Discovery of Novel Agonists and Antagonists of the Free Fatty Acid Receptor 1 (FFAR1) Using Virtual Screening

Irina G. Tikhonova,[†,§] Chi Shing Sum,[‡,§] Susanne Neumann,[‡] Stanislav Engel,[‡] Bruce M. Raaka,[‡] Stefano Costanzi,*[,†] and Marvin C. Gershengorn*[,‡]

Laboratory of Biological Modeling, and Clinical Endocrinology Branch, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892

Received October 1, 2007

The G-protein-coupled receptor free fatty acid receptor 1 (FFAR1), previously named GPR40, is a possible novel target for the treatment of type 2 diabetes. In an attempt to identify new ligands for this receptor, we performed virtual screening (VS) based on two-dimensional (2D) similarity, three-dimensional (3D) pharmacophore searches, and docking studies by using the structure of known agonists and our model of the ligand binding site, which was validated by mutagenesis. VS of a database of 2.6 million compounds followed by extraction of structural neighbors of functionally confirmed hits resulted in identification of 15 compounds active at FFAR1 either as full agonists, partial agonists, or pure antagonists. Site-directed mutagenesis and docking studies revealed different patterns of ligand–receptor interactions and provided important information on the role of specific amino acids in binding and activation of FFAR1.

Introduction

The free fatty acid receptor 1 (FFAR1[a]), previously known as GPR40, is a G-protein-coupled receptor (GPCR) that has been identified as a possible novel target for the treatment of type 2 diabetes. This receptor is highly expressed in the beta cells of pancreatic islets, and its activation by long-chain free fatty acids (FFAs) enhances glucose-stimulated insulin secretion.[1] Thus, this receptor is thought to play a role in the regulation of metabolic processes and glucose homeostasis.[2] It is assumed that synthetic agonists of FFAR1 may mimic the effect of FFAs to enhance glucose-stimulated insulin secretion with the potential to be developed into antidiabetic drugs.[2]

Steneberg et al. have shown that FFAR1 mediates chronic and acute effects of FFAs on beta cells in mice.[3] They found that FFAR1-deficient beta cells secrete less insulin in response to FFAs, indicating the importance of FFAR1 in mediating insulin release. However, enhanced expression of FFAR1 in the long term leads to hypoinsulinemia and to overt diabetes. In contrast, FFAR1-deficient mice were protected from obesity-induced hyperinsulinemia, hepatic steatosis, hypertriglyceridemia, increased hepatic glucose output, hyperglycemia, and glucose intolerance, which are all characteristic of the early stages of type 2 diabetes. Hence, there is no clear understanding to date whether agonists or antagonists of FFAR1 could be applied to the treatment of type 2 diabetes.

To learn more about the pharmacology of FFAR1 and about the implications of receptor activation and inhibition, the development of novel synthetic agonists and antagonists would be helpful. Full agonists based on the 3-(4-([*N*-alkyl]amino)phenyl) propanoic acid scaffold have been discovered recently by high-throughput screening (HTS).[4] The structure–activity relationships of compounds in this series have been explored, leading to the synthesis of agonists with nanomolar potencies, such as **1** (GW9508) and **2**.[4] Subsequently, the first selective antagonist, ethyl 4-[5-{[2-(ethyloxy)-5-pyrimidinyl]methyl}-2-{[(4-fluorophenyl)methyl]thio}-4-oxo-1(4*H*)-pyrimidinyl]benzoate (GW1100), was identified using the same techniques and was shown to completely inhibit the enhancement of glucose-stimulated insulin secretion mediated by **1**,[4] but only partially inhibit the enhancement mediated by linoleic acid.[5] However, this compound has been reported to act as a noncompetitive antagonist,[5] and therefore, it is likely not to interact at the orthosteric ligand binding site, thus preventing us from considering it in our virtual screening (VS). More recently, several bromophenyl derivatives were identified as FFAR1 agonists by HTS, and their chemical optimization led to the discovery of agonists with submicromolar potency.[6]

VS is a complementary approach to HTS that allows discovery of novel ligands from large libraries of diverse compounds by using information about the structure of the protein binding cavity or known ligands. This technique has been successfully employed for the search for novel ligands for several GPCRs.[7–13]

Recently, we published the first structural model of the binding site of FFAR1 in complex with **1**, which was obtained through an iterative approach that combined molecular modeling and receptor mutagenesis.[14] We showed that R183 (5.39), N244 (6.55), and R258 (7.35) are directly involved in interactions with **1** and linoleate,[14,15] and we proposed an NH$-\pi$ interaction between H137 (4.56) and **1** as one of the contributing forces leading to the high potency of **1**. Subsequently, we showed that H86 (3.32) also is able to interact with **1** in a pH-dependent manner while L186 (5.42) plays an important role in the interaction with **1,** but not with linoleic acid.[15]

Using our structural data for FFAR1,[14,15] we have performed VS by means of a two-dimensional (2D) similarity search followed by a three-dimensional (3D) pharmacophore search and docking studies to discover novel compounds that activate or inhibit the receptor. A set of 2 600 000 compounds from the

* Corresponding authors. (S.C.) E-mail: stefanoc@mail.nih.gov. Phone: 301-451-7353. Fax: 301-443-8000. (M.C.G.) E-mail: marving@intra.niddk.nih.gov. Phone: 301-496-4128. Fax: 301-496-9943.

† Laboratory of Biological Modeling.

‡ Clinical Endocrinology Branch.

§ These authors have contributed equally to this work.

[a] Abbreviations: FFAR1, free fatty acid receptor 1; GPCR, G-protein-coupled receptor; VS, virtual screening; HTS, high-throughput screening; MOE, molecular operating environment; XP, extraprecision; MCMM, Monte Carlo multiple minimum.

