based on the physicochemical properties and/or the molecular shape presumably favorable for the interaction with the receptor, without actual bioassay of activity. The technique mainly consists of two methods: one is based on the similarity to a known ligand structure, referred to as ligand-based virtual screening (LBVS); the other is based on the fitting of a molecule to the receptor pocket *in silico*, referred to as high throughput docking (HTD). While HTD is useful in finding ligands of high novelty, especially in the case where no known ligand is present for a receptor, it is rather troublesome and time-consuming, requiring precise structural data of the receptor and long computational time. Meanwhile, LBVS is based on a simpler algorithm, and generally more advantageous in terms of speed and accuracy of calculation when a known ligand for a receptor is available. For this reason, we applied LBVS here, based on the 3-D structure of PonA in the bound complex with EcR. In expecting to find new nonsteroidal ecdysone agonists with a different action from the hitherto known nonsteroidal agonists, we put a high priority on the similarity in terms of shape and the capacity of HB formation with the query compound (PonA) in the screening process, as indicated by the above-mentioned SAR study for the steroidal agonists. The selected compounds were submitted to a binding assay using two insect cell lines, and the active compounds identified were further analyzed for validating their possible binding mode to the ligand binding pocket of EcR by the in silico docking experiments.

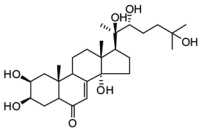
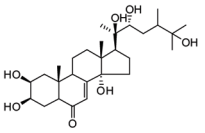
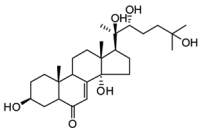
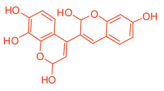
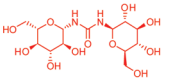
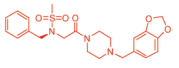
2. RESULTS

2.1. Virtual Screening. A database of 2.1 million compounds was virtually screened for similarity to the structure of PonA with a conformation in the crystalline state, in terms of shape and chemical property using the OpenEye tool ROCS. The compounds were ranked by the sum of their Shape Tanimoto coefficients and Scaled Color scores (Table 1). Unsurprisingly, many of the top ranked compounds contained a steroid skeleton, so we lowered the threshold at this stage for a sufficient level of structure diversity to select a total of 20,000 compounds.

In the subsequent step, more specific interactions that may operate in the complex formation between ecdysteroids and EcR were taken into consideration. It has been shown that the definite space arrangement of a set of hydrogen bond donors/acceptors and a hydrophobic substructure in the ligands is required for the effective interaction with the receptor.[9,10] Thus, we developed an in-house program to count the number of relevant functional groups and the hydrophobic moiety in a molecule, which can be overlapped with the key molecular parts of PonA interacting with EcR. By judging the count of more than 6 to be promising, a total of 237 out of 20,000 compounds were selected. Finally, we narrowed down the selected compounds to 24 for the ligand binding assay by visual inspection, based on 1) chemical diversity, 2) the adequacy of the overlap between template and query structure, and 3) the availability. Structures of the 24 selected compounds are listed in Figure 3.

2.2. Binding Assay. The activity of 24 compounds was evaluated by measuring the inhibition of the incorporation of [^{3}H] PonA into Lepidoptera Sf-9 and Coleoptera BCIRL-Lepd-SL1 cells according to the conventional methods.[11,12] As shown in Figure 4, three compounds (**5**, **12**, **24**) significantly inhibited the incorporation of [^{3}H] PonA (>50%) in both insect cells at their highest concentration (250 μM). For these three potent compounds, the concentration required to give 50% inhibition (IC_{50} in M) was determined from each concentration-response

Table 1. Ranking of the Screened Compounds by the Sum of Shape and Chemical Similarity

Rank	Structures	Shape	Chemical	Σ
1		0.945	0.934	1.878
2		0.931	0.934	1.865
3		0.908	0.786	1.694
•	•	•	•	•
•	•	•	•	•
•	•	•	•	•
16		0.670	0.518	1.188
•	•	•	•	•
•	•	•	•	•
•	•	•	•	•
19		0.628	0.547	1.175
•	•	•	•	•
•	•	•	•	•
•	•	•	•	•
20000		0.651	0.264	0.914

curve (Figure 5). Only one measurement was taken for compound **5** from the BCIRL-Lepd-SL1 cell line, which showed approximately 50% inhibition at 250 μM. This value was therefore used as the IC_{50}. The obtained IC_{50} values are shown in Table 2 with the structures of the compounds.

2.3. Reporter Gene Assay. Three compounds **5**, **12**, and **24** that showed a significant binding activity to EcR were further assayed for their ability to activate the transcription of an ecdysteroid-response element in the promoter of *Drosophila* hsp27 gene. This gene was fused to *Bombyx mori*-derived basal actin promoter, which was followed by a firefly luciferase gene and a termination signal, in order to be transfected in a lepidopteran insect cell line, Sf-9. As shown in Figure 6, none of these three compounds induced luciferase activity even at 100 μM, suggesting that they all bind to EcR in an antagonistic manner.

2.4. Construction and Selection of 3-D Structures of EcR. The primary sequence of the LBD of *Spodoptera frugiperda* EcR (SfEcR) was aligned with that of HvEcR using RPS-BLAST, as shown in Figure 7. The sequence identity between SfEcR and HvEcR was calculated to be 90.6%. Based on this similarity, the 3-D structure of the ligand-binding domain in SfEcR was constructed using a technique of homology modeling. In order to take the flexibility of protein structure upon actual interaction with ligands into consideration, we repeated the simulated annealing to generate multiple models that may cover a range of possible structures arising from a wobble of the protein molecule, and then, the docking was performed for each of the models. In practice, eight representative structures out of 100 were selected to reduce the processing time for docking. The rmsd values of heavy atoms (C,N,O,S) present in the surface of the binding pocket for each pair of these eight structures ranged from 0.368 to 0.602 Å.

