Insights into Ligand-Elicited Activation of Human Constitutive Androstane Receptor Based on Novel Agonists and Three-Dimensional Quantitative Structure−Activity Relationship

Johanna Jyrkkärinne,[†,||] Björn Windshügel,[||,§] Toni Rönkkö,[§] Anu J. Tervo,[§,#] Jenni Küblbeck,[†] Maija Lahtela-Kakkonen,[§] Wolfgang Sippl,[¶] Antti Poso,[§] and Paavo Honkakoski*[,†]

Departments of Pharmaceutics and Pharmaceutical Chemistry, University of Kuopio, P.O. Box 1627, 70211 Kuopio, Finland, and Department of Pharmacy, Martin-Luther-Universität Halle-Wittenberg, Wolfgang-Langenbeck-Strasse 4, 06120 Halle (Saale), Germany

Received June 17, 2008

The human constitutive androstane receptor (CAR, NR1I3) is an important regulator of xenobiotic metabolism and other physiological processes. So far, only few CAR agonists are known and no explicit mechanism has been proposed for their action. Thus, we aimed to generate a 3D QSAR model that could explain the molecular determinants of CAR agonist action. To obtain a sufficient number of agonists that cover a wide range of activity, we applied a virtual screening approach using both structure- and ligand-based methods. We identified 27 novel human CAR agonists on which a 3D QSAR model was generated. The model, complemented by coregulator recruitment and mutagenesis results, suggests a potential activation mechanism for human CAR and may serve to predict potential activation of CAR for compounds emerging from drug development projects or for chemicals undergoing toxicological risk assessment.

Introduction

The 48 members of the human nuclear receptor (NR[a]) superfamily are ligand-dependent transcription factors that control development, cellular homeostasis, and metabolic pathways. Disturbances in NR function underlie several common diseases, and modulation of NR activity by small-molecule ligands constitutes an important mechanism of drug therapy.[1] Agonist- and antagonist-bound NRs will recruit coactivators and corepressors, respectively, to mediate opposing effects on the NR target gene transcription. The type of the coregulator recruited depends on the ligand-dependent shift in the position of the activation helix 12 (H12) within the NR ligand-binding domain (LBD).[2,3] The constitutive androstane receptor (CAR, NR1I3) and the pregnane X receptor (PXR, NR1I2), two members of the NR1I subfamily, play key roles as guardians against xenobiotic challenges. They control the expression of enzymes and proteins responsible for metabolism and elimination of a variety of drugs, environmental contaminants, plant-derived substances, and even endogenous end-products such as bile acids and bilirubin.[4,5] In addition, CAR and PXR modulate glucose and lipid metabolism through their interplay with FoxO and FoxA transcription factors.[6,7] Finally, studies with Car null mice have shown that the presence of *Car* is absolutely necessary for chemically induced liver hypertrophy and tumor promotion.[8−10] These findings emphasize the pivotal and wide-ranging roles for CAR in the control of hepatic metabolism, liver growth, and pathogenesis.

CAR and PXR can be considered as promiscuous sensors because their LBDs can recognize a wide repertoire of structurally dissimilar ligands.[11,12] Moreover, their species differences in ligand recognition are remarkable: for examples, rifampicin is an agonist of human but not rat PXR,[13] and 17α-ethynyl-3,17β-estradiol (EE2) is an inverse agonist of human CAR but an agonist of mouse CAR.[14,15] Such species differences, low specificity of CAR and PXR receptors, and the high structural diversity of their ligands are in profound contrast with steroid hormone receptors that exhibit a high degree of selectivity and affinity for their ligands and remarkable conservation between species.

Ligand activation of CAR and PXR will result in induction of drug metabolism, disturbances in the levels and/or turnover of endogenous substances, and sometimes increased toxicity in a species-dependent manner. Therefore, measurement and prediction of CAR and PXR agonism in humans are of importance for the drug development to anticipate induction and its consequences.[16,17] To this end, we have studied the mechanisms that govern the constitutive activity of mouse and human CAR, their suppression by steroidal inverse agonists, and species differences therein.[14,15,18] Although X-ray structures of the human CAR LBD in complex with agonists 6-(4-chlorophenyl)imidazo[2,1-*b*][1,3]thiazole-5-carbaldehyde *O*-(3,4-dichlorobenzyl)oxime (CITCO) and 5β-pregnane-3,20-dione have been solved,[19] it is not yet understood which structural features of the agonist contribute to CAR activation and to the position of the H12, which is critical for the coactivator binding and NR activation.[1−3] Because the ligand-free CAR has not been crystallized, it is not known whether the diverse activators of CAR mediate their effects via the same mechanism. It has been suggested that CAR can be activated, in some cases, by an indirect mechanism not involving direct ligand binding.[7,16] Moreover, only a few human CAR activators are currently well established, and therefore, no 3D QSAR models have been published.[20] Variability in agonist responses[4] and even opposing effects (e.g., for clotrimazole) have been reported.[13,21,22] Such

* To whom correspondence should be addressed. Telephone: +358-40-3552490. Fax: +358-17-162456. E-mail: paavo.honkakoski@uku.fi.
† Department of Pharmaceutics, University of Kuopio.
|| Both authors (J. J., B.W.) contributed equally to this work.
§ Department of Pharmaceutical Chemistry, University of Kuopio.
# Present address: DECS Computational Chemistry, AstraZeneca R&D Mölndal, SC2, Pepparedsleden 1, 43183 Mölndal, Sweden.
¶ Martin-Luther-Universität Halle-Wittenberg.

[a] Abbreviations: *Car*, constitutive androstane receptor; 3D QSAR, three-dimensional quantitative structure−activity relationship; NR, nuclear receptor; H12, helix 12; LBD, ligand-binding domain; EE2, 17α-ethynyl-3,17β-estradiol; PXR, pregnane X receptor; CITCO, 6-(4-chlorophenyl)imidazo[2,1-*b*][1,3]thiazole-5-carbaldehyde *O*-(3,4-dichlorobenzyl)oxime; VS, virtual screening; LBP, ligand-binding pocket; DMSO, dimethyl sulfoxide; SDEP, standard error of prediction; vdW, van der Waals; PDB, Protein Data Bank; PLS, partial least squares; TMPP, tri(*p*-methylphenyl) phosphate; TPP, triphenyl phosphate; MIF, molecular interaction field.

10.1021/jm800731b CCC: $40.75
Published on Web 11/05/2008

complications in defining CAR activation could be in part due to differential contents of NR coregulators in cell lines used[23,24] or to different assay setups that have not been properly optimized or formally validated.

Because of these uncertainties, we sought to identify additional activators for human CAR. To this end, the Tripos LeadQuest database was screened using two different approaches. So far, a diverse set of virtual screening (VS) methods for identification of novel bioactive molecules[25] have been successfully applied for NRs.[26] On the basis of these ideas, we initially conducted a pharmacophore search in combination with a receptor-based molecular docking approach. In a second approach, we applied a purely ligand-based procedure recently developed in our group.[27,28] Altogether 27 human CAR agonists were identified with a validated CAR activation assay and the results were used to generate a 3D QSAR model. Novel insights into the human CAR activation were obtained through the use of the model coupled with directed mutations within the ligand-binding pocket (LBP), analysis of ligand docking, and coregulator assays. These findings have profound implications for diverse chemicals generated in drug development processes or undergoing toxicological risk assessment.

Results

Virtual Screening. In search of novel CAR agonists we applied two independent approaches. Initially, we employed a combined method consisting of a pharmacophore search and a molecular docking procedure to find novel human CAR agonists. First, the LeadQuest database was screened using a pharmacophore model generated within SYBYL on the basis of our homology model.[18,29] The search query was defined by the volume and shape of the CAR LBP as well as its potential van der Waals (vdW) interactions with ligands tri(*p*-methylphenyl) phosphate (TMPP) and clotrimazole.[18] Retrieved hit compounds from the VS search (about 10 000) were then docked into the LBP using the program GOLD with the GoldScore as fitness function. We rescored all docking poses by applying another scoring function (X-Score) and modifications of the GoldScore and X-Score, followed by redocking of the hits by using an independent docking program (see Experimental Section for details). From each approach, the top-ranked 20 compounds were selected. After correction for molecules common in these subselections, the final VS data set contained 64 compounds, of which 54 were available from the supplier for in vitro testing.

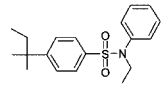
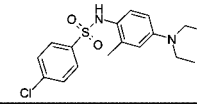
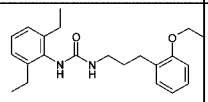
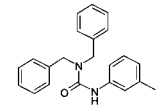
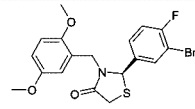
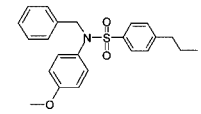
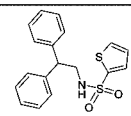
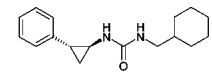
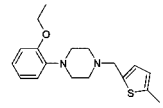
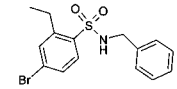
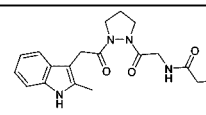
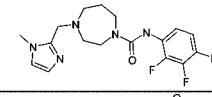
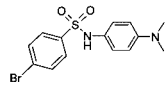
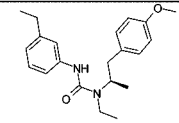
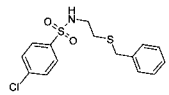
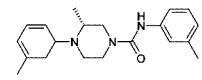
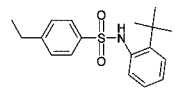
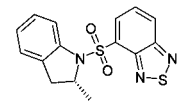
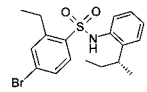
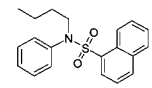
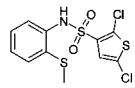
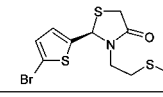
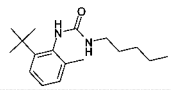
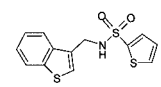
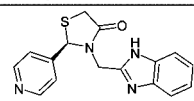
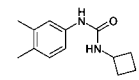
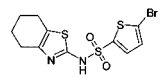
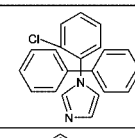
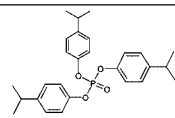
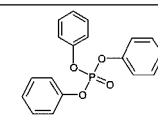
Second, we applied a purely ligand-based search with our BRUTUS software.[27,28] The BRUTUS search is based on automatic molecular field alignment and similarity of selected templates and molecules in a database. The algorithm detects the similarity of steric and electrostatic fields and thus allows VS, which is independent of structural relatedness in 2D fashion. Two structurally dissimilar CAR agonists (CITCO and TMPP) were used as templates to increase the probability to find new chemical scaffolds. On the basis of this VS search, 35 compounds were finally selected for in vitro testing.