10.1021/jm7012425
Published on Web 01/15/2008

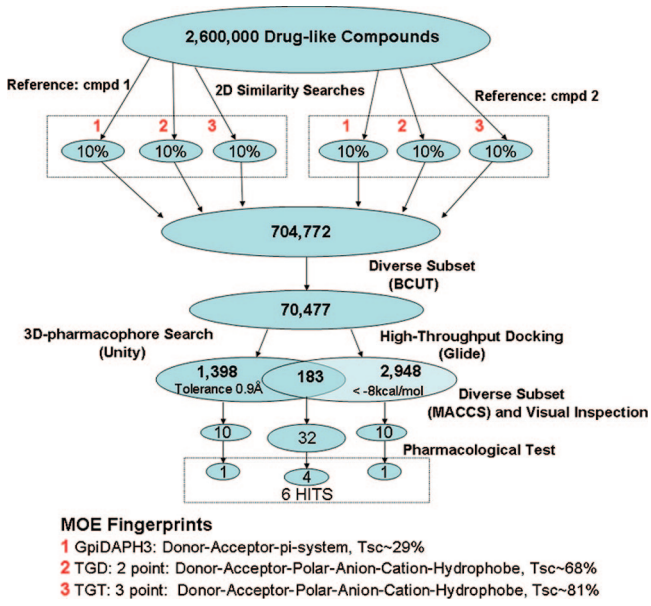


Figure 1. Scheme of the multistep VS protocol for identification of FFAR1 ligands. Initially, the database of commercially available druglike compounds, ZINC,[16] was subjected to 2D fingerprint similarity searches based on the known agonists **1** and **2**. The threshold of the Tanimoto coefficient (Tsc) for each search was chosen to select 10% of the input database. In the next step, a diverse subset of the resulting database of compounds was subjected to 3D pharmacophore search and high-throughput docking in parallel. The database of 1581 compounds obtained from the 3D pharmacophore search and the database of 3131 compounds obtained from docking had 183 identical compounds. In the final step, the selection of compounds for the pharmacological test was done based on their diversity calculated using BIT_MACCS structural key fingerprints and visual inspection. 32 compounds were selected from the overlapping library, and 10 unique compounds were selected from each of the two final databases. In total, 52 compounds were tested experimentally, and 6 hits were found: four hits came from the consensus library of 183 compounds, one hit from the 3D pharmacophore search alone, and one hit from docking alone.

ZINC[16] database of commercially available druglike molecules served as the screening library. A total of 70 compounds identified by VS and a subsequent neighbor search were tested for the ability to modulate the activity of FFAR1, leading to identification of 15 compounds acting as either agonists or antagonists.

Results and Discussion

The multistep VS performed in this study is schematically represented as a flowchart in Figure 1. In summary, using 2D structural fingerprints we initially analyzed a virtual library for similarity to the two known high-potency FFAR1 agonists **1** and **2** (Figure 2). Subsequently, a diverse subset of the compounds selected in this similarity search was subjected to either 3D pharmacophore search or high-throughput flexible docking. A diverse subgroup of the compounds that passed the various VS steps was selected for functional testing.

2D Similarity Search Analysis. In this step, the initial virtual subset of ~2.6 million commercially available druglike compounds retrieved from the database ZINC[16] was reduced by eliminating compounds distant from the two highly potent ligands of FFAR1 **1** and **2** (Figure 2) on the basis of their 2D structural fingerprints.

To enhance the diversity of the resulting database, we computed the structural similarities using three different fingerprints. Specifically, we generated one 2-point (TGD) and two 3-point (TGT and GpiDAPH3) pharmacophore-based fingerprints using the Molecular Operating Environment (MOE) software.[17] Subsequently, we calculated the Tanimoto similarity coefficient (Tsc) with compounds **1** and **2**. To ensure that each individual search yielded 10% of the input database, we adjusted the Tsc to 68% for the TGD, 81% for the TGT, and 29% for the GpiDAPH3 fingerprint.

We analyzed the percentage of identity between the filtered databases of compounds obtained with the six parallel searches (three for compound **1** and three for compound **2**). We found that the TGT-based databases had ~10% of identity to the TGD-based databases and ~15% of identity to the GpiDAPH3-based databases, while the GpiDAPH3-based databases had ~15% of identity to the TGD-based databases. The overall identity between databases screened by using **1** or **2** was 24%. The six filtered databases, each composed of ~260 000 compounds, were merged to yield a library of 704 772 unique compounds. To reduce the time required for the 3D searches, we decreased the size of the resulting database to 70 477 compounds (10%) by generating a diverse subset with the Diverse Solution software as implemented in Sybyl 7.1.[18]

3D Pharmacophore Search. In this step, we subjected the 70 477 compounds obtained in the 2D similarity analysis to a 3D pharmacophore search using the Unity module of Sybyl 7.1.[18] We have recently built a 3D model of the FFAR1 binding pocket complexed with **1**, which was experimentally supported by identification of the residues important for interaction with **1**.[14] After extraction of **1** from this model, we docked **2** to the obtained binding cavity by means of Glide. Only one orientation of **2**, similar to that obtained for **1**, was found and used for generation of a 3D pharmacophore. Compounds **1** and **2** adopted a similar V-shape conformation in the binding site (Figure 3) and established similar contacts with the amino acid residues experimentally identified as important hot spots for agonist recognition and activation.[14,15] A detailed picture of the conformation of **1** and of its interaction with the receptor binding site is provided elsewhere.[14] We defined the 3D pharmacophore on the basis of these docked conformations using partial match constraints and grouping the pharmacophoric features into three regional clusters. In the pharmacophore search, the screened compounds were required to match at least one feature from each cluster to be considered hits (Figure 3). The first cluster included two hydrogen bond acceptor features, which in our model form interactions with R183 (5.39), N244 (6.55), and R258 (7.35). The second cluster, which geometrically links the other two, included two hydrophobic features. Although **1** and **2** feature a secondary amine, our docking experiments did not yield any pose showing a hydrogen bond between the amino group and the protein. For this reason, we did not include a potential hydrogen bond donor in this area in our 3D pharmacophore query. The third cluster included one hydrogen bond acceptor and two hydrophobic features, which in our model form interactions with Y91 (3.37) and H137 (4.56).

In general, the number of compounds that match a pharmacophore query depends on the size of the pharmacophoric features (spheres). Here, we empirically set the radius (tolerance) of the spheres to 0.9 Å to obtain the hit rate of ~2% (1581 compounds).

High-Throughput Docking. In parallel to the 3D search, we subjected the same database of 70 477 compounds, obtained from the 2D similarity search, to high-throughput docking at the model of the ligand binding pocket of FFAR1 using Glide with the GlideScore scoring function.[19,20] We selected 3131 compounds with GlideScores lower than −8 kcal/mol. This