2.5. Docking Simulation. To validate that the ligands selected on the basis of the ROCS overlay with PonA in crystals can bind to EcR in the same manner as PonA, we performed in silico docking experiments for the above-described 24 compounds toward each of the eight EcR structures generated in the previous section, using two docking softwares, GOLD and FRED. Three scoring functions (Goldscore, Chemscore[13] and ASP functions[14]) for GOLD and four scoring functions (Shapegauss,[15] PLP,[16] Chemgauss2, Chemgauss3) for FRED were used to evaluate the ligand−receptor affinity. Since these scoring functions respectively provided different ranges of values, the score was expressed by the ratio of the original value to that of PonA in each series for normalization, and then, the sum of the three (GOLD) or four (FRED) normalized scores were deemed as the affinity level of a compound to be ranked as shown in Table 3. Consequently, compound **24** that had the most potent activity in the receptor binding assay was ranked at fourth (GOLD) and at second (FRED), while the second most active compound **12** was

dx.doi.org/10.1021/ci100400k |*J. Chem. Inf. Model.* 2011, 51, 296–305

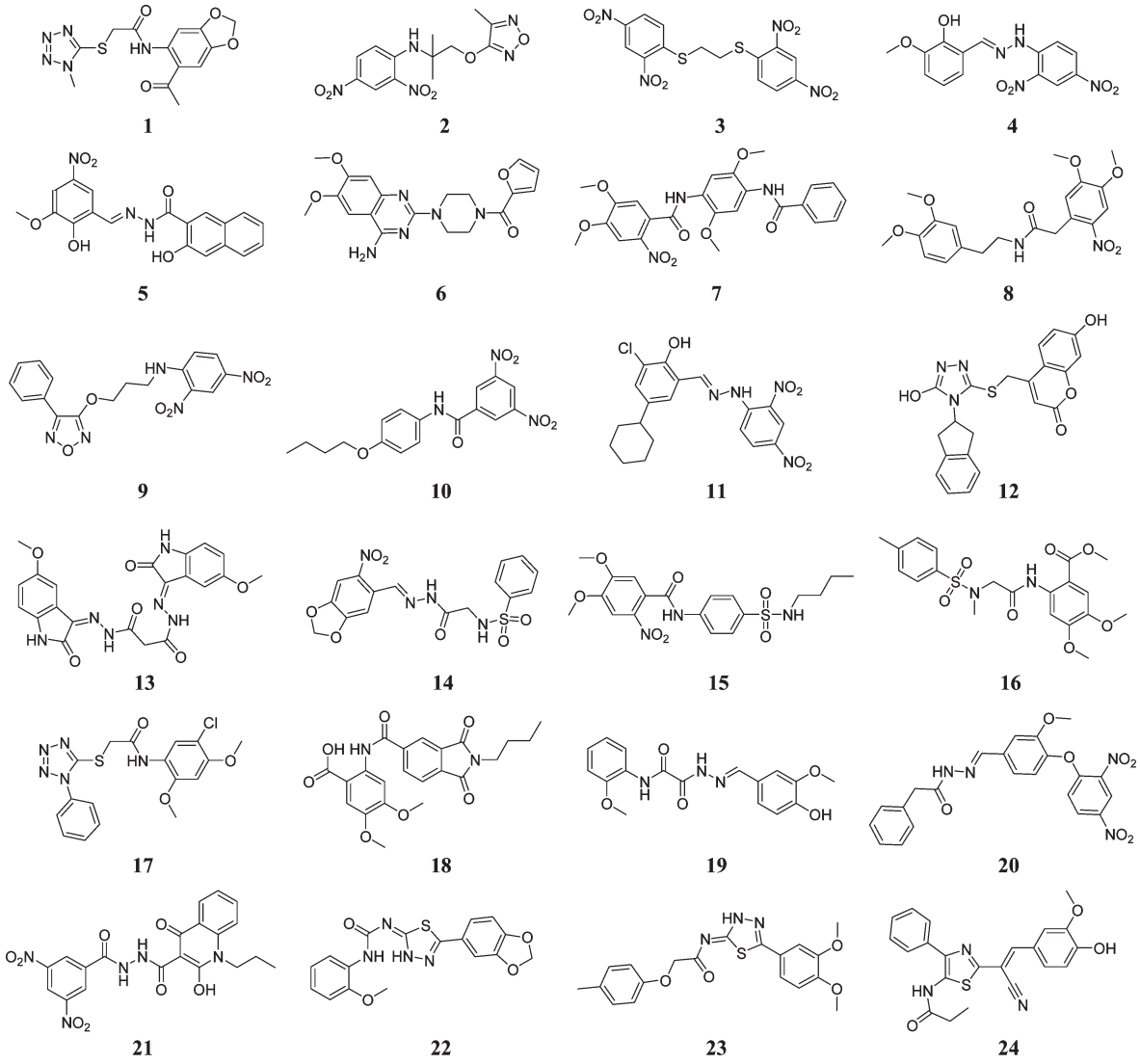


Figure 3. Structures of compounds (**1**−**24**) selected by virtual screening.

at the second (GOLD) and at first (FRED). In both docking, the sum of scores for PonA was larger than any of the selected compounds, and this result is consistent with the fact that PonA has several orders of higher binding activity.

The docking simulation models against SfEcR ligand binding domain for compounds **12** and **24** using Goldscore function are shown in Figure 8, together with that for PonA. The hydrophobic moieties in **12** and **24** are placed in the same cavity as that accommodates the hydrophobic side chain of PonA. The numbers of hydrogen bonds formed in the ligand-receptor complex, which has been thought to be an important factor in determining the activity, amounted to four and three for the cases of **12** and **24**, respectively, while a total of five hydrogen bonds was found in the complex of PonA with receptor (Figure 8). This difference is likely to account, at least in part, for the lower affinity of **12** and **24** compared to PonA.

3. DISCUSSION AND CONCLUSIONS

Three novel ligands that bind to EcR were discovered by LBVS. The inhibitory activity was not high compared to the natural molting hormone (20-hydroxyecdysone) and other nonsteroidal ecdysone agonists; however, the binding potencies for the screened compounds were similar between the receptors of two different insect orders: Lepidoptera and Coleoptera. Accordingly, while not proof of specificity at these high concentrations, these results suggest a similar interaction of the ligand with its receptor, which implies that there is some degree of specificity in the binding. These results were consistent with our original hypothesis that a molecule mimicking the whole molecular structure of an ecdysteroid will possess relatively nonselective activity, as suggested by the similar SARs of ecdysteroids among different insect species. IC_{50} values of the most potent compound **24** against the cells of Sf-9 and BCIRL-Lepd-SL1 are close to the values of ecdysone, a biosynthetic precursor of 20E.[11,17] While compounds **12** and **24** inhibited the binding of PonA at around 10^{-5} M, both of them showed no agonistic activity at 10^{-4} M in the reporter gene assay (Figure 6), suggesting that these EcR ligands are antagonists, or act on receptors by a novel mechanism. Since the structural factors of a molecule that determines whether it acts as agonist or antagonist are generally unclear to date, the ligands newly found in this study may provide interesting materials to study those factors.