Human CAR Activation Assay. Because of the problems described above, the performance of the assay used to detect human CAR agonists is very important. The GAL4-human CAR LBD fusion protein reproduces the ligand specificity of the full-length human CAR,[13,14,21,30,31] but it has the advantage of avoiding cross-activation of the reporter by interfering endogenous NRs.[15,32] In pilot experiments, we optimized this assay by selecting a hepatoma cell line that gave reproducible activating responses with established CAR agonists. After optimization of the protocol, the Z' performance parameter of the assay was 0.67, the coefficient of variation was low (8.4%), and the minimum significant ratio was 2.04, all indicating an excellent and reproducible assay.[33] Next, all 89 compounds obtained from both VS searches were dissolved in dimethyl sulfoxide (DMSO) and initially tested at 10 μM concentration along with reference compounds clotrimazole, TMPP, and TPP.[18] Subsequently, the active compounds were tested with multiple concentrations ranging from <1 to 50 μM. However, at the lowest concentrations, no clear activation and no significant differences between compounds were seen. This suggests that the ligand affinity of CAR is quite low and roughly of similar value for many compounds, in line with the role of CAR as a promiscuous sensor. A few of the compounds displayed some cell toxicity at 25−50 μM concentration (data not shown). Therefore, the precise potency or EC_{50} determinations for all active compounds would have been impossible, and the activities recorded at 10 μM concentrations were chosen as the biological variable. As noted by others,[3] the EC_{50} values do not necessarily reflect the actual transcriptional response of the NR because the ligands may have different coactivator preferences[34,35] and the NR activity recorded in such cell-based assays depends on the total set of coactivators recruited.[23,24]

NR agonists significantly increase the receptor activity, and the criterion (cutoff) for an agonist should be set accordingly high. In addition, reliable reference substances must be used to control for intraday and day-to-day variations in the assay. The human CAR agonist CITCO proved to be unstable upon storage, and it gave highly variable results between assays (data not shown). Therefore, we used the previously studied clotrimazole, TPP, and TMPP as reference CAR agonists.[18,31] However, the generation of a 3D QSAR model requires diverse data, and thus, the cutoff value was set to a minimum of 1.3-fold activation. The cutoff is based on the mean of vehicle control readings added by their standard deviation multiplied by 3 [cut-off = vehicle mean + (3 × SD)]. According to this criterion (Table 1), 11 out of 54 compounds selected by the 3D pharmacophore/docking approach were agonists, which corresponds to a hit rate of 20.4%. The BRUTUS approach was more successful, resulting in 16 agonists out of 35 tested, yielding a hit rate of 45.7%. When a more restrictive cutoff (2-fold activation) was applied, the hit rates dropped to 13.0% (7/54) and 37.1% (13/35), respectively. The identified agonists included molecules with low CAR activation (e.g., compound **27**) but also some with much higher activity (e.g., compounds **1** and **17**). The range of activation was from 1.32- to 7.92-fold, confirming diversity of the data. The mean activation for all 27 agonists was 3.27-fold, falling within the range of the reference compounds (Table 2). The results were also verified by measuring the ligand-dependent recruitment of the human coactivator SRC1 to the CAR LBD in mammalian two-hybrid assays. The extent of CAR activation by compounds **1**−**27** correlated very well with enhanced association of SRC1 ($r^2 = 0.790$), which confirms that the active compounds were indeed human CAR agonists (Table 2, Supporting Information Figure S1).

The novel agonists (Table 1) exhibit a considerable structural and chemical variety, with a predominance of N-substituted and N,N-disubstituted sulfonamides (14 compounds), urea (8), or thiazolidin-4-one (3) derivatives, respectively. The diversity of compounds is also highlighted by their molecular weights (range 218.3−426.3 Da) and the number of rotatable bonds. Some molecules are rather rigid with only two rotatable bonds (e.g., **18**), while others are highly flexible possessing up to 11 rotatable bonds (e.g., **3**). A common feature of all molecules is the dominance of hydrophobic, often aromatic groups.

Table 1. Overview of Human CAR Agonists Identified by the VS Procedures

Compound[a]	CAR Act.[b]	Pred. Act.	Structure	Compound[a]	CAR Act.[b]	Pred. Act.	Structure
1 (1533-12717) B	6.35 ± 0.69	5.48		**2** (1528-04835) B	5.50 ± 0.35	3.69	
3 (1515-04862) B	4.88 ± 0.25	4.36		**4** (1514-12000) B	4.17 ± 0.14	3.62	
5 (1539-37280) B	3.87 ± 0.07	3.90		**6** (1528-08432) B	3.80 ± 0.37	3.75	
7 (1525-01232) B	3.28 ± 0.16	2.92		**8** (1514-02905) B	2.95 ± 0.15	2.73	
9 (1524-11283) B	2.70 ± 0.32	3.10		**10** (1533-05451) B	2.51 ± 0.19	2.17	
11 (1511-17455) B	2.39 ± 0.11	2.73		**12** (1518-16058) B	2.23 ± 0.25	2.86	
13 (1513-05479) B	2.03 ± 0.15	2.58		**14** (1527-31043) B	1.86 ± 0.23	2.89	
15 (1528-02837) B	1.63 ± 0.18	2.17		**16** (1514-13785) B	1.45 ± 0.16	1.54	
17 (1528-03084) M	7.92 ± 0.05	5.06		**18** (1554-05575) M	5.22 ± 0.01	4.83	
19 (1533-04305) M	4.47 ± 0.15	4.57		**20** (1557-04212) M	4.09 ± 0.36	3.78	
21 (1554-03757) M	3.68 ± 0.46	4.63		**22** (1539-32671) M	2.97 ± 0.13	3.50	
23 (1540-00528) M	2.60 ± 0.13	3.45		**24** (1557-04529) M	1.76 ± 0.17	1.66	
25 (1539-00861) M	1.40 ± 0.01	3.19		**26** (1527-12765) M	1.37 ± 0.10	1.91	
27 (1554-08348) M	1.32 ± 0.01	2.31		CLOTR	4.34 ± 0.28		
TMPP	2.45 ± 0.19			TPP	3.96 ± 0.15		

[a] Compound code, LeadQuest code in parentheses, search method (B = BRUTUS, M = based on homology model). [b] CAR activation at 10 μM (*n*-fold over DMSO vehicle; mean ± SD). DMSO vehicle = 1.00 ± 0.08.

Table 2. Compounds and Their Receptor Activities and Occupancy in Coefficient Regions Identified by the 3D QSAR Approach

		interaction with		areas occupied[a]				
compd	CAR activity[b]	SRC1[c]	NCoR[c]	I	II	III	IV	S1
1	6.35 ± 0.69	+	−	+	+	+	−	+
2	5.50 ± 0.35	+	−	+	−	+	−	−
3	4.88 ± 0.25	+	+	+	+	−	+	+
4	4.17 ± 0.14	+	−	+	−	−	−	+
5	3.87 ± 0.07	+	−	+	+	−	−	+
6	3.80 ± 0.37	−	−	+	+	+	+	+
7	3.28 ± 0.16	+	+	+	−	−	+	−
8	2.95 ± 0.15	−	+	+	−	−	+	−
9	2.70 ± 0.32	+	−	+	−	−	−	+
10	2.51 ± 0.19	−	−	+	−	−	+	+
11	2.39 ± 0.11	+	−	+	−	−	+	+
12	2.23 ± 0.25	−	−	+	−	−	+	−
13	2.03 ± 0.15	−	−	−	−	−	−	−
14	1.86 ± 0.23	−	−	+	+	−	+	−
15	1.63 ± 0.18	−	+	+	−	−	+	−
16	1.45 ± 0.16	−	−	+	−	−	+	−
17	7.92 ± 0.05	+	−	+	+	−	−	+
18	5.22 ± 0.01	+	+	+	+	−	−	+
19	4.47 ± 0.15	+	+	+	+	−	−	+
20	4.09 ± 0.36	+	+	+	−	−	−	+
21	3.68 ± 0.46	+	−	+	+	−	−	+
22	2.97 ± 0.13	+	+	+	−	−	−	+
23	2.60 ± 0.13	+	+	+	+	−	+	−
24	1.76 ± 0.17	−	+	+	+	−	+	−
25	1.40 ± 0.01	−	−	+	−	−	−	+
26	1.37 ± 0.10	−	−	−	+	−	+	−
27	1.32 ± 0.01	−	+	−	−	−	−	−
TPP	3.96 ± 0.15	+	−	−	+	+	+	−
TMPP	2.45 ± 0.19	+	−	+	+	−	+	−
clotrimazole	4.34 ± 0.28	+	+	+	−	−	−	+
CITCO	ND[d]	ND[d]	ND[d]	+	+	+	+	+
5β-pregnane-3,20-dione	ND[d]	ND[d]	ND[d]	+	+	−	−	+

[a] Areas I−III correspond to three positive regions, and area IV corresponds to the negative coefficient region. The last column indicates targeting to the subsite S1 in the human CAR LBP. [b] Human CAR activity (mean ± SD) relative to DMSO vehicle (=1.00 ± 0.08). [c] See Supporting Information Figure S1 for actual data on recruitment of SRC1 and NCoR. [d] ND: not determined.

