Identification of Novel Cannabinoid CB1 Receptor Antagonists by Using Virtual Screening with a Pharmacophore Model

Hongwu Wang,*[,†] Ruth A. Duffy,[‡] George C. Boykow,[‡] Samuel Chackalamannil,[§] and Vincent S. Madison[†]

Departments of Structural Chemistry, CNS Biological Research, and CNS/CV Chemistry, Schering-Plough Research Institute, 2015 Galloping Hill Road, Kenilworth, New Jersey 07033

Received December 6, 2007

CB1 receptor antagonists have proven to be clinically effective in treating obesity and related disorders. We report here the identification of a novel class of azetidinone CB1 antagonists by using virtual screening methods. For this purpose, we developed a pharmacophore model based on known representative CB1 antagonists and employed it to screen a database of about a half million Schering-Plough compounds. We applied a stepwise filtering protocol based on molecular weight, compound availability, and a modified rule-of-five to reduce the number of hits. We then combined Bayesian modeling and clustering techniques to select a final set of 420 compounds for in vitro testing. Five compounds were found to have >50% inhibition at 100 nM in a CB1 competitive binding assay and were further characterized by using both CB1 and CB2 assays. The most potent compound has a CB1 K_i of 53 nM and >5-fold selectivity against the CB2 receptor.

Introduction

Selective cannabinoid CB1 receptor antagonists and inverse agonists have been shown to be effective clinically for the treatment of obesity.[1] Obesity has become a major public health challenge and is estimated to be associated with over 100 000 excess deaths every year in the United States alone.[2] Recent surveys show that 28% of men and 34% of women in the United States are obese, and children are among the groups with the greatest obesity rate increase.[3,4] There is an urgent medical need for the effective and safe treatment of obesity to lower the risks of developing cardiovascular disorders, type-2 diabetes, and other serious diseases.[5] The cannabinoid system represents an attractive target for obesity therapy, and extensive research efforts in the past two decades have led to the discovery of rimonabant, an orally active CB1 inverse agonist.[6,7] Rimonabant was recently approved for the treatment of obesity in Europe and has shown clinical efficacy for smoking cessation as well. However, rimonabant was not approved in the United States, and the psychiatric safety concerns may limit its applications. Additionally, CB1 antagonists could be promising treatments for cognitive disorders such as memory impairment and several other diseases.[8]

The known cannabinoid system in humans consists of two cannabinoid receptors, CB1 and CB2, with increasing data suggesting the existence of a third type, CB3.[9–11] The cannabinoid receptors are G-protein coupled receptors (GPCRs).[a] The CB1 and CB2 receptors share 44% sequence identity, which rises to 68% in the transmembrane domains. The CB1 receptor is abundantly expressed in the central nervous system as well as in peripheral tissues, for example, testis, eye, ileum, and adipocytes. It is involved in the regulation of cognition, memory, and motor activity and inhibits transmitter release through its coupling to ion channels. By contrast, the CB2 receptor is almost exclusively expressed in cells of the immune system and is assumed to participate in the regulation of immune responses and inflammatory reactions.

The crystal structure of the CB1 receptor remains unsolved, but homology models based on the X-ray structure of rhodopsin[12] have provided insights into the ligand−receptor interactions for CB1 antagonists.[13–16] These models were usually developed and validated by using available mutagenesis data along with structure–activity relationship (SAR) information from known ligands. The ligands were first placed into a putative binding site in the transmembrane domain, and energy minimization or molecular dynamics simulations were then applied to relax the complex structures to produce the ligand−receptor models. As with other signal transduction GPCRs, the CB1 receptor can exist in two states, the activated state (R*) and inactive state (R). It has been suggested that antagonists bind to the two states with equal potency, whereas inverse agonists bind preferentially to the inactive state. CB1 receptor models for both states have been proposed, and the key ligand−receptor interactions have been described.[17,18] Despite their usefulness in explaining experimental observations and revealing binding interactions, these homology models have had limited success in structure-based drug design or virtual screening applications.[19,20]

In the absence of a reliable target structure, ligand-based molecular modeling tools can be used to derive the structural requirements crucial for receptor binding. Pharmacophore modeling belongs to this approach and has been widely used in drug discovery. Pharmacophore models are usually developed with a collection of structurally diverse compounds that are known to bind to the same active site. Two types of pharmacophore models have been reported in the literature. A 3D quantitative structure–activity relationship (QSAR)-like model can be developed by using a training set of compounds with biological activities spread over 3 orders of magnitude.[21,22] This approach is represented by the Hypogen method in the Catalyst

* Corresponding author. Tel: (908) 740-2924. Fax: (908) 740-4640. E-mail: hongwu.wang@spcorp.com.
† Department of Structural Chemistry.
‡ Department of CNS Biological Research.
§ Department of CNS/CV Chemistry.