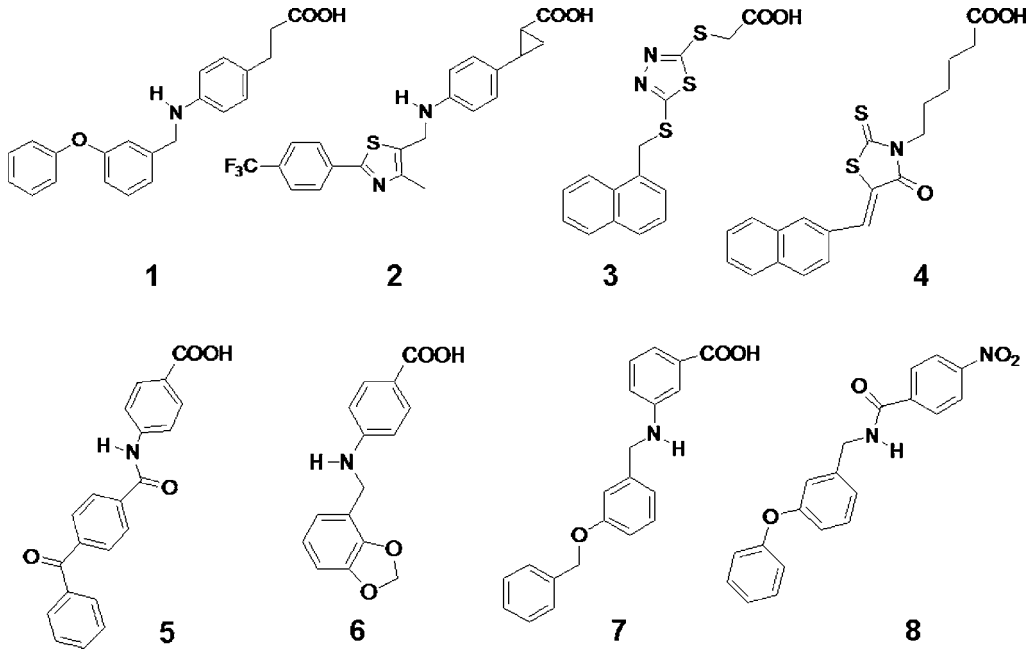


Figure 2. Structures of the potent synthetic agonists of FFAR1 **1** and **2**, which were used for 2D similarity and 3D pharmacophore searches, and of the novel ligands identified by virtual screening. Among the six hits, **3**, **4**, **5**, and **7** were retrieved by both the 3D pharmacophore search and docking studies, while **8** was found only by the pharmacophore search, and **6** was identified only by docking.

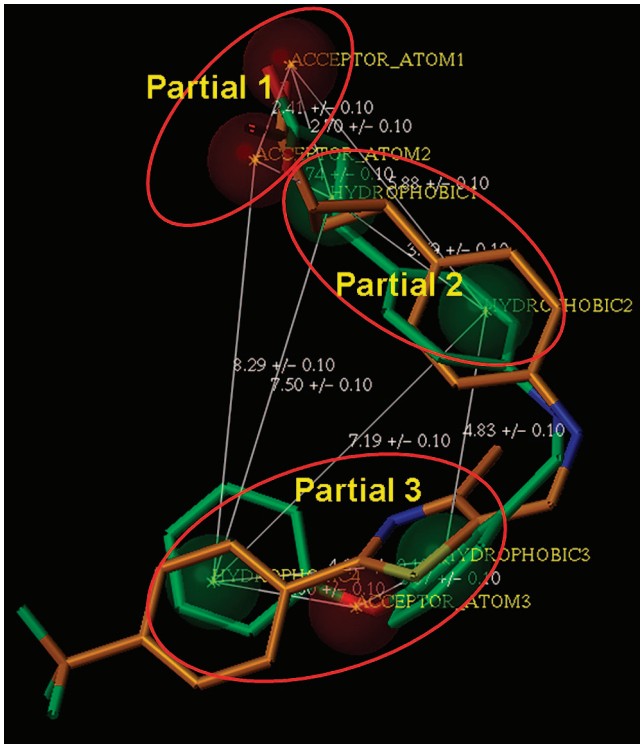


Figure 3. 3D pharmacophore query generated based on the common structural features of **1** and **2** (shown in green and orange correspondently). By using partial match constraints, the pharmacophoric features were grouped into three regional clusters (Partials 1, 2, and 3). During the pharmacophore search, the hits were required to match at least one feature from each cluster.

threshold was chosen in consideration of the fact that the GlideScores of **1** and **2** are −10 and −9 kcal/mol, respectively.

Choice of Compounds for the Experimental Test. The database of 1581 compounds obtained from the 3D pharmacophore search and the database of 3131 compounds obtained from docking contained 183 identical compounds. Compounds to be tested experimentally were selected based on their structural diversity calculated using BIT_MACCS structural keys (MOE) and visual inspection. Preference was given to the subset of the compounds (183) obtained by both the 3D pharmacophore search and docking through selection of 32 of these compounds and 10 unique compounds from each of the two final databases, for a total of 52 compounds (Figure 1 and Supporting Information).

Experimental Validation

The 52 compounds selected from VS were tested for both agonistic and antagonistic activities in a functional assay measuring receptor-mediated calcium flux. Compounds were screened by using both FFAR1-expressing cells and parental cells to discriminate between receptor-specific and nonspecific activity. Experimental testing identified six active compounds (Figure 2). Five compounds displayed agonistic activity with EC_{50} around 10 μM or less in FFAR1-expressing cells but showed no activation of parental cells. Based on the response observed with **1**, among the five compounds, two were full agonists (**3**, **4**), and three were partial agonists (**5**, **6**, and **7**). The concentration–response curves and the EC_{50} for the two most potent agonists (**3** and **4**) are shown in Figure 4a and Table 1. Further assay of the compounds by using cells expressing thyrotropin-releasing hormone receptor 1 (TRH-R1), a GPCR which couples to Gq, as does FFAR1, did not result in any response. The lack of response in non-FFAR1-expressing cells indicated that the agonists act on FFAR1 rather than on downstream signaling pathways.

To examine antagonistic activity, FFAR1-expressing cells were preincubated for 30 min with a test compound followed by addition of a near EC_{50} concentration of **1** (100 nM). In this assay, each of the identified full agonists and partial agonists diminished the stimulated response by **1**, probably due to receptor desensitization or competition for binding. In addition, **8** inhibited the stimulated response by **1** in the absence of detectable agonistic activity. The concentration–response curve for the inhibitory effect of **8** against **1** is shown in Figure 4b,

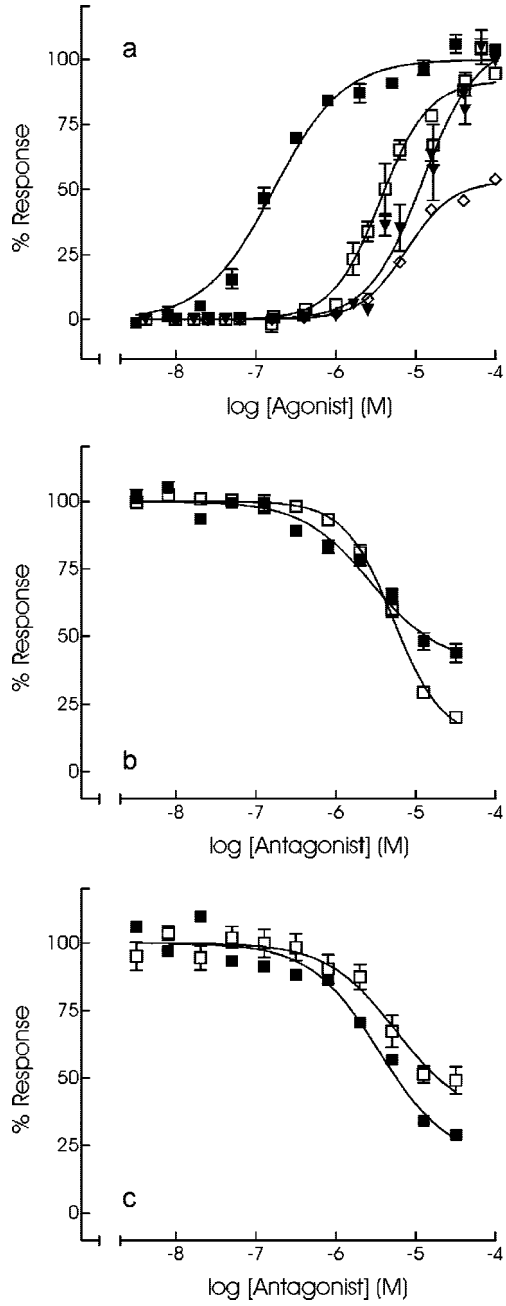


Figure 4. Effect of new compounds on the activity of FFAR1. (a) Receptor-mediated flux of intracellular calcium upon incubation with increasing concentrations of agonists **3** (□), **4** (▼), **16** (◇), and **1** (■). (b, c) Inhibitory effect of antagonists **8** (■) and **21** (□) on the agonistic activity of **1** (b) and linoleic acid (c). Compounds were incubated for 30 min prior to stimulation with EC_{50} concentrations of either **1** (100 nM) or linoleic acid (10 μM).