3D QSAR of Human CAR Agonists. The GRID/GOLPE approach was applied to generate a 3D QSAR model based on the identified CAR agonists. When the modeling was initiated, the coordinates of the human CAR X-ray structure became available.[19] Because the alignment of the compounds is critical for the quality of the 3D QSAR model, the X-ray structure (PDB entry 1XVP) was used for a receptor-based alignment. Human CAR agonists were docked into the CAR LBP using GOLD and subsequently rescored with X-Score.[36] Molecular descriptors were generated by calculating interaction energies between both aromatic and phenolic probes and the data set using GRID,[37] and 3D QSAR models were generated with GOLPE. Initial data processing yielded a relatively low quality two-component model with $q^2 = 0.31$, standard error of prediction (SDEP) = 1.36, and $r^2 = 0.91$ (correlation between experimental and predicted activity). Because the data set used is structurally diverse and different areas of the LBP are occupied by the CAR agonists (see Table 2), a large number of molecular field properties (variables) do not contribute to the explanation of differences in the activity and hence can be characterized as noise. In order to remove such variables, smart region definition and fractional factorial design procedures[37] were utilized consecutively. The first procedure joins variables close in space and collapses the groups of variables containing virtually the same information to "smart regions". Next, smart regions are evaluated with the second procedure by which a large number of models are calculated with some regions excluded to evaluate the importance of the individual regions. Only those regions that contribute in a positive manner to the predictivity are retained. Using the "leave-one-out" cross-validation method, we obtained much improved statistics for the two-component model:

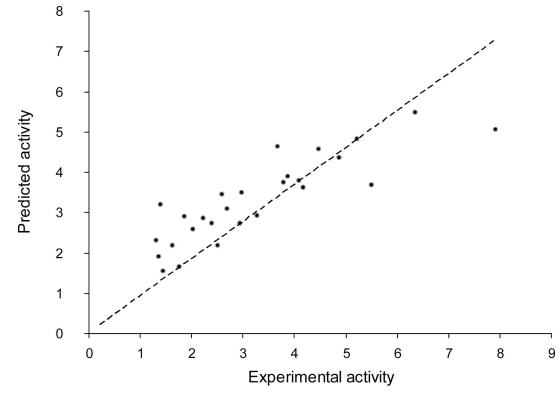


Figure 1. PLS plot of the 3D QSAR model based on the 27 novel human CAR agonists.

$q^2 = 0.73$, SDEP = 0.85, and $r^2 = 0.95$. By application of a more restrictive validation method using five random groups ("leave-20%-out"), the values declined only marginally ($q^2 = 0.68$, SDEP = 0.93, $r^2 = 0.94$). Both validation methods indicate a reliable model ($q^2 > 0.5$) without any clear outlier (Figure 1).

In order to interpret the 3D QSAR model, the PLS coefficient maps from both components were plotted together with the human CAR X-ray structure used to generate the alignment (Figure 2A). The contour maps show molecular field properties that are correlated with changes in human CAR activation. These maps reveal three larger positive regions (orange fields, I−III) and only one distinct negative region (yellow field, IV). Positive coefficient regions correlate well with small pockets or clefts

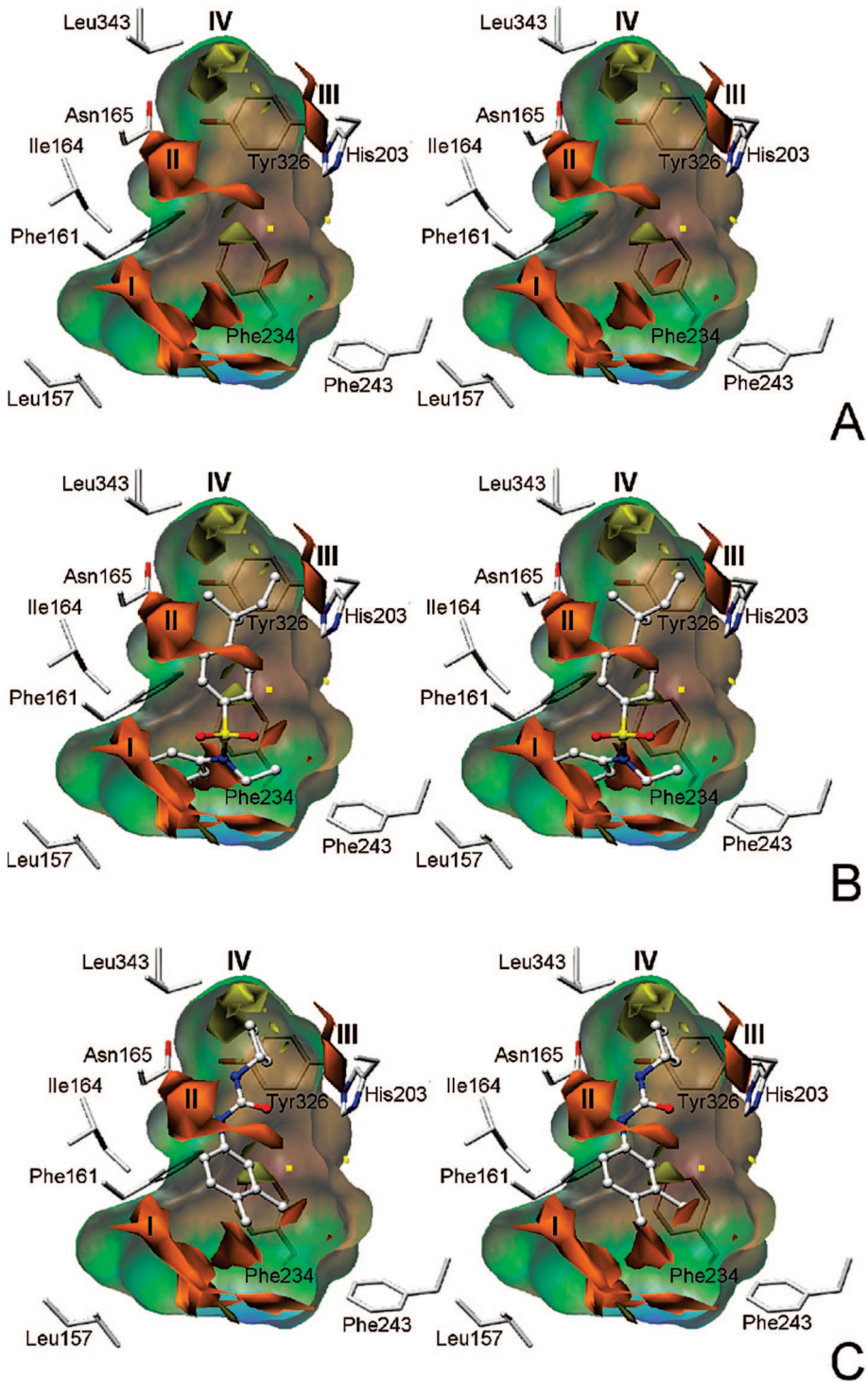


Figure 2. Regions important for human CAR activity with examples of agonist occupancy. (A) PLS coefficient grid plot for the aromatic probe. Coefficient levels −0.003 and 0.003 are shown in yellow and orange, respectively. The LBP is shown as a surface colored by the lipophilic (brown) and hydrophilic (blue) potentials. Surrounding amino acids are shown in capped sticks. (B) Most efficacious agonists (compound **1** shown) occupy the subsite S1 overlapping region I and partially contact other positive coefficient regions II and III. (C) Compounds with lower activation (compound **26** shown) are not located in region I but tend to occupy parts of the negative coefficient region IV.

within the LBP as depicted in Figure 2A. The largest coefficients are found in region I, which largely overlaps with a crevice (termed S1), framed by eight amino acids located on helices H2′ (Leu141), H3 (Leu157, His160, Phe161, Ile164), H6 (Asp228, Val232), and the β_3 strand (Tyr224) (Figures 2A and 4). For all agonists, interactions with each region have been examined extensively (Table 2). Generally, compounds occupying the positive regions (I−III) feature a higher CAR activation than those lacking that property or those associated with the negative region IV (Table 3). As an example, the strong agonist **1** overlaps with all three positive areas I−III but not with the negative region IV (Figure 2B). By contrast, the weak activator **26** appears to interact with the negative region IV (Figure 2C).

Then we examined whether the PLS coefficient plots match the data from the molecular interaction field (MIF) calculations based on the human CAR LBD. Therefore, GRID interaction fields between the aromatic probe and the structure 1XVP were calculated and compared with coefficient maps yielded by

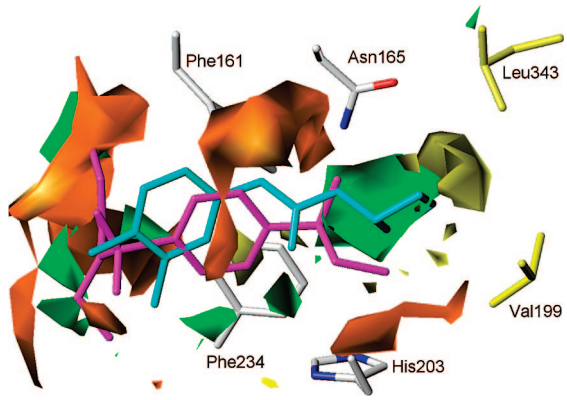


Figure 3. Comparison of positive (orange, 0.003) and negative (yellow, −0.003) coefficient regions with GRID fields (green; aromatic probe, contour level −2.7 kcal mol^{-1}) and binding modes of compounds **1** (magenta) and **26** (cyan). Amino acids whose vdW radii (not shown) overlap with negative coefficient regions are colored in yellow.