The method of virtual screening employed in this study was based only on the structure of ligands, and the structure of

dx.doi.org/10.1021/ci100400k |*J. Chem. Inf. Model.* 2011, 51, 296–305

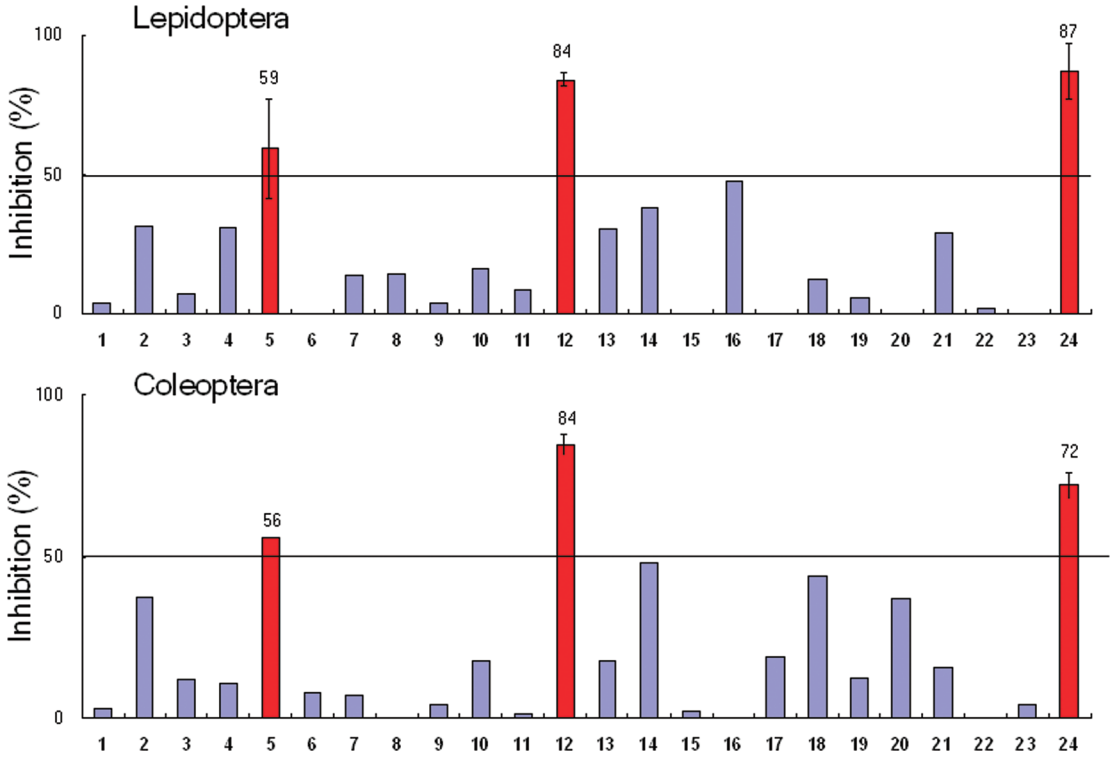


Figure 4. Inhibition of the incorporation of [^{3}H]PonA by selected compounds (**1**-**24**) to (A) Sf-9 and (B) BCIRL-Lepd-SL1 cells. Vertical axis indicates % relative to PonA. Numbers on the horizontal axis correspond to the compound numbers given in Figure 1. Insect cells were incubated for 30 min with [^{3}H]PonA and each test compound at 250 μM or 80 μM. Cells were collected by filtration through a glass filter and the radioactivity of the glass filter was counted by LSC.

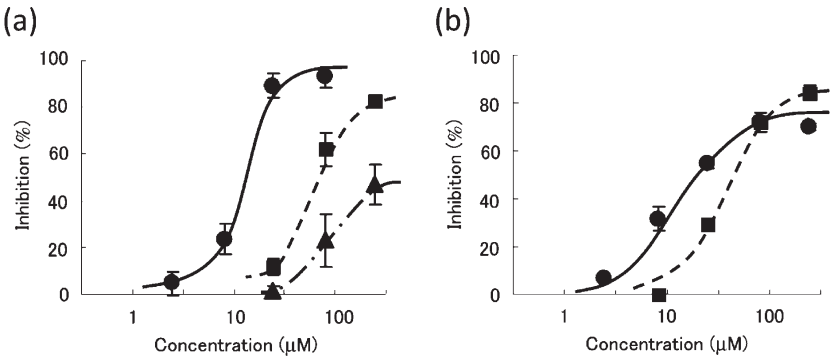
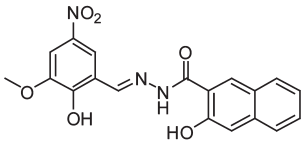
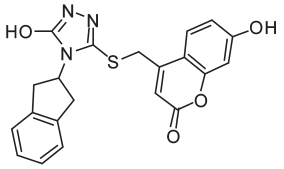
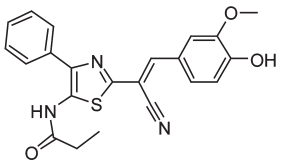


Figure 5. Concentration−response curves for inhibition of the incorporation of [^{3}H]PonA by compounds, **5** (▲), **12** (■), and **24** (●) to Sf9 (a) and BCIRL-Lepd-SL1 (b) cells.