[a] Abbreviations: AR, aromatic ring; GPCR, G-protein coupled receptor; HBA, hydrogen bond acceptor; HBD, hydrogen bond donor; HPO, hydrophobic; MCMM, Monte Carlo tortional sampling; MDDR, MDL Drug Data Report; NG, negatively charged/ionizable group; PG, positively charged/ionizable group; QSAR, quantitative structure–activity relationship; SAR, structure–activity relationship.

10.1021/jm701519h CCC: $40.75
Published on Web 03/26/2008

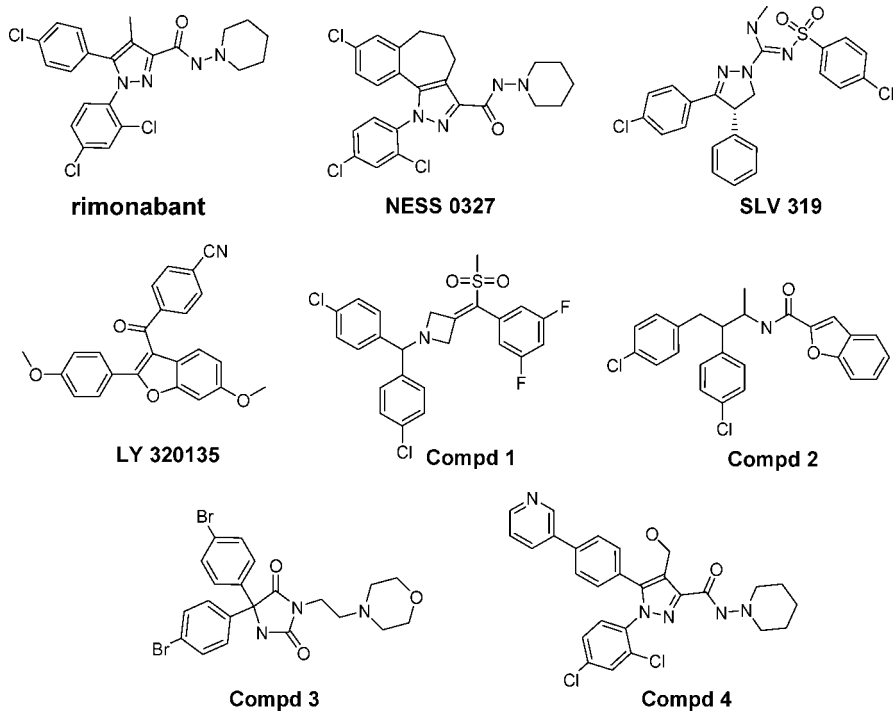


Figure 1. Structures of the compounds used for pharmacophore model development. The compounds were labeled by their generic name or company codes when available.

software.[23] The potencies of new compounds can be quantitatively predicted by evaluating how well each compound maps onto the pharmacophore model. Generally speaking, the predictions are heavily affected by the conformational models and the methods used for pharmacophore mapping, that is, Best or Fast. On the other hand, a pharmacophore model can be derived with a training set including only the active compounds. This approach is represented by the HipHop method in the Catalyst software. The biological activities of new compounds can be evaluated qualitatively by whether they match the pharmacophore model.

A pharmacophore model captures the 3D arrangement of the structural features common to the active molecules that are presumably essential for the desired pharmacological activity. Most of the pharmacophore modeling methods consider functional features such as hydrogen bond donors (HBD), hydrogen bond acceptors (HBA), hydrophobic groups (HPO), aromatic rings (AR), positively charged/ionizable groups (PG), and negatively charged/ionizable groups (NG). Additionally, shape and excluded volumes can be incorporated into the pharmacophore models to represent the framework of the active site. Pharmacophore models have many applications in drug design. They can be used to align structurally diverse leads; as a consequence, the groups in each structure that are crucial for the biological activity can be identified. Nonessential parts of the molecule can be modified to improve the physical and pharmacokinetic properties, or new molecular scaffolds can be designed to establish novel patent space. Equally importantly, pharmacophore models have been applied successfully in virtual screening to select molecules for biological testing from large databases.[22,24–29] Pharmacophore screening is much faster than protein structure-based virtual screening, therefore making it possible to work with very large databases. In addition, focusing on 3D pharmacophore features also enables retrieval of structurally diverse hits as compared to methods using 2D similarity indices.