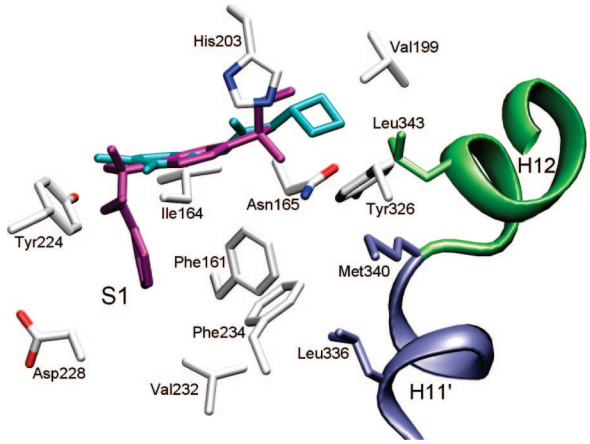


Figure 4. Binding modes of compounds **1** (magenta) and **26** (cyan) within the human CAR LBP. H12 is colored green, and the preceding H11′ is shown in blue. Single amino acids of the LBP forming the subsite S1 and the LBD/H12 interface are shown as capped sticks.

GOLPE (Figure 3). The most important positive coefficients (region I) and the GRID fields are in good agreement, but we could not detect MIFs that correspond to regions II and III. Instead, a large favorable MIF is located between the regions II and III, in the immediate vicinity of the negative region IV. This field coincides with a small channel (termed C1) connecting the LBP with H12, while the negative region IV overlaps slightly with amino acids in H5 (Val199) and H12 (Leu343). Upon docking, most human CAR agonists were found to (partially) occupy the C1 channel by placing an aromatic or aliphatic group therein, thus matching the favorable GRID fields.

Activation of Mutated Human CAR Receptors. We tested the activities of the 27 novel agonists with CAR receptors bearing mutations in the key LBP residues (Table 4, Supporting Information Table S1). The selected mutations differ in their responses to reference agonists TMPP and clotrimazole.[18] In agreement with our previous data, all seven mutations decrease the constitutive activity by 40−78%.

Mutation of Phe161 into alanine decreased the ligand-mediated CAR activation drastically. Altogether, 17 agonists failed to activate this mutant and marginal induction was observed for seven compounds. Only three compounds (**10**, **19**, **20**) exhibited more than 2-fold activation over the basal values (Table 4, Supporting Information Table S1). Four out of ten ligands with some activation of the Phe161Ala mutant (**3**, **10**, **16**, **24**) belonged to the group of compounds that partially overlap the negative coefficients region in the 3D QSAR model (Table 2). Only clotrimazole could reactivate the Phe161Ala mutant to the levels of the wild-type CAR.

Similarly, replacement of Tyr326 by alanine resulted in a profound decrease in ligand-induced activation. Only two agonists (**1**, **19**) could activate the Tyr326Ala mutant significantly, while the remaining ligands were either inactive or marginal activators (14 and 11 compounds, respectively). Notably, of the latter compounds, most (8 out of 11) contacted the negative coefficient region IV (Tables 2 and 4). Among the reference compounds, only TPP activated the Tyr326Ala mutant by 2.1-fold.

Results for the residue Asn165, which resides in proximity to Tyr326, are also dramatic. Fifteen compounds showed no effect on the Asn165Ala mutant, whereas seven ligands (**6**, **12**, **14**, **22**−**24**, **26**) displayed a marginal activation (Table 4, Supporting Information Table S1). Of the latter, all molecules except **22** showed some overlap with negative coefficients of region IV by placing a ring system in the C1 channel (Table 2). Five novel CAR agonists activated the Asn165Ala mutant 2- to 3.5-fold (**9**, **19**, **20**) or greater (**4, 21**). All reference compounds were able to activate this mutant.

Substitution of Phe234 by alanine resulted in loss of activity for 18 compounds, whereas six agonists retained marginal receptor activation. Only three compounds elicited over 2-fold activation (**1**, **19**, **20**). The responses of Phe234Ala were similar to those with the Tyr326Ala mutant in that wild-type activity levels were not reached, and of reference compounds, only TPP retained some activation.

The mutation His203Ala had the least impact on both basal activity (40% reduction) and activation potential of most ligands. Only four compounds (**3**, **22**, **25**, **27**) were devoid of any activity. Ten compounds elicited a marginal activation, and seven of these were partially located in the negative region IV. The rest of the ligands (13) activated His203Ala mutant at least 2-fold above the basal activity, and three of them reached an activation level comparable to wild-type CAR or higher (**15**, **19**, **23**). Clotrimazole appeared to inhibit the His203Ala mutant, acting as a reverse agonist, in line with our previous results.[18]

Mutating Ile164 into alanine reduced the basal activity by almost 80%, but most ligands were still capable of activating this mutant to some extent. Only five weak CAR agonists (**11**, **16**, **23**, **25**, **27**) displayed no activation, while 14 compounds showed marginal activation of the Ile164Ala mutant. Four compounds (**8**, **10**, **13**, **15**) were found to activate the Ile164Ala mutant significantly (2- to 3.5-fold). Of the latter, three compounds appeared to overlap with region IV, and five ligands were found to be highly efficacious activators (**2**, **4**, **18**−**20**).

Compared to the above mutations, results for the Phe243Ala mutant are in striking contrast. The basal activity was reduced by 40%, but all compounds could induce this mutant by more than 2-fold over the basal activity. In the majority of cases, the activation level either reached (10 compounds) or surpassed (7 compounds) the corresponding wild-type CAR activation (Table 4). The reference compounds showed the same trend. The highest potential was observed for compound **9**, which activated the Phe243Ala mutant by more than 5-fold above the level observed with wild-type CAR.

The tested mutations seem to form clusters of related phenotypes that display concordances in responses to the 27 agonists. We used principal component analysis to statistically validate the above observations. Principal component analysis is often used in drug discovery to reduce multidimensional data

Table 3. Average Human CAR Activity of Compounds Occupying Different Coefficient Regions[a]

area occupied	I	II	III	IV	S1	all compds
yes	3.66 ± 0.33 [24]	3.98 ± 0.54 [12]	5.22 ± 0.61 [3]	2.52 ± 0.27 [15]	4.03 ± 0.41 [15]	3.27 ± 0.32 [27]
no	1.57 ± 0.18 [3]	2.71 ± 0.29 [15]	3.03 ± 0.31 [24]	3.97 ± 0.48 [14]	2.33 ± 0.33 [12]	NA[b]

[a] Human CAR activity is indicated as the mean ± SEM, with the number of compounds in brackets. [b] NA: not applicable.

sets to lower dimensions for analysis.[39−41] The graphical plot of the first two principal components confirms that mutants Phe161Ala, Phe234Ala, Tyr326Ala, and Asn165Ala form a tightly clustered group while the others (Ile164Ala + His203Ala and Phe243) are clearly separated (Figure 5).

The first cluster formed by residues Phe161, Asn165, Phe234, and Tyr326 shares 10−17 concordances in pairwise comparisons of mutant receptors' responses to CAR agonists. Very strong correlations can be seen between Phe234 and Tyr326 as well as between Asn165 and Tyr326. All the amino acids within this cluster are fully concordant in response to seven ligands (**5**, **7**, **11**, **13**, **15**, **17**, **27**). Mutation of these residues significantly impaired the activation potential of human CAR agonists. In particular, Phe161 and Phe234 emerged as central amino acids for ligand-dependent activation. Depending on the amino acid, 50−63% of the ligands tested were unable to activate these mutant receptors. It is noteworthy that these residues belong either to the crevice S1 (Phe161) or to the LBD/H12 interface. The second cluster is formed by Ile164 and His203, which belong to residues framing the regions II and III. Both mutants share 12 concordances in the trends of ligand-dependent activation. Both amino acids have less impact on human CAR activation because most compounds (83%) are able to activate these mutants at least partially and several ligands reach or surpass the levels of the wild-type CAR. In contrast, Phe243 seems to be unrelated to the above clusters. Consistently, Phe243 is not part of amino acids that frame the positive regions I−III or the negative region IV. The Phe243Ala mutation further enhanced CAR activation for 18 agonists, and a decrease was observed for only two compounds (**3**, **21**). We have noted earlier that this mutation also converted EE2 from an inverse agonist to an agonist.[18] Phe243Ala thus appears to be a gain-of-function mutation with a decrease in basal activity.

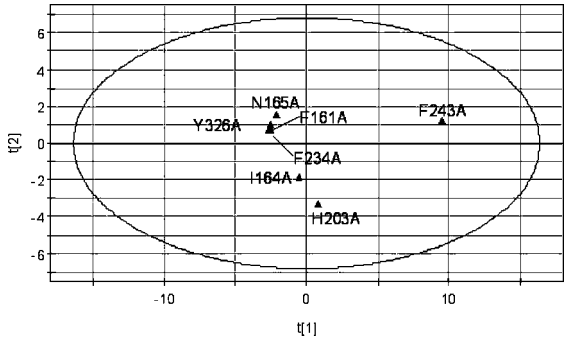


Figure 5. Two-dimensional plot of the principal component analysis showing the first (t[1]) and second (t[2]) principal components.

Ligand-Elicited Human CAR Activation. All docked human CAR agonists contact at least one of the amino acids constituting the LBD/H12 interface (Asn165, Phe234, Tyr326). Those vdW interactions alone may exert some activation by stabilization of the residues contacting H12. However, our data suggest that this is not sufficient for efficient CAR activation. The subsite S1 is targeted by most compounds that activate CAR efficaciously (Table 2). Most molecules with 3.5-fold activation or more appear to place a rigid aromatic group in S1 (**1**, **4**, **17**−**20**; see also Figure 2B). MIF calculations using GRID revealed favorable interactions between amino acids constituting S1 and the aromatic probe that colocalize with positive coefficients in S1 (Figure 3). Although several less active compounds (<2-fold activation) overlap with the positive coefficient region I (**14**−**16**, **24**, **25**), only one of these (**25**) actually seems to enter S1. Remarkably, the measured and predicted activities show the largest relative deviation for this particular compound (Table 1). The importance of S1 is highlighted by the fact that the average ligand-induced CAR activity of S1-targeting compounds (Table 3) clearly exceeds that of molecules not occupying this subsite (4.03-fold versus 2.33-fold). In addition, 11 out of 15 compounds occupying S1 were able to significantly enhance the interaction between the CAR LBD and the coactivator SRC1. Remarkably, compounds **1**, **4**, and **17**−**20**, placing an aromatic group in S1, had high SRC1 activity whereas compounds **14**−**16** and **24** did not enter S1 and exhibited only modest SRC1 activity in this assay (Table 2, Supporting Information Figure S1).