binding site in the receptor was not taken into account for simplicity. So we validated whether the selected compounds are able to bind to the EcR-LBD pocket in a reasonable manner by docking study. In the docking, we constructed the model of SfEcR-LBD based on the crystal structure of HvEcR, in order to correlate the interaction of ligands with the receptor with their binding activity against Sf-9 cells. To evaluate the stability of a docking model, we developed a consensus scoring approach that may not depend on any particular tendency of individual scoring function.[18−21] Namely, the sum of normalized scores given by three types of built-in scoring functions of GOLD (Goldscore, Chemscore, and ASP scoring functions) or four types of built-in scoring functions of FRED (Shapegauss, PLP, Chemgauss2, Chemgauss3) was employed as the index for ranking the selected compounds, in order to reduce the number of false positives generated by each of the scoring functions. The obtained ranking results were reasonable on the whole as described in the Results section, although the relative ranks did not necessarily reflect the actual difference in the activity of the compounds: the rank of weakly active compound **5** was not so high (at eighth in GOLD, at 16th in FRED), while inactive compound **23** was ranked at relatively higher positions in both GOLD and FRED calculations. Another inactive compound **13** was predicted to be the highest in GOLD docking analysis, whereas it is ranked lower in FRED analysis. The structure of weakly active compound **5** appears rather similar to that of a well-known ecdysone agonist, DBH, which was shown to be accommodated in EcR-LBP in a different manner from PonA. Thus, it is likely that this hit is false

Table 2. Structures of the Three Potent Compounds and Their Binding Activity to Insect Cells and Similarity Index Values

Compound		IC_{50} (μM)		Similarity	
No.	Structure	Lepidoptera	Coleoptera	Shape	Chemical
5		250 [a)]	250[a)]	0.649	0.361
12		74	55	0.587	0.405
24		13	32	0.564	0.364

[a] The 50% inhibition was observed at 250 μM.

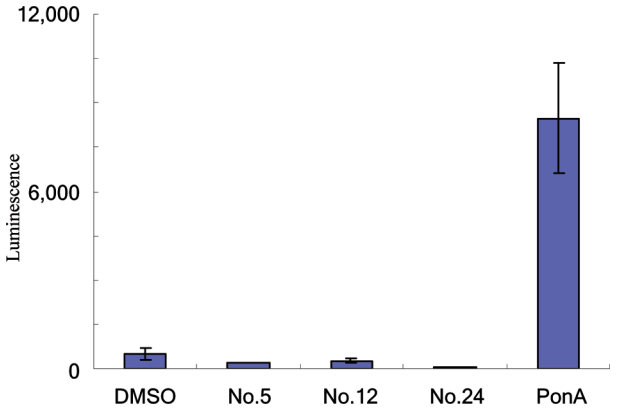


Figure 6. Molting hormonal activity evaluated by the reporter gene assay. The vertical axis is the chemical luminescence. Compounds **5**, **12**, and **24** were assayed at 100 μM, and PonA was assayed at 100 nM.

positive in the present LBVS. On the other hand, the high scores for inactive compounds **13** and **23** are probably related to their relatively high structural flexibility, which often causes an overestimation of binding affinity in the docking calculations by generating a rather unrealistic molecular conformation to fit the binding pocket.

The conformation of PonA in the modeled SfEcR complex after the docking was almost identical with that in the crystallized complex with HvEcR (data not shown). In the docking model of the complex with SfEcR-LBD, PonA formed five HBs with the amino acid residues Glu23 (two HBs), Thr58, Ala113, and Tyr123 (Figure 8). Compounds **12** and **24** also formed four and three HBs, respectively, in the docking models, although the amino acid residues involved in the HB formation were different from those in the complex with PonA: Arg102, Leu111, Ala113, and Asn219 for the case of **12**, and Thr58, Leu111, and Ala113 for the case of **24**. No HBs corresponding to those formed for 2- and 3-OH groups of PonA were found in the complexes of **12** or **24**, which is likely to account for their lower activity compared to PonA.[9] While the hydrophobic pocket accommodating the side chain of PonA is also occupied by the hydrophobic region of **12** or **24**, it is notable that the propanamide moiety of **24** interacts with an extra pocket on the receptor that is not utilized by **12** or PonA. It is likely that this additional interaction contributes to the relatively high activity of **24**, and this pocket may provide a basis for designing new ligands with improved activity.

4. METHODS

4.1. Chemicals. Unless otherwise specified, all solvents and reagents were obtained from commercial suppliers. The selected compounds listed in Figure 3 were purchased from Namiki Shoji Co., Ltd. (Tokyo, Japan), and their purity is guaranteed as >90%. Chemical structures of compounds **5**, **12**, and **24** which are detected to be active in the binding assay were confirmed by [^{1}H]NMR (AVANCE-400, BRUKER) and high resolution mass spectrometry (HRMS; JMS-700, JEOL) in Kyoto University. Compound **5** (Scientific Exchange, Inc., Center Ossipee, NH, USA), HRMS (FAB) *m/z*: $C_{19}H_{16}N_3O_6$ $[M+H]^+$, calcd 382.1039, found 382.1053. Compound **12** (MAYBRIDGE, Cornwall, UK), HRMS (FAB) *m/z*: $C_{21}H_{18}N_3O_4S$ $[M+H]^+$, calcd 408.1018, found 408.1012. Compound **24** (Pharmeks, Moscow, RUSSIA), HRMS (FAB) *m/z*: $C_{22}H_{19}N_3O_3S$ $[M^+]$ calcd 405.1147, found 405.1145. Ponasterone A (Invitrogen Corp.; Carlsbad, CA, USA) was used as a standard molting hormone agonist in the binding and reporter gene assays. The tritium labeled ponasterone A ([^{3}H] PonA; 140 Ci/mmol) was purchased from ARC Inc. (St. Louis, MO, USA).

4.2. Software and Structure Database. FILTER (ver. 2.0), OMEGA (ver. 2.1.0) ROCS (ver. 2.2), and FRED (ver. 2.2.5) of OpenEye Co. Ltd. (Santa Fe, NM) were used for *in silico* screening. GOLD 4.1.1 (Genetic Optimization for Ligand Docking) developed by Jones et al.[22] was purchased from Cambridge Crystallographic Data Center (Cambridge, U.K.). PDFAMS and PDFAMS-Pro were purchased from In-Silico Sciences, Inc. (Tokyo, Japan).[23] The cavity surface of the binding site of the SfEcR was visualized by the SYBYL

Figure 7. Alignment of sequence of Sf-9 to that of HvEcR. Numbers for HvEcR correspond to those given in PDB.