and the IC_{50} is presented in Table 1. Compound **8** also inhibited the linoleic acid-stimulated response with a similar IC_{50} (Figure 4c). Additionally, when we assayed the antagonist using TRH-R1-expressing cells stimulated with TRH, the antagonist **8** did not inhibit the TRH-stimulated activity except at 30 μM, where inhibition was about 30%. This suggests that the antagonist did not act as a downstream inhibitor of the signaling pathway; the effect at 30 μM could be due to partial nonselectivity or toxicity at the high concentration.

Analysis of the VS Scheme in Light of the Pharmacological Results. Following identification of the active compounds, we evaluated the effectiveness of the pharmacophore-based fingerprint searches. TGD, with the chosen Tsc of 68%, was most effective in hit isolation, selecting all identified actives based on similarity to either **1** or **2**. TGT, with the chosen Tsc of 81%, identified all actives with the exception of **4** based on similarity to **1** and with the exceptions of **4**, **6**, and **8** based on similarity to **2**. GpiDAPH3, with the chosen Tsc of 29%, did not select **4** based on similarity to **1**, nor did it select **5**, **7**, and **8** based on similarity to **2**. Notably, the Tsc between **1** and **2** is 84% for TGD, 87% for TGT, and 45% for GpiDAPH3.

Among the six hits, compounds **3**, **4**, **5**, and **7** were retrieved by both the 3D pharmacophore search and docking studies, while the pure antagonist **8** was found only by the pharmacophore search, and the partial agonist **6** was identified only by docking. Hence, the hit rate for the library of compounds selected by both the 3D pharmacophore search and docking was 12.5% (4/32) while the hit rate for the library of compounds selected only by one approach was 10% (1/10). This suggests that the hit rate could be slightly improved when the compounds are selected based on consensus between the 3D pharmacophore search and docking.

Our FFAR1 3D model was built based on the structure of the ground-state of rhodopsin and subsequently optimized with the agonist **1**. The docking part of our VS procedure yielded full agonists and partial agonists, but not pure antagonists, suggesting that in our model the ligand binding site is biased toward the active state of the receptor.

Neighbor Search. Using BIT_MACCS structural key fingerprint, we scanned the ZINC library[16] for close neighbors of the six experimentally confirmed hits. Eighteen compounds (Table 1, compounds **10**−**28**) were selected for experimental testing. The activity of these compounds was measured by intracellular calcium release experiments. Of the 18 compounds tested, nine were active with EC_{50} of 40 μM or lower (Table 1). Eight of these were partial agonists, and one was a pure antagonist. The neighbor search for full agonists **3** and **4** and for partial agonists **5** and **7** did not yield additional full agonists but only partial agonists, the most potent of which was **16** (the concentration–response curve is shown in Figure 4a). The neighbor search for the pure antagonist **8** led to the identification of **20**, another pure antagonist that was shown to inhibit the responses stimulated by **1** and linoleic acid with IC_{50} values similar to those of **8** (Figure 4b and c).

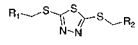
Structure–Activity Relationships (SARs) of Novel FFAR1 Ligands. Compounds **3** and **4**, which contain an aliphatic carboxylate, are the only full agonists identified by our VS. According to our modeling hypothesis, the carboxylate is coordinated by the positively charged residues R183 (5.39) and R258 (7.35) and by the hydrophilic residues S247 (6.58) and N244 (6.55) (Figure 1S of the Supporting Information). Compounds **5**, **6**, and **7**, which bear an aromatic carboxylate, are partial agonists. Regiochemistry seems to be an important factor, since movement of the carboxylate of **5** from the *para* to the *ortho* position led to complete inactivity of **24**. The pure antagonists **8** and **20** are both characterized by the presence of a nitro group instead of the carboxylate, suggesting that this substitution might contribute to their antagonistic properties. This hypothesis is currently under investigation. Docking of antagonist **8** is discussed as a separate section and is presented in Figure 1S of the Supporting Information. For the antagonists, regiochemistry is also important because moving the nitro group from the *para* to the *meta* position resulted in the inactivity of **23**. Our docking studies suggest that the derivatives with functional groups at the *ortho* or *meta* position cannot adopt the orientation in the binding site required to establish contacts

Table 1. Structures and Activities of the Six Virtual Screening Hits and the Structurally Related Compounds Retrieved with the Neighbor Search[c]

a

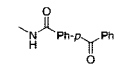
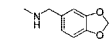
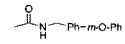
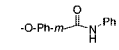
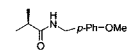
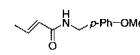
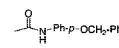
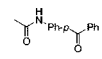
cmpd	R_1	R_2	Relative Maximum Response	EC_{50}	IC_{50}	Type
4	*n*-C_5H_{10}-COOH	1-naphthyl	100%	12.2		full agonist
9	*n*-C_2H_4-COOH	4-pyridyl		inactive		
10	*n*-C_3H_6-COOH	2-furyl	44%	16.8		partial agonist
11	*n*-C_3H_6-COOH	5-benzo[d][1,3]dioxole		ns[b]		(partial) agonist[a]
12	*n*-C_3H_6-COOH	*p*-fluorophenyl	26%	22.6		partial agonist[a]
13	*n*-C_3H_6-COOH	2-thiophenyl	43%	19		partial agonist
14	*n*-C_2H_4-COOH	1-naphthyl		inactive		
15	*n*-C_2H_4-COOH	3-indolyl		inactive		
16	*n*-C_5H_{10}-COOH	2-thiophenyl	59%	7.6		partial agonist

b



cmpd	R_1	R_2	Relative Maximum Response	EC_{50}	IC_{50}	Type
3	COOH	1-naphthyl	100%	3.6		full agonist
17	*p*-C_6H_4-COOH	*p*-metoxyphenyl	40%	15.1		partial agonist
18	COOMe	phenyl	41%	16.7		partial agonist
19	COOMe	*p*-metoxyphenyl		ns[b]		[d]

c

cmpd	R_1	R_2	R_3	R_4	Relative Maximum Response	EC_{50}	IC_{50}	Type
5	COOH	H	H		56%	16.1		partial agonist
6	COOH	H	H		50%	24.6		partial agonist
8	NO_2	H	H				2.8	pure antagonist
20	NO_2	H	H				5.3	pure antagonist
21	NO_2	H	H			inactive		
22	NO_2	H	H			inactive		
23	H	NO_2	H			inactive		
24	H	H	COOH			inactive		

d

cmpd	R_1	R_2	Relative Maximum Response	EC_{50}	IC_{50}	Type
7	-O-CH_2-*p*-Ph	H	52%	8.2		partial agonist
25	H	-O-CH_2-CH=CH_2	49%	14.5		partial agonist
26	-O-CH_2-CH=CH_2	H		ns[b]		[d]

[a] Compound with partial activity on non-FFAR1-expressing cells, and the maximal response is given as that after subtracting from non-FFAR1expressing cells. [b] Not saturating even at 30 μM. [c] EC_{50} and IC_{50} values (μM) are shown for agonists and antagonists, respectively. IC_{50} values were determined using 100 nM of **1**. [d] Compounds showing only very slight agonistic activity.

with the residues predicted to be involved in ligand recognition (Figure 2S of the Supporting Information).