Earlier MD simulations of ligand-free CAR indicated that the Phe161 side chain occupies S1 as a result of an X_1 torsion angle rotation.[42] By contrast, in X-ray structures of human CAR, the cocrystallized agonists CITCO and 5β-pregnane-3,20-dione either block or occupy the S1 pocket.[19] As a result, Phe161 is prevented from occupying S1. Instead, it is located at the LBD/H12 interface, in contact with amino acids from H3 (Asn165), the loop connecting H6 and H7 (Phe234), H11 (Tyr326), as well as the H11′−H12 region (Leu336, Met340). As described above, site-directed mutagenesis confirmed Phe161 as an important factor for ligand-elicited CAR activation, as most compounds were unable to activate the mutant (Table 4).

While almost all molecules target the positive region I, positive areas II and III are less frequently occupied (Table 2). Because these particular fields are framed by five (Phe132, Ile164, Asn165, Met168, Leu206) or three (Val199, Cys202, His203) amino acids, respectively, they correspond to significantly smaller clefts than the subsite S1. Therefore, access of ligands to both areas is spatially much more restricted and strongly dependent on structural properties of the compound. Although molecules contacting either region activate CAR to a greater extent than the average agonist activity (Table 2), the minor impact of regions II and III is evident from the relatively modest effects by mutations of Ile164 and His203, respectively. These residues do not form a part of the LBD/H12 interface, and although the overall extent of CAR activation was reduced, both mutants were still activated >2-fold by 9 and 13 compounds, respectively.

Some ligands were found to occupy the small channel C1 (e.g., **3**, **6**, **7**) that connects the LBP with the H12 (Figures 3 and 4). This channel provides sufficient space to accommodate even substituted aromatic rings. GRID fields and the positive coefficients areas are in good agreement with aromatic functions located therein (Figure 3). However, the most active compounds (>4-fold activation) usually exhibit no direct contact with the activation helix H12 (e.g., **1**, **2**, **17**, **18**; average minimum distance of 6.4 Å). In contrast, less active molecules (<2-fold activation) are located closer to H12 (e.g., **14**−**16**, **26**; average minimum distance of 4.6 Å).

Table 4. Summary of Activation Data for Wild-Type and Mutated Human CAR by 27 Novel Agonists[a]

classified activity[b]	human CAR	Phe161Ala	Ile164Ala	Asn165Ala	His203Ala	Phe234Ala	Phe243Ala	Tyr326Ala
>6-fold (strong)	1, 17 [2]		18, 20 [2]		**19** [*1*]	1 [*1*]	**1**, **4**, **6**, **7**, **9**, **10**, 17 [7]	
3.5- to 6-fold (moderate)	2−6, 18−21 [9]	19, 20 [2]	2, 4, 19 [3]	4, 21 [2]	1, 6, 20, **23** [*4*]	19 [*1*]	2, **8**, **12**, **16**, 18, 19, 20, **22** [*8*]	1 [*1*]
2- to 3.5-fold (significant)	7−13, 22, 23 [9]	10 [*1*]	8, 10, 13, 15 [*4*]	9, 19, 20 [3]	2, 4, 5, 8, 11, 13, **15**, 17 [*8*]	20 [*1*]	3, 5, 11, **13-15**, 21, 23, **24−27** [*12*]	19 [*1*]
1.3- to 2-fold (marginal)	14−16, 24−27 [7]	1, 3, 4, 16, 18, 24, 25 [7]	1, 3, 5−7, 9, 12, 14, 17, 21, 22, 24, 26 [*13*]	6, 12, 14, 22−24, 26 [7]	7, 9, 10, 12, 14, 16, 18 21, 24, 26 [*10*]	4, 6, 8, 14, 25, 26 [6]		2−4, 6, 8, 12, 14, 16, 20, 23, 24 [*11*]
<1.3-fold (no response)		2, 5−9, 11−15, 17, 21−23, 26, 27 [*17*]	11, 16, 23, 25, 27 [5]	1−3, 5, 7, 8, 10, 11, 13, 15−18, 25, 27 [*15*]	3, 22, 25, 27 [*4*]	2, 3, 5, 7, 9, 10−13, 15−18, 21−24, 27 [*18*]		5, 7, 9, 10, 11, 13, 15, 17, 18, 21−22, 25−27 [*14*]

[a] The italicized numbers in brackets indicate the number of compounds within a group. Those ligands that reach or surpass the absolute activation level of wild-type human CAR in the mutants are marked in bold. [b] The ligand-induced activities are expressed as relative to the vehicle (DMSO) control of the respective construct and classified into five groups. See the original data for these ligands and reference compounds in the Supporting Information Table S1.

Tyr326 directly interacts with Leu343 of H12 and is believed to represent a key element of basal activity.[19,42] In agonist-bound CAR X-ray structures, the Tyr326 side chain is stabilized by a hydrogen bond formed with Asn165. Furthermore, this side chain is contacted by Phe234 and Phe161. The latter residue also interacts with Asn165, probably stabilizing the hydrogen bond between Asn165 and Tyr326. Mutation of Phe161, Asn165, or Phe234 is expected to destabilize the Tyr326 side chain and to diminish its contacts with H12. Data from site-directed mutagenesis for all four amino acids correlate to a large extent, in particular, for Phe234 and Tyr326 (Table 4, Supporting Information Table S1).

The single negative coefficient region IV is located at the very end of channel C1, and it corresponds to the LBD/H12 interface. Because the negative region partly overlaps with vdW radii of amino acids Val199 (H5) and Leu343 (H12), full occupation of region IV by a ligand may result in unfavorable steric clashes that may, in the extreme, culminate in displacing H12 from its active conformation and prevent the association of the LBD with coactivators. This view gains support from the finding that out of 13 compounds occupying region IV, 6 were able to recruit the corepressor NCoR and the remaining 7 were unable to further enhance the association between SRC1 and the CAR LBD (Table 2, Supporting Information Figure S1). Ligands that only partially occupy region IV (e.g., **16** and **26**) generally have lower activating effects (Tables 1 and 3, Figure 2C). The putative mechanisms by which these ligands may retain some CAR activation include provision of additional vdW contacts between the ligand and aliphatic amino acid residues in H12, in particular Leu343, and stabilization of the Tyr326 side chain.

Discussion

In the present study we have successfully applied both ligand- and protein-based approaches in the search for novel CAR agonists. Although several compounds activate CAR less than 2-fold, the overall extent of activation is comparable to activity measurements of other CAR agonists.[4,43,44] Compared to the protein-based screening, the BRUTUS approach had a 2-fold higher success rate (45.7% vs 20.4%) and revealed additional scaffolds besides the substituted sulfonamides, urea derivatives, and some thiazolidin-4-ones found by both screening approaches. The performance of the protein-based approach strongly depends on the quality of the target structure; in this respect, the torsion angles of amino acid side chains defining the LBP are particularly critical. In prior analyses of this quality, we found a good agreement between our homology model and the available X-ray structures.[42]

On the basis of the newly identified CAR ligands, we present the successful generation of a robust 3D QSAR model. To our knowledge, this is the first 3D QSAR study performed for a NR that is based exclusively on compounds identified by a VS approach. In spite of the substantial structural diversity and varying binding modes of the ligands, a good correlation between experimental and predicted activity is observed. Often, the biological data are not of sufficiently high quality to generate a valid 3D QSAR model. Before this study, generation of 3D QSAR models for CAR agonists was impeded by the lack of known ligands and problems in measuring the CAR activity.[20] The alignment procedure used to superpose the agonists also has consequences on the quality of the resulting model. A receptor-based alignment has been previously conducted for estrogen and progesterone receptor ligands.[45,46] The GRID/GOLPE approach has already been successfully applied for the

generation of 3D QSAR models for murine CAR inhibitors and estrogen receptor agonists.[15,45] Related methods such as CoMFA have been used for peroxisome proliferators-activated receptors α and γ, bile acid receptor, and androgen receptor.[47−49] Although NRs are of high pharmaceutical interest and hundreds of X-ray and NMR structures are available, only few successful VS reports have been published, for example, for estrogen receptor modulators[26,50,51] or, very recently, for a potent pregnane X receptor agonist.[52]

Our results give new insight into molecular processes of ligand-dependent CAR activation. In classical NRs, agonists usually directly contact H12 which is generally believed being responsible for keeping H12 attached to the LBD in the active conformation.[53,54] In contrast, possibilities for agonists to interact with H12 in CAR are limited due to the amino acids almost completely separating the LBP from H12. Only a small channel (C1) connects the LBP with H12, and therefore, a different mechanism of activation for most ligands must be involved.

Our results suggest that efficient CAR activation requires the ligand to enter the subsite S1, which is likely occupied by Phe161 in ligand-free CAR. Thus, the Phe161 side chain is forced to rotate toward the H11′ region and vdW contacts with Leu336 and Met340 are established. Moreover, the Phe161 side chain interacts with Asn165 and Tyr326 that are integral residues in keeping H12 in its active position. We propose that novel contacts with Leu336 and Met340 are introduced and existing interactions of Asn165 and Tyr326 with H12 are enhanced by restrictions on side chain movements. This is supported by our findings that mutation of Tyr326 and Met340 abolishes the agonist-dependent CAR activity[18] and that SRC1 interactions are enhanced by S1-occupying agonists.

Generally, NR ligands optimally fit in their binding pocket, as the LBP is shaped accordingly (e.g., as for vitamin D and its receptor).[53] The xenobiotic sensors CAR and PXR possess large, spherical, and mostly hydrophobic LBPs, thereby enabling binding of chemically diverse compounds, albeit with lower affinity compared with steroid hormone receptors.[55] As indicated here, ligands that enter certain clefts within the LBP may not only induce the conformational change of Phe161 as described above but also bind more stably and better promote CAR activation when compared to molecules unable to occupy these crevices. Because CAR ligand binding is governed almost exclusively by vdW interactions due to limited number of polar LBP residues, these subsites represent relatively unselective anchor points for a broad spectrum of ligands.