Table 3. Ranking of Screened Compounds Based the Summation of Three Scoring Function Values (GOLD) or Four Scoring Function Values (FRED)

Ranking	GOLD		FRED	
	Compound	Score	Compound	Score
1	**13**	2.85	**12**	**3.89**
2	**12**	**2.78**	**24**	**3.75**
3	**23**	2.77	**23**	3.71
4	**24**	**2.74**	**20**	3.69
5	**20**	2.73	**6**	3.68
6	**17**	2.56	**22**	3.68
7	**8**	2.52	**11**	3.64
8	**5**	**2.49**	**18**	3.62
9	**22**	2.49	**9**	3.61
10	**18**	2.48	**13**	3.59
11	**19**	2.44	**3**	3.58
12	**16**	2.37	**14**	3.57
13	**6**	2.37	**17**	3.54
14	**3**	2.37	**8**	3.49
15	**9**	2.37	**15**	3.48
16	**15**	2.33	**5**	**3.47**
17	**14**	2.31	**4**	3.41
18	**10**	2.27	**19**	3.39
19	**1**	2.17	**16**	3.39
20	**2**	2.07	**21**	3.32
21	**21**	2.05	**7**	3.29
22	**7**	1.82	**1**	3.25
23	**4**	1.78	**2**	3.23
24	**11**	1.55	**10**	3.16
	ponA	3.00	ponA	4.00

(ver. 7.3; Tripos, St. Louis, MI) module, MOLCAD, using the Fast Connolly algorithm. pK_a values of compounds were evaluated by ACD laboratories pK_a database software (Advanced Chemistry Development, Toronto, Canada).

The chemical structure database used for the virtual screening of about 3 million compounds was supplied from Namiki Shoji Co. Ltd. (Tokyo, Japan). The number of structures of the original database was reduced to 2.1 million based on molecular weight (MW: 300 to 500) using FILTER. For these 2.1 million compounds, we generated about 170 conformations per each compound with local minimum energies (350 million conformations in total) using OMEGA. Energies of all conformations are within 25 kcal/mol of the most stable conformation for each compound, and a root-mean-square deviation (rmsd) cutoff is set as 0.8 Å. The conformation of the query molecule, PonA, was obtained from EcR-bound PonA structure (PDB code: 1R1K).

4.3. Virtual Screening. The initial screening was performed by ROCS for all 350 million conformations using combination of two parameters, Shape Tanimoto[24] and Scaled Color score.[25] The Shape Tanimoto and Scaled Color score of each query structure were calculated against the template structure. Shape Tanimoto ranges from 0 to 1, and Scaled Color score ranges from −1 to 1. The Shape Tanimoto coefficient of 20E to PonA is 0.95. The Scaled Color score was calculated using the Implicit Mills-Dean color force field, focusing on the overlapping of HB donor/acceptor, anionic/cationic charge, hydrophobic region, and ring structure. The Scaled Color score is equivalent to the chemical similarity. When the Shape Tanimoto score alone was used to perform the ranking, the top hits were too hydrophobic. Accordingly, both the Shape Tanimoto and Scale Color scores were used to determine the ranking. Both shape and chemical similarity values of some screened compounds are listed in Table 1. The first nonsteroidal compound is ranked at 16 as shown in Table 1.

To further reduce the number of the initial screening hits, we used a Python script that is described below. In this second step, the overlaps of six HB acceptors and one hydrophobic moiety are considered as schematized in Figure 9. Among multiple overlaps, it was assumed that the presence of a HB acceptor at C-2 was essential, because the oxygen atom of 2-OH group is working as a multiple HB acceptor.[9] With respect to the hydrophobic interaction, the carbon atom at 25-position structure is selected as the central position of hydrophobic region. If a query compound has more than 6 overlaps including the HB acceptor matching to 2-OH against the template structure, it is selected (Figure 9). Finally, 237 compounds were submitted to visual inspection to select 24 compounds focusing on the chemical diversity, adequacy of overlap, and cost as described in Results. The flowchart of the virtual screening procedure is shown in Figure 10.

4.4. In-House Program Using Python Script. After superimposing each query to the template molecule with ROCS, the matching of selected heavy atoms of the database molecules (O, N, S, Cl, Br, F, I, and P) with the 6 oxygen atoms of PonA was checked using a custom Python script. The ROCS hits were checked for matching hydrogen bonding features in PonA by ensuring that the corresponding heavy atom of each database compound was within a 1 Å cube around the oxygen atoms bound to C-2, C-3, C-6, C-14, C-20, and C-22 of PonA (Figure 9). A 3 Å cube and a 6 Å cube around C-25 were then used to analyze the hydrophobic feature matching. Accordingly, in order to match the hydrophobic moiety of PonA, a hit molecule must place a C-atom within the 3 Å cube around C-25 but may not place any of the selected heavy atoms in the 6 Å cube around C-25; resulting in a region of reduced hydrophilicity in the 6 Å cube (Figure 9).

4.5. Construction of the Ligand Binding Domain of EcR. The ligand binding domain (LBD) of SfEcR was constructed using a homology modeling software PDFAMS.[23] The primary