Among the 4-thiazolidinone derivatives tested as neighbors of **4**, compounds bearing a long carboxylic aliphatic chain (five carbon atoms) demonstrated higher potency as compared to those having short chains. Notably, **4** and **16**, which are identical except for the substituent at the R2 position, exhibited similar potencies but different efficacies, while compounds **9**, **14**, and **15** were inactive (Table 1). According to our 3D model, the 4-thiazolidinone derivatives featuring long chain carboxylates adopt a semifolded conformation and are capable of interacting with H86 (3.32) or H137 (4.56) (Figure 5c), while the derivatives featuring short chain carboxylates cannot establish interactions with either of the two His residues. As an example, the docking pose of **12** is presented in Figure 1S of the Supporting Information. The binding mode of **16** is discussed in detail in the next section.

The SARs of the 1,3,4-thiadiazole derivatives of **3** suggest that the potency and efficacy are higher for the free carboxylates than for the methyl esters **18** and **19**. According to our docking results, the lower potencies of the ester derivatives are due not only to the loss of the negative charge but also to a steric hindrance exerted by the methyl group on R183 (5.39) and on R258 (7.35). The *para*-methoxyphenyl substituent at the R2 position resulted in the significant decreased potency of **19** with respect to **18**, which bears an unsubstituted phenyl group at this position. Docking of **18** and **19** revealed unfavorable interactions between the methoxy group and N241(6.52) (Figure 2S of the Supporting Information).

The substitution of the *meta*-benzyloxy group of **7** with a *para*-propenoxy group was well tolerated (**25**). However, the presence of a *meta*-propenoxy group resulted in an inactive compound (**26**). According to our model, the greatly diminished activity of **26** is probably due to steric clashes with L90 (3.36) and with L190 (5.46) (Figure 3S of the Supporting Information).

The analysis of the inactive compounds (structures in the Supporting Information) is also useful to study the determinants of ligand recognition. In particular, we noted that the carboxylate could not be substituted by imidazolidine-2,4-dione, tetrazole, piperazine-2,3-dione, pyrimidine-2,4-dione, 1,2,4 triazole, or

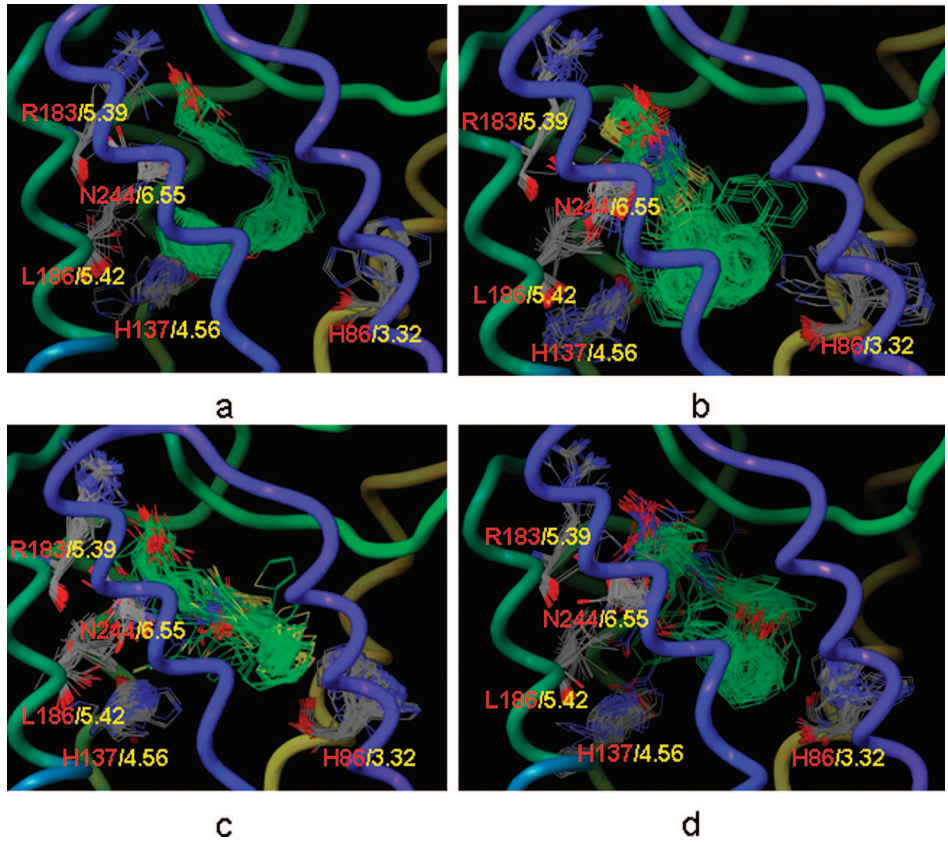


Figure 5. Conformational analysis of the FFAR1 binding pocket complexed with the full agonist **1** (a), the newly identified full agonist **3** (b), the partial agonist **16** (c), and the pure antagonist **8** (d). Only the ligands (colored in green) and residues R183, N244, L186, H137, and H86 are shown.

dihydrofuran-2-one. However, because these compounds have additional structural differences, this conclusion should be tested further.