Compounds with lower activation generally do not occupy the subsite S1. It is not yet known whether their modest CAR activity relies on (more subtle) stabilizing effects of H12-contacting amino acids or on a different mechanism. By not occupying S1, those compounds might bind less stably to the LBP, and therefore, any effect on H12 or the interacting amino acids (Asn165, Tyr326) might be less pronounced. When compared to the more active compounds, these weaker agonists are located considerably closer to the LBD/H12 interface and tend to occupy the C1 channel. It is possible that direct interactions of ligands with Asn165 and Tyr326 may activate CAR slightly, but also direct ligand contacts with H12, as commonly observed in other NR/ligand complexes,[53,54] may contribute. First, CAR activation is sensitive to mutations at the LBD/H12 interface, where even small amino acid exchanges (e.g., Val199Ala) completely abolished the CAR activity.[18] Therefore, stabilization of the hydrophobic contact surface for H12 may improve its attachment to the LBD, resulting in increased CAR activity. Second, an appropriately positioned aromatic or aliphatic moiety in C1 may enlarge the hydrophobic contact surface for H12, and interactions with hydrophobic amino acids on H12 (Leu343) might be improved. Our hypothesis is supported by findings that clotrimazole, which is buried deep in the LBP, is unable to activate CAR carrying a Tyr326Ala substitution, whereas TMPP contacting Leu343 at least partially activates this mutant receptor.[18,29] On the other hand, association of a ligand with region IV attenuates human CAR activity, most likely by displacement of H12 from its active position.

Although these novel human CAR agonists may serve as templates in the development of drugs for hyperbilirubinemia or cholestasis,[56−58] the main emphasis of this work was to provide a model describing the key molecular aspects of human CAR activation. Understanding the structural basis for ligand-dependent CAR activation is of high importance in drug development. A predictive model may be applied to reduce or avoid CAR-activating potential in drug candidates and to increase drug efficacy as well as to reduce the risk of drug/drug interactions. Moreover, many compounds positive in the rodent carcinogenicity bioassay are capable of activating CAR, causing cell proliferation and decrease in apoptosis.[9,59,60] Even though there is little evidence of carcinogenicity of phenobarbital (an indirect activator of human and rodent CAR) in humans,[61] the 3D QSAR model and in vitro CAR assay developed here provide improved screening tools to better assess the toxicological role of human CAR activation.

Conclusions

Virtual screening with ligand- and protein structure-based methods provided 27 novel agonists for the human nuclear receptor CAR, an important regulator of drug and xenobiotic metabolism, liver growth, and pathogenesis. A 3D QSAR based on these agonists, and analysis of molecular fields indicated key regions within the CAR LBD that control the agonist activity. Analysis of ligand binding modes, site-directed mutagenesis, and assays of coactivator and corepressor recruitment provided clues to main mechanisms of activation by human CAR. Importantly, occupancy of agonists within subsite S1 forces the reorientation of a key residue Phe161 and enhancement of interactions between LBD and the activation helix 12. The 3D QSAR model may help in identifying human CAR agonists among compounds in drug discovery libraries or those undergoing toxicological risk evaluations.

Experimental Section

3D Database Search. The Tripos LeadQuest database containing approximately 85 000 compounds was used as the starting point for the 3D database search. The homology model of human CAR was deployed as input structure for the search because CAR X-ray structures were not available at that time. The details of model generation, docking of known agonists, as well as a comparison with X-ray structures published later have been previously described.[29,41]

A search query within the UNITY module of SYBYL (Tripos Inc., St. Louis, MO) was defined consisting of the Connolly surface[62] of the LBP and two sites of hydrophobic character. The tolerance of the surface was set to 0.5 Å. The selection of the hydrophobic features with a tolerance of 1.0 Å was based on both vdW interactions between the docked ligands clotrimazole and TMPP and the LBP as well as favorable sites of lipophilic interactions within the LBP detected by calculating the lipophilic potential.[63] Compounds violating Lipinski's "rule of five" were excluded from the search.[64]

Ligand-based search was based on BRUTUS software[27,28] with CITCO and TMPP as templates. A set of 25 conformations for

CITCO were generated using Confort 7.0.1 within SYBYL. For TMPP, Confort did not work and we used instead the systematic search function within SYBYL. The search procedure resulted in over 2000 conformations that were classified using an in-house field-based classification method by which 90 unique conformations of TMPP were retrieved. All the selected conformations of TMPP and CITCO were aligned and ranked by BRUTUS, and the top-ranking conformations were evaluated by aligning them onto several other CAR ligands.[4] One independent conformation of CITCO and TMPP and one conformation of TMPP retrieved from molecular docking experiment were finally selected as templates for BRUTUS field-based virtual screening. From all of these analyses, 200 best compounds were retrieved and field similarities were further analyzed with self-organizing maps (Visual Data) as described.[27]

Molecular Docking. All compounds emerging from the 3D database search were subjected to a molecular docking approach using GOLD, version 2.2 (CCDC, Cambridge, U.K.) with GoldScore as the scoring function.[65] All molecules were docked within a sphere of 20 Å radius around atom CD1 of amino acid Leu206. For each compound, a maximum of 10 conformations were allowed. Ligand docking was terminated once the first three emerging conformations were all within rmsd of 1.5 Å when compared to each other.

In order to take the coherence of the GoldScore and the molecular weight into account, the GoldScore value was modified by dividing the original score over the square-root of the number of heavy atoms.[66] Additionally, all conformations obtained from the GOLD docking were rescored by applying an external scoring function X-Score.[36] These results were also modified.[66] Finally, the 50 best ranked conformations resulting from modified GoldScore and X-Score were redocked and rescored using the Lamarckian genetic algorithm implemented in AutoDock, version 3.0.5.[67] From each docking/scoring approach, 20 top-ranked results were selected. Altogether, 64 compounds were chosen for further investigations of which 54 were commercially available.

3D QSAR. In this study, 3D QSAR is used to quantitatively explain structure−activity relationships of human CAR agonists, thus allowing the visualization of biologically important structural properties. Structures of the ligands were generated within SYBYL and minimized using the Tripos force field and Gasteiger−Hückel charges. Subsequently, all molecules were docked into the CAR LBP using GOLD with settings as described before and afterward rescored using X-Score. The top-scoring docking poses determined by X-Score were selected and minimized within the CAR LBP as described above. As the measured ligand-induced receptor activities already represent fairly symmetrically distributed data, no further transformation of the data was required.[68]

For the 3D QSAR, the GRID/GOLPE approach[69] was used with GRID, version 22 (Molecular Discovery, Pinner, U.K.), and GOLPE, version 4.5 (Multivariate Informetric Analysis, Perugia, Italy). The GRID/GOLPE approach correlates molecular interaction fields with biological activities using a multivariate partial least squares (PLS) method. Here, aromatic (C1═) and phenolic (OH) probe−ligand interactions fields are correlated with ligand-elicited human CAR activity values. The grid spacing was 1.0 Å, and the standard deviation threshold for exclusion of data columns from the PLS analysis was set at 0.1 kcal/mol. All two- and three-level variables were removed, and values between −0.05 and 0.5 were set at zero. Smart region definition with fractional factorial design variable selection[38] was used with two PLS components in the selection and validation phase. The PLS validation using 5 random groups was repeated 100 times. The maximum number of PLS components was kept at 2 in order to ensure simplicity of the model. Compared to the "leave-one-out" procedure, the "leave-20%-out" method has been shown to yield better indices for the robustness of a model.[45,70] From cross-validation predictions, a q^2-value above 0.5 indicates an internally consistent and valid model.[71]

Principal Component Analysis. The principal component analysis was used to identify functional similarities and differences of tested mutations. Mathematically, this method is defined as an orthogonal linear transformation that transforms the data to a new coordinate system such that the greatest variance by any projection of the data comes to lie on the first coordinate (called the first principal component), the second greatest variance on the second coordinate, and so on.[71] Calculations were performed using the SimcaP-10.0.0 software (Umetrics, Inc., Kinnelon). The data matrix consisted of measured human CAR activity for each of the tested mutations.

Chemicals. All steroids were bought from Steraloids, Inc. (Newport, RI). TMPP and TPP were synthesized as described,[31] and clotrimazole and CITCO were from Sigma (St. Louis, MO) and Biomol (Plymouth, PA), respectively. Chemicals ordered from Tripos were diluted in DMSO, and their activities were measured at 10 μM.

Plasmids and Human CAR Activation Assays. The CMX-GAL4-human CAR LBD (residues 108−348) plasmid and its various mutants, the GAL4-dependent luciferase reporter, and the pCMVβ control plasmids have been described.[18,21] The mutated and wild-type CAR receptors were expressed at similar levels in mammalian cells (Jyrkkärinne et al., unpublished experiments). All plasmids were purified with Qiagen columns (Hilden, Germany). Human hepatoma C3A cells (ATCC CRL-10741) were cultured overnight in 48-well plates to 50% confluence in phenol red-free DMEM medium (Invitrogen, Gaithersburg, MD) containing 10% fetal bovine serum (Bio-Whittaker, Belgium), 1% L-glutamine (Euroclone, Pero (Milano), Italy), 100 U/mL penicillin plus 100 μg/mL streptomycin (Euroclone). The cells were then transfected with the plasmids CMX-GAL4-LBD (450 ng/well), luciferase reporter (300 ng/well), and pCMVβ (600 ng/well) using the calcium phosphate method. After transfection, the medium was replaced with fresh medium that was supplemented with 5% delipidated serum (HyClone, Logan, UT) with 10 μM EE2 to suppress the high CAR basal activity and with the DMSO vehicle (0.1% v/v) or the test chemical (10 μM). We used clotrimazole (4 μM), TMPP (10 μM), and TPP (10 μM) as reference compounds[18] in all assays because the most efficacious activator CITCO proved to be unstable upon storage. After 24 h, the treated cells were lysed and the luciferase and β-galactosidase activities were measured.[32] All luciferase activities were normalized to β-galactosidase activities, and the results are expressed as the mean ± standard deviation (SD) of at least three independent experiments. The assay was validated with clotrimazole as the positive control substance, and performance parameters Z', the coefficient of variation, and the minimum significant ratio were calculated according to Iversen et al.[33] The ligand-dependent association of CAR with the human coactivator SRC1 was measured in C3A cells with expression and reporter plasmids described elsewhere,[72] and the assay for human corepressor NCoR recruitment was done as reported.[21]

Acknowledgment. This work was supported by grants from the Academy of Finland and the Finnish Funding Agency for Technology and Innovation. The research groups of A.P. and P.H. are associated with Biocenter Kuopio.