In this paper, we describe the development of a CB1 antagonist pharmacophore model and its application as a virtual screening tool to identify a novel class of CB1 selective antagonists. We evaluated more than a half million compounds from our corporate database by using this pharmacophore model and found a large number of hits that had never been tested before in cannabinoid receptor bind assays. We applied a combination of filtering techniques to select a subset of 420 compounds among these hits; the subsequent in vitro assays revealed a new class of CB1 selective antagonists.

Methods

Training Set for Pharmacophore Model. Eight representative CB1 selective antagonists and inverse agonists were collected from the literature, most of them from drug discovery programs at different pharmaceutical companies.[30–37] When multiple structures were reported in the original literature, only one representative compound (usually the most potent one) from each structural class was selected. Structural redundancy is not necessary for the generation of common feature pharmacophore models. The structures of these compounds are shown in Figure 1.

Pharmacophore Model Generation. A 3D structure for each compound in the training set was generated from its 2D SMILES representation by using the Concord program.[38] The structures were then subjected to energy minimization until the gradient dropped below 0.05. The minimized structures were used as the starting points for subsequent conformational searches. A 5000-step Monte Carlo tortional sampling (MCMM) conformational search was conducted for each compound. Unique low-energy conformations within 40 kJ/mol of the corresponding global energy minimum were collected for each molecule. A conformation was considered unique only when the maximum displacement of at least one heavy atom was larger than 0.5 Å compared to all the low-energy conformations already collected for that molecule. A maximum of 250 unique conformations were collected for each compound. Energy minimization and conformational search were conducted with MacroModel[39] using the MMFF94s force field.

Pharmacophore models were generated with the Catalyst HipHop method. On the basis of the structural characteristics of the training

set molecules, only HBA, AR, and HPO features were considered for model development. Ten HipHop pharmacophore models were generated, and each model was visually evaluated by mapping the molecules in the training set. The model that yielded the most reasonable mapping for rimonabant was selected for the database search.

Database Search. A database search was conducted on a pregenerated Catalyst 3D database of about a half million compounds from the Schering-Plough collection. Standard procedures such as stripping salts from the structures, discarding compounds with inorganic atoms, and removing compounds with very high molecular weight (>900) or too many rotatable bonds (>15) were applied to select compounds for Catalyst database generation. Conformational models were generated by using the Catalyst FAST approach from its SMILES representation, and a maximum of 150 diverse conformations were retained for each compound in the database. The database search was conducted with Catalyst BEST flexible search method by using the above-mentioned pharmacophore model. A molecule must match all features of the pharmacophore model to be retrieved as a hit.

Hits Selection. At the beginning of the project, a target was set to assay 420 compounds. As expected, the number of hits obtained from pharmacophore screening was well above our target (~22 000 hits). To narrow down the number of hits, a stepwise postsearch filtering scheme was developed. We first filtered out compounds with molecular weight higher than 550 because they were not desirable for lead optimization. The remaining hits were then filtered by sample availability; any compound with less than 2 mg available at the time of selection was discarded. Third, we filtered out compounds with potential pharmacokinetic problems by using a modified Lipinski's rule-of-five.[40,41] Compounds that did not meet at least three of the following criteria were left out:

MW $\leq$ 500
Clog $P \leq 5$
Number of HBA $\leq$ 10
Number of HBD $\leq$ 3
Number of rotatable bonds $\leq$ 10

These three steps together eliminated about two-thirds of the original hits.

We next employed Bayesian modeling and clustering techniques offered in Pipeline Pilot software[42] to further reduce the number of hits. A Bayesian model was developed at this stage to rank the hits for their likelihood of being CB1 antagonists. Two data sets are required for Bayesian modeling, one containing active molecules and the other one containing inactive reference compounds. We turned to the MDL Drug Data Report (MDDR)[43] for these data sets. This database holds the structures and activities for drug candidates and was compiled from published literature such as patents and research journals. The ACTION field reports the biological activities for a molecule. Accordingly, we searched the MDDR database with the keyword CB1 in this field. All compounds listed as selective CB1 antagonist, CB1 inverse agonist, or CB1 modulator useful for the treatment of obesity were selected as actives for the Bayesian modeling. We then generated an inactive data set (negative control) from the rest of the compounds in the MDDR database by treating and filtering them with the following steps: remove salts from each structure, keep the largest piece for entries with multiple fragments, remove compounds with inorganic atoms, and filter out compounds with MW > 650 or MW < 300. Pipeline Pilot software was then used to develop a CB1 Bayesian model by using the FCFP6 structural descriptors. This Bayesian model was applied to rank the remaining hits.