Structure–Function Study of FFAR1 Using Novel Agonists. The search for new ligands of FFAR1 was based on the structure of the known ligands and on the model of the receptor complexed with **1**. Here, we used the newly discovered agonists to study structure–activity relationships in FFAR1, hypothesizing that the variety of their structural and functional properties might provide additional insights on the process of molecular recognition in FFAR1. Hence, we tested full agonist **3** and partial agonist **16** for interactions with residues known to be important for the activity of **1** (R183A, N244A, L186F, H137F, and H86F).[14,15]

R183 and N244 have been shown to be important as anchor residues for the carboxylate of **1**, and mutation of these residues to alanine greatly reduced receptor activation by this ligand.[14,15] Compounds **3** and **16** also were less effective at these mutants. In particular, the potency of **3** was reduced by about 100-fold, while that of **16** was reduced 10-fold or more in the R183A and N244A mutants (Figure 6a and c). H137 and, to a lesser extent, H86 also were shown to interact with **1**, presumably through aromatic contacts. In our model, **1** was proposed to make amino−aromatic interactions with H137.[14] Experimentally, **1** activated H137F with about 10-fold lower potency and reduced efficacy, whereas only potency was affected in H86F.[14,15] Compound **3** appeared to interact both with H86 (3.32) and with H137 (4.56) because H86F showed minimal response while H137F showed an efficacy diminished by about 50% without a significant change in potency (Figure 6b). Compound **16** had only slightly reduced potency and efficacy at H137F, but the potency was reduced to a greater extent at H86F (7-fold) (Figure 6d). Both **3** and **16** therefore appeared to make stronger interactions with H86 (3.32) than with H137 (4.56), whereas **1** interacted more strongly with H137 (4.56). Interestingly, L186F, which significantly reduced both the potency and efficacy of **1**, decreased the efficacy by about 50% with **3** but exhibited little change with **16** (Figure 6b and d).

In order to study the potential binding orientation of **3** and **16**, we docked these compounds to the model of the FFAR1 binding pocket and subsequently optimized the complexes by a Monte Carlo conformational search using MacroModel.[21,22] The results were compared with our previous conformational analysis of the complex of FFAR1 with **1**[15] (Figure 5). A total of 62, 67, and 70 conformations were obtained with **1**, **3**, and **16**, respectively. The results of the conformational analysis support the mutagenesis data. Compounds **1**, **3**, and **16** interact similarly with R183 (5.39) and N244 (6.55). However, there is a clear difference in the interactions with the two His residues located in the binding pocket. All conformers of **1** are consistently coordinated by H137 (4.56). The full agonist **3** can interact with either H137 (4.56) or H86 (3.32), with the position of the naphthyl group showing almost equal distribution between these two residues. The partial agonist **16** is mostly oriented toward H86 (3.32). These data suggest that although the agonists bind within the same binding pocket, different ligands interact differently with the residues forming the pocket. In particular, the effect of H86F on the activity of **3** revealed an important role for H86 (3.32) that was not clearly appreciated in the case of **1**.

Binding Mode of the Novel Antagonist in FFAR1. In order to explore the binding mode of the pure antagonist **8**, we performed a docking experiment followed by a Monte Carlo conformational search of the ligand within the binding pocket

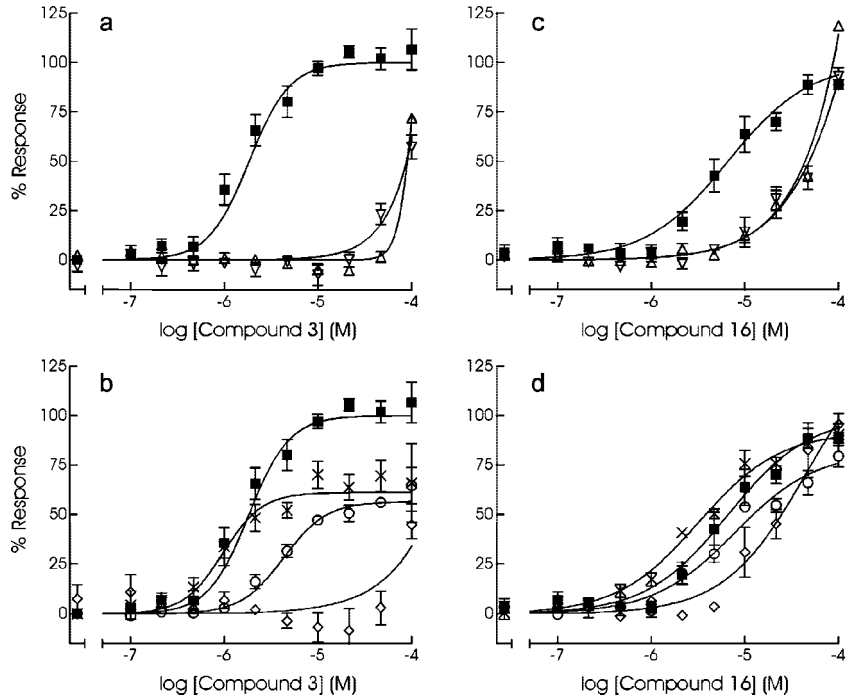


Figure 6. Structure–function study of FFAR1. HEK-EM 293 cells transiently expressing wild-type FFAR1 (■), R183A (△), N244A (▽), H86F (◇), H137F (○), or L186F (×) were treated with **3** (a and b) or **16** (c and d), and the intracellular calcium flux was measured in the FLIPR assay. With compound **3**, the EC_{50} values were 1.8 μM (wild-type), ≥100 μM (R183A), ≥100 μM (N244A), ≥100 μM (H86F), 4.3 μM (H137F), and 0.97 μM (L186F). With compound **16**, the EC_{50} values were 6.6 μM (wild-type), >30 μM (R183A), >30 μM (N244A), >10 μM (H86F), 7.8 μM (H137F), and 3.2 μM (L186F).

of FFAR1. The docking pose of **8** generated by Glide resulted in a GlideScore of −7.7 kcal/mol, which was higher than the threshold that we adopted in our VS, explaining why compound **8** was retrieved by the 3D pharmacophore search but not by docking. The subsequent conformational search generated 68 complexes.

Although the binding mode of **8** remains to be confirmed experimentally, molecular modeling suggests an overall orientation similar to that of the agonists. Generally, the nitro group of compound **8** docked similarly to the carboxylate group of the agonists (Figure 5d). However, in 16 complexes the nitro group flipped away from the positively charged residues, suggesting that its interactions with R183 (5.39) and R258 (7.35) were weaker than those established by the negatively charged carboxylate of the agonists. Moreover, the aromatic moiety of **8** did not establish contacts with H137 (4.56) and Y91 (3.37), which were shown to be important for receptor activation in our previous studies. With a manner similar to that of the partial agonist **16**, the pure antagonist **8** seemed to lean toward H86 (3.32).

Conclusions. Computer-assisted drug discovery (CADD) is an attractive alternative to high-throughput screening for the identification of novel and diverse ligands for biological targets of pharmaceutical interest. In particular, virtual screenings allow a quick evaluation of large databases of compounds. Only a limited number of compounds are purchased and experimentally evaluated, with a conspicuous saving of time and financial resources.

Here, we applied VS to the identification of ligands for the GPCR FFAR1, a potential therapeutic target for the regulation of insulin secretion. We devised a multistep VS protocol which included 2D similarity searches, 3D pharmacophore searches, and molecular docking. Parallel fingerprint similarity searches based on the structure of two potent FFAR1[4] agonists served as the first step of our VS protocol. Subsequently, we performed a 3D pharmacophore search in parallel with high-throughput docking. Our experimental results suggest that compounds consensually retrieved by the 3D pharmacophore search and by docking have more chances to be hits than those retrieved by only one of the two techniques.