Supporting Information Available: Data on activation of wild-type and mutant human CAR receptors and recruitment of coregulators SRC1 and NCoR by 27 novel agonists and reference compounds in Table S1 and Figure S1, respectively, in the first Supporting Information file; individual HPLC−MS run outputs for the agonists, provided by the chemical supplier, in the second Supporting Information file. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Gronemeyer, H.; Gustafsson, J. Å.; Laudet, V. Principles for modulation of the nuclear receptor superfamily. *Nat. Rev. Drug Discovery* **2004**, *3*, 950–964.

(2) Rosenfeld, M. G.; Glass, C. K. Coregulator codes of transcriptional regulation by nuclear receptors. *J. Biol. Chem.* **2001**, *276*, 36865–36868.

(3) Nettles, K. W.; Greene, G. L. Ligand control of coregulator recruitment to nuclear receptors. *Annu. Rev. Physiol.* **2005**, *67*, 309–333.

(4) Stanley, L. A.; Horsburgh, B. C.; Ross, J.; Scheer, N.; Wolf, C. R. PXR and CAR: nuclear receptors which play a pivotal role in drug disposition and chemical toxicity. *Drug Metab. Rev.* **2006**, *38*, 515–597.

(5) Timsit, Y. E.; Negishi, M. CAR and PXR: the xenobiotic-sensing receptors. *Steroids* **2007**, *72*, 231–246.

(6) Kodama, S.; Koike, C.; Negishi, M.; Yamamoto, Y. Nuclear receptors CAR and PXR cross talk with FOXO1 to regulate genes that encode drug-metabolizing and gluconeogenic enzymes. *Mol. Cell. Biol.* **2004**, *24*, 7931–7940.

(7) Nakamura, K.; Moore, R.; Negishi, M.; Sueyoshi, T. Nuclear pregnane X receptor cross-talk with FoxA2 to mediate drug-induced regulation of lipid metabolism in fasting mouse liver. *J. Biol. Chem.* **2007**, *282*, 9768–9776.

(8) Wei, P.; Zhang, J.; Egan-Hafley, M.; Liang, S.; Moore, D. D. The nuclear receptor CAR mediates specific xenobiotic induction of drug metabolism. *Nature* **2000**, *407*, 920–923.

(9) Yamamoto, Y.; Moore, R.; Goldsworthy, T. L.; Negishi, M.; Maronpot, R. R. The orphan nuclear receptor constitutive active/androstane receptor is essential for liver tumor promotion by phenobarbital in mice. *Cancer Res.* **2004**, *64*, 7197–7200.

(10) Columbano, A.; Ledda-Columbano, G. M.; Pibiri, M.; Cossu, C.; Menegazzi, M.; Moore, D. D.; Huang, W.; Tian, J.; Locker, J. Gadd45beta is induced through a CAR-dependent, TNF-independent pathway in murine liver hyperplasia. *Hepatology* **2005**, *42*, 1118–1126.

(11) Kretschmer, X. C.; Baldwin, W. S. CAR and PXR: xenosensors of endocrine disrupters. *Chem. Biol. Interact.* **2005**, *155*, 111–128.

(12) Chang, T. K. H.; Waxman, D. J. Synthetic drugs and natural products as modulators of constitutive androstane receptor (CAR) and pregnane X receptor (PXR). *Drug Metab. Rev.* **2006**, *38*, 51–73.

(13) Moore, L. B.; Parks, D. J.; Jones, S. A.; Bledsoe, R. K.; Consler, T. G.; Stimmel, J. B.; Goodwin, B.; Liddle, C.; Blanchard, S. G.; Willson, T. M.; Collins, J. L.; Kliewer, S. A. Orphan nuclear receptors constitutive androstane receptor and pregnane X receptor share xenobiotic and steroid ligands. *J. Biol. Chem.* **2000**, *275*, 15122–15127.

(14) Mäkinen, J.; Reinisalo, M.; Niemi, K.; Viitala, P.; Jyrkkärinne, J.; Chung, H.; Pelkonen, O.; Honkakoski, P. Dual action of oestrogens on the mouse constitutive androstane receptor. *Biochem. J.* **2003**, *376*, 465–472.

(15) Jyrkkärinne, J.; Mäkinen, J.; Gynther, J.; Savolainen, H.; Poso, A.; Honkakoski, P. Molecular determinants of steroid inhibition for the mouse constitutive androstane receptor. *J. Med. Chem.* **2003**, *46*, 4687–4695.

(16) Wilson, T. M.; Kliewer, S. A. PXR, CAR and drug metabolism. *Nat. Rev. Drug Discovery* **2002**, *1*, 259–266.

(17) Lin, J. H. CYP induction-mediated drug interactions: in vitro assessment and clinical implications. *Pharm. Res.* **2006**, *23*, 1089–1116.

(18) Jyrkkärinne, J.; Windshügel, B.; Mäkinen, J.; Ylisirniö, M.; Peräkylä, M.; Poso, A.; Sippl, W.; Honkakoski, P. Amino acids important for ligand specificity of the human constitutive androstane receptor. *J. Biol. Chem.* **2005**, *280*, 5960–5971.

(19) Xu, R. X.; Lambert, M. H.; Wisely, B. B.; Warren, E. N.; Weinert, E. E.; Waitt, G. M.; Williams, J. D.; Collins, J. L.; Moore, L. B.; Willson, T. M.; Moore, J. T. A structural basis for constitutive activity in the human CAR/RXRα heterodimer. *Mol. Cell* **2004**, *16*, 919–928.

(20) Poso, A.; Honkakoski, P. Ligand recognition by drug-activated nuclear receptors PXR and CAR: structural, site-directed mutagenesis and molecular modeling studies. *Mini-Rev. Med. Chem.* **2006**, *6*, 937–947.

(21) Mäkinen, J.; Frank, C.; Jyrkkärinne, J.; Gynther, J.; Carlberg, C.; Honkakoski, P. Modulation of mouse and human phenobarbital-responsive enhancer module by nuclear receptors. *Mol. Pharmacol.* **2002**, *62*, 366–378.

(22) Toell, A.; Kroncke, K. D.; Kleinert, H.; Carlberg, C. Orphan nuclear receptor binding site in the human inducible nitric oxide synthase promoter mediates responsiveness to steroid and xenobiotic ligands. *J. Cell. Biochem.* **2002**, *85*, 72–82.

(23) Smith, C. L.; Nawaz, Z.; O'Malley, B. W. Coactivator and corepressor regulation of the agonist/antagonist activity of the mixed antiestrogen, 4-hydroxytamoxifen. *Mol. Endocrinol.* **1997**, *11*, 657–666.

(24) Liu, Z.; Auboeuf, D.; Wong, J.; Chen, J. D.; Tsai, S. Y.; Tsai, M. J.; O'Malley, B. W. Coactivator/corepressor ratios modulate PR-mediated transcription by the selective receptor modulator RU486. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 7940–7944.

(25) Stoermer, M. J. Current status of virtual screening as analysed by target class. *Med. Chem.* **2006**, *2*, 89–112.

(26) Schapira, M.; Abagyan, R.; Totrov, M. Nuclear hormone receptor targeted virtual screening. *J. Med. Chem.* **2003**, *46*, 3045–3059.

(27) Tervo, A. J.; Rönkkö, T.; Nyrönen, T. H.; Poso, A. BRUTUS: optimization of a grid-based similarity function for rigid-body molecular superposition. 1. Alignment and virtual screening applications. *J. Med. Chem.* **2005**, *48*, 4076–4086.

(28) Rönkkö, T.; Tervo, A. J.; Parkkinen, J.; Poso, A. BRUTUS: optimization of a grid-based similarity function for rigid-body molecular superposition. II. Description and characterization. *J. Comput.-Aided Mol. Des.* **2006**, *20*, 227–236.

(29) Windshügel, B.; Jyrkkärinne, J.; Poso, A.; Honkakoski, P.; Sippl, W. Molecular dynamics simulations of the human CAR ligand-binding domain: deciphering the molecular basis for constitutive activity. *J. Mol. Model.* **2005**, *11*, 69–79.

(30) Kawamoto, T.; Kakizaki, S.; Yoshinari, K.; Negishi, M. Estrogen activation of the nuclear orphan receptor CAR (constitutive active receptor) in induction of the mouse *Cyp2b10* gene. *Mol. Endocrinol.* **2000**, *14*, 1897–1905.

(31) Honkakoski, P.; Palvimo, J. J.; Penttilä, L.; Vepsäläinen, J.; Auriola, S. Effects of triaryl phosphates on mouse and human nuclear receptors. *Biochem. Pharmacol.* **2004**, *67*, 97–106.

(32) Honkakoski, P.; Jääskeläinen, I.; Kortelahti, M.; Urtti, A. A novel drug-regulated gene expression system based on the nuclear receptor constitutive androstane receptor (CAR). *Pharm. Res.* **2001**, *18*, 146–150.

(33) Iversen, P. W.; Eastwood, B. J.; Sittampalam, G. S.; Cox, K. L. A comparison of assay performance measures in screening assays: signal window, Z′ factor, and assay variability ratio. *J. Biomol. Screening* **2006**, *11*, 247–252.