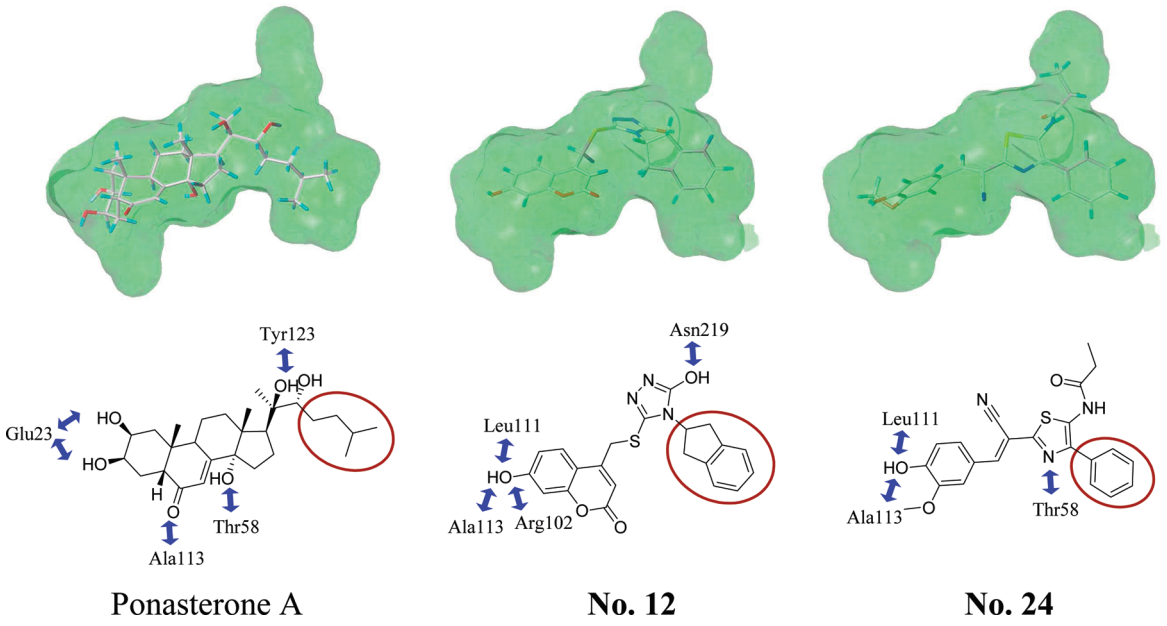


Figure 8. The docking simulation model against SfEcR ligand binding domain constructed using Goldscore function. Blue arrows and brown circle indicate the HBs and hydrophobic interactions observed in the models.

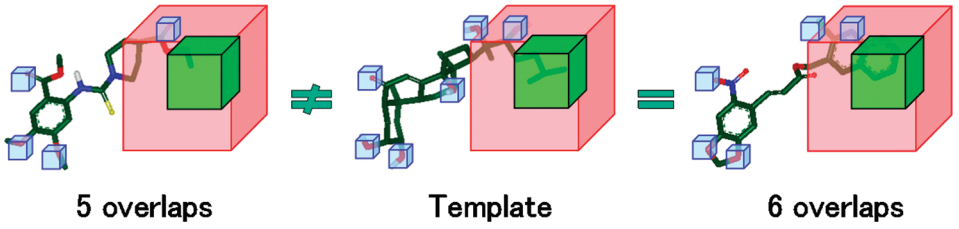


Figure 9. Schematic expression of the similarity in the 2nd screening using in-house program. Cyan is a cube with 1 Å on a side in which heavy atoms exist as HB acceptor. Red and green are the cubes with 6 Å and 3 Å on a side with the same center, respectively.

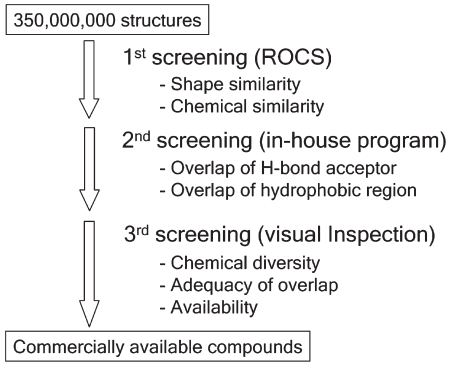


Figure 10. Flowchart for the virtual screening.

sequence of SfEcR-LBD was derived from two partial sequences which are reported by two groups, Chen et al.[26] and Pascal et al.[27] (NCBI accession number; AAM54494 and CAD58232). The sequence of the ligand binding domain of SfEcR is aligned with that of HvEcR as shown in Figure 7.

The coordinates of the crystal structure of PonA bound to HvEcR-LBD/USP-LBD was downloaded from the Protein Data Bank, code 1R1K. We selected the chain D (EcR) of this apo structure as the template structure for homology modeling. The structure of HvEcR-LBD was replaced with that of SfEcR-LBD followed by optimization of the domain using the simulated annealing method in PDFAMS-Pro. One hundred trials of simulated annealing were performed in parallel to construct 100 independent conformations of the SfEcR-LBD. These 100 conformations of SfEcR-LBD were examined by round-robin analysis to select representative conformations where gaps among conformations were evaluated. The gaps were calculated based on the rmsd of 22 heavy atoms making up the ligand binding pocket. The cavity is defined by the volume surrounding the ligand in the crystal structure at a distance of 4.5 Å. Similar structures were eliminated using the cutoff value of rmsd (0.35 Å).

4.6. GOLD Docking Simulation. The binding pocket was defined by 51 residues surrounding PonA of the crystallographic structure, 1R1K, at the distance of 7 Å. The docking experiment was replicated 10 times using following parameters of the GA (Population size = 100, Selection pressure = 1.1, Operation number = 150,000, Island number = 5, Niche Size = 2, Migrate = 10, Mutate = 95, Crossover = 95). If the top three solutions were within 1.5 Å of the rmsd value, the docking process was terminated. Three scoring functions, Goldscore, Chemscore,[13] and ASP scoring function,[14] were used. In this study, the sum of the normalized three scoring functions was used as the ranking metric.

In the docking calculations, ionization states, tautomeric changes, and isomers were carefully checked for 11 compounds (**4**, **5**, **11**-**14**, **19**, **20**, **22**-**24**). The ionization of the compounds at

dx.doi.org/10.1021/ci100400k |*J. Chem. Inf. Model.* 2011, 51, 296–305

neutral pH (6.8-7.2) was defined using ACD laboratories pK_a database software. Tautomeric forms (**13**, **22**, **23**), *E*/*Z* isomers (**4**, **5**, **11**, **13**, **14**, **19**-**24**), and optical isomers (**11** and **12**) were considered, and all 43 tautomers/isomers were submitted to the docking analysis.

4.7. FRED Docking Simulation. Various conformations were pregenerated for each ligand by OMEGA using parameters described at Section 4.2. The binding pocket was defined by a "box", which was constructed by extending the size of a cocrystallized ligand by 5 Å (addbox parameter of FRED).

All pregenerated conformations of each ligand were docked into a rigid target protein. All the generated docking poses were ranked by each of the four implemented functions (Shapegauss, PLP, Chemgauss2, Chemgauss3). In this study, the sum of the normalized four scoring functions was used as the ranking metric.