Our molecular docking was based on a 3D model of FFAR1 built on the ground-state structure of rhodopsin and optimized with the full agonist **1**. It yielded only full and partial agonists, suggesting that the model is biased toward agonist recognition. However, site-directed mutagenesis and docking studies suggest that subtle changes in the ligand–receptor interactions are sufficient to affect the spectrum of efficacy of the ligands.

Given that FFAR1 is a promising target for the treatment of type 2 diabetes, the novel ligands and SAR provided here offer ground for the generation of potential antidiabetic agents. Additionally, the discovery of agonists and antagonists may provide the pharmacological tools necessary to understand the role of FFAR1 activation and/or inhibition in the regulation of insulin and glucagon secretion.

Experimental Section

2D Similarity Search and Nearest Neighbor Search. The 2D similarity searches were performed with the software MOE[17] using two-point (TGD) and three-point (TGT and GpiDAPH3) pharmacophore-based fingerprints, all calculated from a 2D molecular graph. Each atom was given a type among donor, acceptor, polar, anion, cation, or hydrophobe for the calculation of TGD and TGT, and among donor, acceptor, or π system for the calculation of GpiDAPH3. Subsequently, pairs (two-point fingerprint) or triplets (three-point fingerprints) of types were formed by graph distances and coded as sparse features in a fingerprint. Distance or triangle-based fingerprints are powerful tools for the identification of similar compounds endowed with diverse scaffolds. Furthermore, being graph-based, TGD, TGT, and GpiDAPH3 do not require conformational databases. The Tanimoto coefficient was used as the similarity metric.

For the nearest neighbor searches, we used a bit-packed version of the MACCS fingerprints. In the MACCS fingerprints, each feature represents one of the 166 public MDL MACCS structural keys calculated from the molecular graph. Being based on chemical substructures, MACCS fingerprints are less likely to retrieve diverse compounds than distance or triangle-based fingerprints, and they are more suited for neighbor searches.

Diverse Subset. A diverse subset of the 704 772 compounds retrieved by the 2D similarity searches was selected using the Diverse Solutions application available with Sybyl 7.1.[18] The selection of this subset was carried out based on the position of the compounds in 6D chemical space, defined by BCUT descriptors. These descriptors consider various physicochemical parameters related to ligand–receptor binding, including atomic charges, atomic polarizability, and atomic hydrogen-bond donor and acceptor abilities, as well as topological properties, including 2D connectivity and exposed surface area scaling factors.[23] The remaining diverse subset selections were carried out with MOE using a bit-packed version of the MACCS fingerprints

3D Pharmacophore Search. A 3D pharmacophore query with partial match constraints was defined on the basis of the common structural features of the agonists **1** and **2**, by using the putative bound conformation of the ligands derived from the docking. The 3D pharmacophore search was performed by using the Unity flexible search protocol as implemented in Sybyl 7.1, with all options set as default.[18] In the Unity search, the conformations of the screening database were generated on the fly by means of the Directed Tweak method.[24] The maximum time per structure, which affects the thoroughness of the conformational search, was set to 60 s. Given that our database includes only druglike compounds, Lipinski's rule of 5 was turned off during the search.

High-Throughput Docking and Scoring. Automated docking was performed by using Glide (Schrödinger)[19,20] with extraprecision (XP) settings. The binding site was defined as the residues located within 6 Å from the ligand. To compensate for the rigid representation of the receptor, the Van der Waals radii of the atoms with partial charges lower than 0.25 were scaled by the factor 0.6 to generate the receptor grid. All compounds were preprocessed with LigPrep at pH 7. Default settings were used for docking. The final ligand poses were selected based on the Glide empirical docking score (GlideScore). The OPLS2005 force field was used for all calculations.

Conformational Search. The complexes of **3** and **16** with FFAR1 obtained by docking were subsequently subjected to conformational search using the Monte Carlo Multiple Minimum (MCMM) protocol as implemented in MacroModel.[21,22] Ligands and residues located within 6 Å of the ligand were subjected to an extended torsional sampling and were surrounded by a shell of frozen atoms of 3 Å; 1000 steps of MCMM were performed, and the resulting structures with a potential energy lower than 2000 kJ/mol were saved. The OPLS2005 force field was used with a dielectric constant of one and with the GB/SA model for treatment of solvation.

Cell Culture and cDNA Constructs. Wild-type HEK-EM 293 cells were cultured in DMEM (Mediatech, Herndon, VA) supplemented with 10% fetal bovine serum (Hyclone, Logan, UT). HEK-EM 293 cells stably expressing FFAR1[14] or TRH-R1[25] were grown in the same medium further supplemented with 200 μg/mL hygromycin. cDNA constructs encoding R183A, N244A, H86F, H137F, and L186F FFAR1 mutants have been described previously.[14,15] To characterize the effects of mutations on the agonist-stimulated response, cDNA constructs encoding the mutant receptors or the wild-type receptor were transiently transfected into HEK-EM 293 cells. Transfections were carried out using lipofectamine LTX and PLUS reagent (Invitrogen, Carlsbad, CA) according to the manufacturer's protocol.

Receptor Assay with a Fluorometric Imaging Plate Reader (FLIPR). Receptor signaling assays were carried out essentially as described previously by measuring the calcium flux in response to addition of compounds.[14] Cells were seeded at 60 000 cells/well in 96-well plates one day prior to experiments. The cells were loaded with the calcium fluorescent indicator Calcium4 (Molecular Devices, Sunnyvale, CA) by replacing the growth medium with the dye dissolved in HEPES (20 mM)-buffered Hank's balanced salt solution (HBSS). Compounds were screened for agonistic and antagonistic activity in tandem in a single experiment. First, agonistic activity was measured by monitoring the fluorescence signals for 5 min upon addition of compounds. Subsequently, the cells were incubated for 25 min with test compounds. Thereafter, the antagonistic activity was monitored for the inhibitory effect upon the addition of an agonist. Compound **1** and linoleic acid were used as the agonists for FFAR1, whereas thyrotropin-releasing hormone was used for TRH-R1. Data were analyzed using data sets using GraphPad Prism 4 (San Diego, CA) as described.[15]

Acknowledgment. This research was supported by the Intramural Research Program of the National Institute of Diabetes and Digestive and Kidney Diseases, NIH.

Supporting Information Available: Table enlisting the 52 compounds selected by virtual screening and experimentally tested. Docking poses of selected compounds at FFAR1. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Itoh, Y.; Kawamata, Y.; Harada, M.; Kobayashi, M.; Fujii, R.; Fukusumi, S.; Ogi, K.; Hosoya, M.; Tanaka, Y.; Uejima, H.; Tanaka, H.; Maruyama, M.; Satoh, R.; Okubo, S.; Kizawa, H.; Komatsu, H.; Matsumura, F.; Noguchi, Y.; Shinohara, T.; Hinuma, S.; Fujisawa, Y.; Fujino, M. Free fatty acids regulate insulin secretion from pancreatic beta cells through GPR40. *Nature* **2003**, *422*, 173–176.