(34) Routledge, E. J.; White, R.; Parker, M. G.; Sumpter, J. P. Differential effects of xenoestrogens on coactivator recruitment by estrogen receptor (ER) alpha and ERbeta. *J. Biol. Chem.* **2000**, *275*, 35986–25993.

(35) Bramlett, K. S.; Wu, Y.; Burris, T. P. Ligands specify coactivator nuclear receptor (NR) box affinity for estrogen receptor subtypes. *Mol. Endocrinol.* **2001**, *15*, 909–922.

(36) Wang, R.; Lai, S.; Wang, S. Further development and validation of empirical scoring functions for structure-based binding affinity prediction. *J. Comput.-Aided Mol. Des.* **2002**, *16*, 11–26.

(37) Goodford, P. J. A computational procedure for determining energetically favorable binding sites on biologically important macromolecules. *J. Med. Chem.* **1985**, *28*, 849–857.

(38) Pastor, M.; Cruciani, G. Smart region definition: a new way to improve the predictive ability and interpretability of three-dimensional quantitative structure−activity relationships. *J. Med. Chem.* **1997**, *40*, 1455–1464.

(39) Migliavacca, E. Applied introduction to multivariate methods used in drug discovery. *Mini-Rev. Med. Chem.* **2003**, *3*, 831–843.

(40) Steindl, T. M.; Crump, C. E.; Hayden, F. G.; Langer, T. Pharmacophore modeling, docking, and principal component analysis based clustering: combined computer-assisted approaches to identify new inhibitors of the human rhinovirus coat protein. *J. Med. Chem.* **2005**, *48*, 6250–6260.

(41) Matero, S.; Lahtela-Kakkonen, M.; Korhonen, O.; Ketolainen, J.; Lappalainen, R.; Poso, A. Chemical space of orally active compounds. *Chemom. Intell. Lab. Syst.* **2006**, *84*, 134–141.

(42) Windshügel, B.; Jyrkkärinne, J.; Vanamo, J.; Poso, A.; Honkakoski, P.; Sippl, W. Comparison of homology models and X-ray structures of the nuclear receptor CAR: assessing the structural basis of constitutive activity. *J. Mol. Graphics Modell.* **2007**, *25*, 644–657.

(43) Maglich, J. M.; Parks, D. J.; Moore, L. B.; Collins, J. L.; Goodwin, B.; Billin, A. N.; Stoltz, C. A.; Kliewer, S. A.; Lambert, M. H.; Willson, T. M.; Moore, J. T. Identification of a novel human constitutive androstane receptor (CAR) agonist and its use in the identification of CAR target genes. *J. Biol. Chem.* **2003**, *278*, 17277–17283.

(44) Kobayashi, K.; Yamanaka, Y.; Iwazaki, N.; Nakajo, I.; Hosokawa, M.; Negishi, M.; Chiba, K. Identification of HMG-CoA reductase inhibitors as activators for human, mouse and rat constitutive androstane receptor. *Drug Metab. Dispos.* **2005**, *33*, 924–929.

(45) Sippl, W. Receptor-based 3D QSAR analysis of estrogen receptor ligands−merging the accuracy of receptor-based alignments with the computational efficiency of ligand-based methods. *J. Comput.-Aided Mol. Des.* **2000**, *14*, 559–572.

(46) Söderholm, A. A.; Lehtovuori, P. T.; Nyrönen, T. H. Docking and three-dimensional quantitative structure−activity relationship (3D QSAR) analyses of nonsteroidal progesterone receptor ligands. *J. Med. Chem.* **2006**, *49*, 4261–4268.

(47) Khanna, S.; Sobhia, M. E.; Bharatam, P. V. Additivity of molecular fields: CoMFA study on dual activators of PPARalpha and PPARgamma. *J. Med. Chem.* **2005**, *48*, 3015–3025.

(48) Honorio, K. M.; Garratt, R. C.; Polikarpov, I.; Andricopulo, A. D. 3D QSAR comparative molecular field analysis on nonsteroidal farnesoid X receptor activators. *J. Mol. Graphics Modell.* **2007**, *25*, 921–927.

(49) Tamura, H.; Ishimoto, Y.; Fujikawa, T.; Aoyama, H.; Yoshikawa, H.; Akamatsu, M. Structural basis for androgen receptor agonists and antagonists: interaction of SPEED 98-listed chemicals and related compounds with the androgen receptor based on an in vitro reporter gene assay and 3D QSAR. *Bioorg. Med. Chem.* **2006**, *14*, 7160–7174.
(50) Yang, J. M.; Shen, T. W. A pharmacophore-based evolutionary approach for screening selective estrogen receptor modulators. *Proteins* **2005**, *59*, 205–220.
(51) Wang, C. Y.; Ai, N.; Arora, S.; Erenrich, E.; Nagarajan, K.; Zauhar, R.; Young, D.; Welsh, W. J. Identification of previously unrecognized antiestrogenic chemicals using a novel virtual screening approach. *Chem. Res. Toxicol.* **2006**, *19*, 1595–1601.
(52) Lemaire, G.; Benod, C.; Nahoum, V.; Pillon, A.; Boussioux, A. M.; Guichou, J. F.; Subra, G.; Pascussi, J. M.; Bourguet, W.; Chavanieu, A.; Balaguer, P. Discovery of a highly active ligand of human pregnane X receptor: a case study from pharmacophore modeling and virtual screening to "in vivo" biological activity. *Mol. Pharmacol.* **2007**, *72*, 572–581.
(53) Rochel, N.; Wurtz, J. M.; Mitschler, A.; Klaholz, B.; Moras, D. The crystal structure of the nuclear receptor for vitamin D bound to its natural ligand. *Mol. Cell* **2000**, *5*, 173–179.
(54) Watkins, R. E.; Wisely, G. B.; Moore, L. B.; Collins, J. L.; Lambert, M. H.; Williams, S. P.; Willson, T. M.; Kliewer, S. A.; Redinbo, M. R. The human nuclear xenobiotic receptor PXR: structural determinants of directed promiscuity. *Science* **2001**, *292*, 2329–2333.
(55) Honkakoski, P.; Sueyoshi, T.; Negishi, M. Drug-activated nuclear receptors CAR and PXR. *Ann. Med.* **2003**, *35*, 172–182.
(56) Guo, G. L.; Lambert, G.; Negishi, M.; Ward, J. M., Jr.; Kliewer, S. A.; Gonzalez, F. J.; Sinal, C. J. Complementary roles of farnesoid X receptor, pregnane X receptor, and constitutive androstane receptor in protection against bile acid toxicity. *J. Biol. Chem.* **2003**, *278*, 45062–45071.
(57) Wagner, M.; Halilbasic, E.; Marschall, H. U.; Zollner, G.; Fickert, P.; Langner, C.; Zatloukal, K.; Denk, H.; Trauner, M. CAR and PXR agonists stimulate hepatic bile acid and bilirubin detoxification and elimination pathways in mice. *Hepatology* **2005**, *42*, 420–430.
(58) Zollner, G.; Marschall, H. U.; Wagner, M.; Trauner, M. Role of nuclear receptors in the adaptive response to bile acids and cholestasis: pathogenetic and therapeutic considerations. *Mol. Pharmaceutics* **2006**, *3*, 231–251.
(59) Huang, W.; Zhang, J.; Washington, M.; Liu, J.; Parant, J. M.; Lozano, G.; Moore, D. D. Xenobiotic stress induces hepatomegaly and liver tumors via the nuclear receptor constitutive androstane receptor. *Mol. Endocrinol.* **2005**, *19*, 1646–1653.
(60) Köhle, C.; Schwarz, M.; Bock, K. W. Promotion of hepatocarcinogenesis in humans and animal models. *Arch. Toxicol.* **2008**, *82*, 623–631.
(61) Whysner, J.; Ross, P. M.; Williams, G. M. Phenobarbital mechanistic data and risk assessment: enzyme induction, enhanced cell proliferation, and tumor promotion. *Pharmacol. Ther.* **1996**, *71*, 153–191.
(62) Connolly, M. L. Solvent-accessible surfaces of proteins and nucleic acids. *Science* **1983**, *221*, 709–713.
(63) Ghose, A. K.; Viswanadhan, V. N.; Wendoloski, J. J. Prediction of hydrophobic (lipophilic) properties of small organic molecules using fragmental methods: an analysis of ALOGP and CLOGP methods. *J. Phys. Chem. A* **1998**, *102*, 3762–3772.
(64) 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 Delivery Rev.* **2001**, *46*, 3–26.
(65) Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol.* **1997**, *267*, 727–748.
(66) Pan, Y.; Huang, N.; Cho, S.; Mackerell, A. D. Consideration of molecular weight during compound selection in virtual target-based database screening. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 267–272.
(67) Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. Automated docking using a Lamarckian genetic algorithm and and empirical binding free energy function. *J. Comput. Chem.* **1998**, *19*, 1639–1662.
(68) Wold, S.; Johansson, E.; Cocchi, M. PLS—Partial Least Squares Projections to Latent Structures. In *3D QSAR in Drug Design*; Kubinyi, H., Ed.; ESCOM Science Publishers B.V.: Leiden, The Netherlands, 1993; pp 523−550.
(69) Baroni, M.; Constantino, G.; Cruciani, G.; Riganelli, D.; Valigi, R.; Clementini, S. Generating optimal linear PLS estimations (GOLPE): an advanced chemometric tool for handling 3D QSAR. *Quant. Struct.−Act. Relat.* **1993**, *12*, 9–20.
(70) Clark, M.; Cramer, R. D. The probability of chance correlation using partial least squares. *Quant. Struct.−Act. Relat.* **1993**, *12*, 137–145.
(71) Jolliffe, I. T. *Principal Component Analysis*; Springer: New York, 1992.
(72) Küblbeck, J.; Jyrkkärinne, J.; Poso, A.; Turpeinen, M.; Sippl, W.; Honkakoski, P.; Windshügel, B. *Biochem. Pharmacol.* **2008**, *76*, 1288−1297.

JM800731B