4.8. Binding Assay. Sf-9 and BCIRL-Lepd-SL1 cells were cultured in EX-CELL 401 (SAFC Bioscience, Lenexa, KS, USA) containing 10% of FBS (Fatal Bovine Serum; CAMBREX Co., North Brunswick, NJ, USA) at 25 °C. BCIRL-Lepd-SL1 cell line which is established from female pupae of *L. decemlineata* is kindly gifted from Dr. Smagghe with the permission of Biological Control of Insect Research Laboratory of USDA Midwest Area (Columbia, MS) for the utilization of cells.

The ligand-receptor binding assay was executed against intact cells according to the previously reported methods.[11,12] In brief, 400 μL of cell suspension (ca. 4 × 10^6 cells/mL) containing 1 μL of DMSO solution of the test compound and 2 μL of the 70% ethanol solution of [^{3}H] PonA (ca. 60,000 dpm) was incubated for 30 min at 25 °C. The reaction mixture was immediately filtered through a glass filter (GF/F; ADVANTEC, Japan) and washed three times with water (1 mL). The filter was dried under infrared light and transferred to the vial for a liquid scintillation counter (LSC). The radioactivity collected in the filter was counted with LSC in 3 mL of Aquasol-2 (Packard Instrument Co., Meriden, CT, USA). The binding assay was performed at 250 μM, but the lower concentration (80 μM) was chosen for some compounds **3**, **4**, **6**, **7**, **11**, **18**, **21**, and **22** due to the limited solubility in DMSO. For the compounds which inhibited the [^{3}H]PonA incorporation more than 50%, the concentration-response curves were drawn and IC_{50} values were determined by probit analysis.

4.9. Reporter Gene Assay. The reporter gene assay was performed in Sf-9 cells. A reporter plasmid for ecdysone agonists-dependent activation of transcription is constructed from pBmbA/hsp27/firefly luciferase (hsp27/Fluc) which is referred as ERE-b.act.luc.[28] It contains seven repeats of hsp27 EcRE sequence of *D. melanogaster* on the upper stream of a promoter region of actin A3 gene of *Bombyx mori* followed by a firefly luciferase coding gene. This plasmid is a kind gift from Dr. Swevers.

The plasmid was amplified in *Escherichia coli* and purified by QIAGEN Plasmid Midi Kit (QIAGEN GmbH, Hilden, Germany). The mixture of plasmids and Tfx^{TM}-20 Transfection Reagent (Promega; Madison, WI) were added to cell suspension (ca. 2.0 × 10^6 cells/mL; 200 μL) in 24-well plate. After incubating for 1 h, 400 μL of medium and 6 μL of solution of test compound (1% v/v) was added and incubated. After 48 h, cells in each well were scraped and transferred to a 1.5 mL microtube. After discarding the medium by centrifugation (300 × g, 4 °C, 8 min), 250 μL/tube of PBS was added to harvested cells to wash. PBS was removed again by centrifugation (300 × g, 4 °C, 8 min). The cells were submitted to the luciferase assay. Cells were lysed by adding 100 μL/tube of Passive Lysis Buffer and incubated for 15 min at 25 °C to lyse cells. The lysate (20 μL) was transferred to a 96 well microplate for the luciferase assay. Chemical luminescence was measured by the 96 microplate luminometer GLOMAX (Promega; Madison, MI) according to the manufacturer's instruction.

AUTHOR INFORMATION

Corresponding Author

*Phone: +81-75-753-6117. Fax: +81-75-753-6123. E-mail: naka@kais.kyoto-u.ac.jp.

ACKNOWLEDGMENT

We are thankful to Dr. Paul Hawkins (OpenEye Scientific Software) for carefully reviewing this manuscript and to Dr. Craig E. Wheelock (Karolinska Institute) for the helpful suggestion to revise manuscript. The study was supported, in part, by the 21st century COE program for Innovative Food and Environmental Studies Pioneered by Entomomimetic Sciences, from the Ministry of Education, Culture, Sports, Science and Technology of Japan. Dr. Toshiyuki Harada is a recipient of a Research Fellowship of the Japan Society for the Promotion of Science for Young Scientists. We also thank Dr. Guy Smagghe for the provision of BCIRL-Lepd-SL1 cells. The plasmid, ERE-b.act.luc, was kindly provided by Dr. Luc Swevers of NCSR "DEMOKRITOS", Greece. We thank OpenEye Scientific Software for their provision of a free license for all their software. A part of experiments was performed in the RI center of Kyoto University.

ABBREVIATION

20E, 20-hydroxyecdysone; EcR, ecdysone receptor; USP, ultraspiracle; RXR, retinoid X receptor; DBH, *N*-*t*-butyl-*N*,*N*′-dibenzoylhydrazine; PonA, ponasterone A; Hv, *Heliothis virescens*; QSAR, quantitative structure-activity relationship; HB, hydrogen bond; LBVS, ligand-based virtual screening; HTD, high throughput docking; Sf, *Spodoptera frugiperda*; LBD, ligand binding domain.

REFERENCES

(1) Nakagawa, Y.; Henrich, V. C. Arthropod nuclear receptors and their role in molting. *FEBS J.* **2009**, *276*, 6128–6157.

(2) Wing, K. D. RH 5849, a nonsteroidal ecdysone agonist: effects on a *Drosophila* cell line. *Science* **1988**, *241*, 467–469.

(3) Mikitani, K. An automated ecdysteroid receptor binding assay using a 96-well microplate. *J. Seric. Sci. Jpn.* **1996**, *65*, 141–144.

(4) Smith, H. C.; Cavanaugh, C. K.; Friz, J. L.; Thompson, C. S.; Saggers, J. A.; Michelotti, E. L.; Garcia, J.; Tice, C. M. Synthesis and SAR of cis-1-benzoyl-1,2,3,4-tetrahydroquinoline ligands for control of gene expression in ecdysone responsive systems. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 1943–1946.

(5) Birru, W.; Fernley, R.; Graham, L. D.; Grusovin, J.; Hill, R. J.; Hofmann, A.; Howell, L.; James, P.; Jarvis, K. E.; Johnson, W. M.; Jones, D. A.; Leitner, C.; Liepa, A. J.; Lovrecz, G. O.; Lu, L.; Nearn, R. H.; O'Driscoll, B. J.; Phan, T.; Pollard, M.; Turner, K. A.; Winkler, D. A. Synthesis, binding and bioactivity of γ-methylene γ-lactam ecdysone receptor ligands: Advantages of QSAR models for flexible receptors. *Bioorg. Med. Chem.* **2010**, *18*, 5647–5660.