(2) Rayasam, G. V.; Tulasi, V. K.; Davis, J. A.; Bansal, V. S. Fatty acid receptors as new therapeutic targets for diabetes. *Expert Opin. Ther. Targets* **2007**, *11*, 661–671.

(3) Steneberg, P.; Rubins, N.; Bartoov-Shifman, R.; Walker, M. D.; Edlund, H. The FFA receptor GPR40 links hyperinsulinemia, hepatic steatosis, and impaired glucose homeostasis in mouse. *Cell Metab.* **2005**, *1*, 245–258.

(4) Garrido, D. M.; Corbett, D. F.; Dwornik, K. A.; Goetz, A. S.; Littleton, T. R.; McKeown, S. C.; Mills, W. Y.; Smalley, T. L., Jr.; Briscoe, C. P.; Peat, A. J. Synthesis and activity of small molecule GPR40 agonists. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 1840–1845.

(5) Briscoe, C. P.; Peat, A. J.; McKeown, S. C.; Corbett, D. F.; Goetz, A. S.; Littleton, T. R.; McCoy, D. C.; Kenakin, T. P.; Andrews, J. L.; Ammala, C.; Fornwald, J. A.; Ignar, D. M.; Jenkinson, S. Pharmacological regulation of insulin secretion in MIN6 cells through the fatty acid receptor GPR40: identification of agonist and antagonist small molecules. *Br. J. Pharmacol.* **2006**, *148*, 619–628.

(6) Song, F.; Lu, S.; Gunnet, J.; Xu, J. Z.; Wines, P.; Proost, J.; Liang, Y.; Baumann, C.; Lenhard, J.; Murray, W. V.; Demarest, K. T.; Kuo, G. H. Synthesis and biological evaluation of 3-Aryl-3-(4-phenoxy)-propionic acid as a novel series of G protein-coupled receptor 40 agonists. *J. Med. Chem.* **2007**, *50*, 2807–2817.

(7) Becker, O. M.; Marantz, Y.; Shacham, S.; Inbal, B.; Heifetz, A.; Kalid, O.; Bar-Haim, S.; Warshaviak, D.; Fichman, M.; Noiman, S. G-protein-coupled receptors: in silico drug discovery in 3D. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 11304–11309.

(8) Becker, O. M.; Dhanoa, D. S.; Marantz, Y.; Chen, D.; Shacham, S.; Cheruku, S.; Heifetz, A.; Mohanty, P.; Fichman, M.; Sharadendu, A.; Nudelman, R.; Kauffman, M.; Noiman, S. An integrated in silico 3D model-driven discovery of a novel, potent, and selective amidosulfonamide 5-HT1A agonist (PRX-00023) for the treatment of anxiety and depression. *J. Med. Chem.* **2006**, *49*, 3116–3135.

(9) Evers, A.; Klebe, G. Successful virtual screening for a submicromolar antagonist of the neurokinin-1 receptor based on a ligand-supported homology model. *J. Med. Chem.* **2004**, *47*, 5381–5392.

(10) Evers, A.; Hessler, G.; Matter, H.; Klabunde, T. Virtual screening of biogenic amine-binding G-protein coupled receptors: comparative evaluation of protein- and ligand-based virtual screening protocols. *J. Med. Chem.* **2005**, *48*, 5448–5465.

(11) Vaidehi, N.; Schlyer, S.; Trabanino, R. J.; Floriano, W. B.; Abrol, R.; Sharma, S.; Kochanny, M.; Koovakat, S.; Dunning, L.; Liang, M.; Fox, J. M.; de Mendonca, F. L.; Pease, J. E.; Goddard, W. A., III; Horuk, R. Predictions of CCR1 chemokine receptor structure and BX 471 antagonist binding followed by experimental validation. *J. Biol. Chem.* **2006**, *281*, 27613–27620.

(12) Varady, J.; Wu, X.; Fang, X.; Min, J.; Hu, Z.; Levant, B.; Wang, S. Molecular modeling of the three-dimensional structure of dopamine

3 (D3) subtype receptor: discovery of novel and potent D3 ligands through a hybrid pharmacophore- and structure-based database searching approach. *J. Med. Chem.* **2003**, *46*, 4377–4392.
(13) Kellenberger, E.; Springael, J. Y.; Parmentier, M.; Hachet-Haas, M.; Galzi, J. L.; Rognan, D. Identification of nonpeptide CCR5 receptor agonists by structure-based virtual screening. *J. Med. Chem.* **2007**, *50*, 1294–1303.
(14) Tikhonova, I. G.; Sum, C. S.; Neumann, S.; Thomas, C. J.; Raaka, B. M.; Costanzi, S.; Gershengorn, M. C. Bidirectional, iterative approach to the structural delineation of the functional "chemoprint" in GPR40 for agonist recognition. *J. Med. Chem.* **2007**, *50*, 2981–2989.
(15) Sum, C. S.; Tikhonova, I. G.; Neumann, S.; Engels, M.; Raaka, B. M.; Costanzi, S.; Gershengorn, M. C. Identification of residues important for agonist recognition and activation in GPR40. *J. Biol. Chem.* **2007**, *282*, 29248–29255.
(16) Irwin, J. J.; Shoichet, B. K. ZINC−a free database of commercially available compounds for virtual screening. *J. Chem. Inf. Model.* **2005**, *45*, 177–182.
(17) Molecular Operating Environment (MOE). [2008.05]; Chemical Computing Group, Inc.: Montreal, Quebec, Canada, 2007.
(18) SYBYL Molecular Modeling System. [7.1]; TRIPOS, Assoc., Inc.: St-Louis, MO, 2006.
(19) Glide. [4.0]; Schrödinger, LLC.: New York, 2005.
(20) Friesner, R. A.; Murphy, R. B.; Repasky, M. P.; Frye, L. L.; Greenwood, J. R.; Halgren, T. A.; Sanschagrin, P. C.; Mainz, D. T. Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein−ligand complexes. *J. Med. Chem.* **2006**, *49*, 6177–6196.
(21) Macromodel. [9.1]; Schrödinger, LLC.: New York, 2007.
(22) Mohamadi, F.; Richards, N. G. J.; Guida, W. C.; Liskamp, R.; Lipton, M.; Caufield, C.; Chang, G.; Hendrickson, T.; Still, W. C. MacroModel−An integrated software system for modeling organic and bioorganic molecules using molecular mechanics. *J. Comput. Chem.* **1990**, *11*, 440–467.
(23) Mason, J. S.; Beno, B. R. Library design using BCUT chemistry-space descriptors and multiple four-point pharmacophore fingerprints: simultaneous optimization and structure-based diversity. *J. Mol. Graphics Modell.* **2000**, *18*, 438–451.
(24) Hurst, T. Flexible 3D searching - the directed tweak technique. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 190–196.
(25) Engel, S.; Neumann, S.; Kaur, N.; Monga, V.; Jain, R.; Northup, J.; Gershengorn, M. C. Low affinity analogs of thyrotropin-releasing hormone are super-agonists. *J. Biol. Chem.* **2006**, *281*, 13103–13109.

JM7012425