(6) Palli, S. R.; Tice, C. M.; Margam, V. M.; Clark, A. M. Biochemical mode of action and differential activity of new ecdysone agonists against mosquitoes and moths. *Arch. Insect Biochem. Physiol.* **2005**, *58*, 234–242.

(7) Billas, I. M.; Iwema, T.; Garnier, J. M.; Mitschler, A.; Rochel, N.; Moras, D. Structural adaptability in the ligand-binding pocket of the ecdysone hormone receptor. *Nature* **2003**, *426*, 91–96.

(8) Arai, H.; Watanabe, B.; Nakagawa, Y.; Miyagawa, H. Synthesis of ponasterone A derivatives with various steroid skeleton moieties and evaluation of their binding to the ecdysone receptor of Kc cells. *Steroids* **2008**, *73*, 1452–1464.

(9) Harada, T.; Nakagawa, Y.; Akamatsu, M.; Miyagawa, H. Evaluation of hydrogen bonds of ecdysteroids in the ligand-receptor interactions using a protein modeling system. *Bioorg. Med. Chem.* **2009**, *17*, 5868–5873.

(10) Wheelock, C. E.; Nakagawa, Y.; Harada, T.; Oikawa, N.; Akamatsu, M.; Smagghe, G.; Stefanou, D.; Iatrou, K.; Swevers, L. High-throughput screening of ecdysone agonists using a reporter gene assay followed by 3-D QSAR analysis of the molting hormonal activity. *Bioorg. Med. Chem.* **2006**, *14*, 1143–1159.

(11) Nakagawa, Y.; Minakuchi, C.; Ueno, T. Inhibition of [^{3}H]ponasterone a binding by ecdysone agonists in the intact Sf-9 cell line. *Steroids* **2000**, *65*, 537–542.

(12) Nakagawa, Y.; Minakuchi, C.; Takahashi, K.; Ueno, T. Inhibition of [^{3}H]ponasterone A binding by ecdysone agonists in the intact Kc cell line. *Insect Biochem. Mol. Biol.* **2002**, *32*, 175–180.

(13) Eldridge, M. D.; Murray, C. W.; Auton, T. R.; Paolini, G. V.; Mee, R. P. Empirical scoring functions: I. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. *J. Comput.-Aided Mol. Des.* **1997**, *11*, 425–445.

(14) Mooij, W. T.; Verdonk, M. L. General and targeted statistical potentials for protein-ligand interactions. *Proteins* **2005**, *61*, 272–287.

(15) McGann, M. R.; Almond, H. R.; Nicholls, A.; Grant, J. A.; Brown, F. K. Gaussian docking functions. *Biopolymers* **2003**, *68*, 76–90.

(16) Verkhivker, G. M.; Bouzida, D.; Gehlhaar, D. K.; Rejto, P. A.; Arthurs, S.; Colson, A. B.; Freer, S. T.; Larson, V.; Luty, B. A.; Marrone, T.; Rose, P. W. Deciphering common failures in molecular docking of ligand-protein complexes. *J. Comput.-Aided Mol. Des.* **2000**, *14*, 731–751.

(17) Ogura, T.; Minakuchi, C.; Nakagawa, Y.; Smagghe, G.; Miyagawa, H. Molecular cloning, expression analysis and functional confirmation of ecdysone receptor and ultraspiracle from the Colorado potato beetle Leptinotarsa decemlineata. *FEBS J.* **2005**, *272*, 4114–4128.

(18) Wang, R.; Wang, S. How does consensus scoring work for virtual library screening? An idealized computer experiment. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 1422–1426.

(19) Clark, R. D.; Strizhev, A.; Leonard, J. M.; Blake, J. F.; Matthew, J. B. Consensus scoring for ligand/protein interactions. *J. Mol. Graphics Modell.* **2002**, *20*, 281–295.

(20) Yang, J. M.; Chen, Y. F.; Shen, T. W.; Kristal, B. S.; Hsu, D. F. Consensus scoring criteria for improving enrichment in virtual screening. *J. Chem. Inf. Model.* **2005**, *45*, 1134–1146.

(21) Cheng, T.; Li, X.; Li, Y.; Liu, Z.; Wang, R. Comparative assessment of scoring functions on a diverse test set. *J. Chem. Inf. Model.* **2009**, *49*, 1079–1093.

(22) Jones, G.; Willett, P.; Glen, R. C. Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation. *J. Mol. Biol.* **1995**, *245*, 43–53.

(23) Ogata, K.; Umeyama, H. An automatic homology modeling method consisting of database searches and simulated annealing. *J. Mol. Graphics Modell.* **2000**, *18* (258−272), 305–256.

(24) Rush, T. S., 3rd; Grant, J. A.; Mosyak, L.; Nicholls, A. A shape-based 3-D scaffold hopping method and its application to a bacterial protein-protein interaction. *J. Med. Chem.* **2005**, *48*, 1489–1495.

(25) Hawkins, P. C.; Skillman, A. G.; Nicholls, A. Comparison of shape-matching and docking as virtual screening tools. *J. Med. Chem.* **2007**, *50*, 74–82.

(26) Chen, J. H.; Turner, P. C.; Rees, H. H. Molecular cloning and induction of nuclear receptors from insect cell lines. *Insect Biochem. Mol. Biol.* **2002**, *32*, 657–667.

(27) Pascal, E. J.; Valentine, S. A.; Brown, J. A.; Cockress, A. S.; Johnson, B. D. 2002 WO 02/061102 A2.

(28) Swevers, L.; Kravariti, L.; Ciolfi, S.; Xenou-Kokoletsi, M.; Ragoussis, N.; Smagghe, G.; Nakagawa, Y.; Mazomenos, B.; Iatrou, K. A cell-based high-throughput screening system for detecting ecdysteroid agonists and antagonists in plant extracts and libraries of synthetic compounds. *FASEB J.* **2004**, *18*, 134–136.