Clustering was employed to trim down the structural redundancy among the hits. The Pipeline Pilot maximum dissimilarity clustering algorithm was applied to cluster the structures. This algorithm starts with a randomly chosen structure as the first cluster center. The molecule with the maximal distance from the first compound is selected as the next cluster center. The compound with the maximal distance from both current cluster centers is selected after that. The process is repeated until the desired number of cluster centers is reached. The nonselected objects are then assigned to the nearest cluster center to determine the cluster membership. The structural clustering was based on the pairwise Tanimoto distances calculated with the FCFP6 fingerprint descriptors.

Table 1. Cannabinoid Receptor Binding Affinities for Compounds in the Training Set

compound name	CB1 K_i (nM)	CB2 K_i (nM)	ref	number of conformers[a]
Rimonabant	1.98	>1000	30	62
NESS 0327	0.00035	21	31	130
SLV 319	7.8	7943	32	26
LY 320135	141	>10 000	33	35
Compd 1	<100	not reported	34	25
Compd 2	not reported	not reported	35	32
Compd 3	70.3	not reported	36	82
Compd 4	5.8	2312	37	250

[a] Number of conformers collected from the MacroModel conformational search. A maximum number of 250 conformers with conformational energy <40 kJ/mol above the corresponding global energy minimum for each compound was used for pharmacophore model generation.

Cannabinoid Receptor Binding Assays. cDNA expressing the human CB1 or CB2 receptor in pcDNA3.1 was transfected into HEK-293 cells and CHO cells, respectively, by using Lipofectamine 2000 reagent (Invitrogen, Carlsbad, CA). Once stable expression was achieved, membranes from cells expressing the appropriate receptor were prepared into membranes for competition and saturation studies by homogenizing in 50 mM Tris HCl, pH 7.5, containing 2.5 mM EDTA, 5 mM MgCl, and protease inhibitors (Complete, Roche Biochemicals, Baltimore, MD) (buffer A). The homogenate was centrifuged at 40 000*g* for 30 min. The pellet was resuspended in buffer A and centrifuged again. The final pellet was resuspended in buffer A at a concentration of 1 mg/mL protein and stored at −80 °C.

Receptor binding assays were performed in 96-well plates in a 200 μL final volume. Schering compounds and standards were made at a stock concentration of 10 mM in 100% DMSO. For single concentration determinations, the stock solution was diluted to a final concentration of 100 nM. For concentration curves, the stock solution was diluted by using 100% DMSO with a Multiprobe robotics device (Perkin-Elmer, Boston, MA), so that the final concentrations ranged from 0.1 to 3 μM. Drugs were diluted 20× into assay buffer A plus 0.4% BSA to yield a final DMSO concentration of 5%. Ligand concentrations of ^{3}H-CP-55,940 for these studies ranged from 0.2 to 0.4 nM, and protein concentrations for CB1 and CB2 membranes ranged from 2 to 20 μg protein. For both CB1 and CB2 receptor binding studies, nonspecific binding was defined in the presence of 10 μM final CP-55,940. All assays were incubated for 1.5 h and were terminated by vacuum filtration by using a Brandel cell harvester (Brandel Instruments, Gaithersburg, MD) over GF/C filter plates (Packard Instruments, Downers Grove, IL) presoaked in 0.3% polyethylenimine. Filter plates were washed with ice-cold buffer A plus BSA and dried, and after addition of Microscint 20 scintillation cocktail (Packard Instruments), the bound radioactivity was determined in a Topcount scintillation counter (Packard Instruments).

For data analysis, Microsoft Excel Fit 4.0 was used to determine IC_{50} values for each compound by using nonlinear regression analysis. These were converted to K_i values based upon predetermined affinity (K_d) determinations for each receptor preparation. K_d determinations were made by performing saturation analysis for each ligand at each receptor by using the same procedures as above, except that increasing concentrations of each ligand were incubated with the appropriate membranes in the presence and absence of CP-55,940. The K_d and Bmax values were then determined by using nonlinear regression analysis with GraphPad Prism software.[44]

Results

Training Set. Figure 1 shows the molecular structures of the eight CB1 selective compounds used for pharmacophore model development. The generic names or company codes are listed

as compound names when available, and the others are simply assigned as Compd *n*. Table 1 lists their corresponding CB1 and CB2 affinities as reported in the original literature where available. In some cases, for example, Compd 2, the exact K_i values were not reported in the original patents; hence, the affinities were listed as not reported. These compounds were included in the training set because they were not only shown in the original patents but also cited in other literature as potent selective CB1 antagonists. Furthermore, they represent structurally distinct classes that are useful for pharmacophore modeling.

A conformational search was carried out for each compound by using MacroModel as outlined in the Methods section. Table 1 also lists the number of unique conformations for each compound. Only one compound (Compd 3) reached the maximal unique conformation limit of 250; the energy of the last conformation was 29.9 kJ/mol above that of the minimum energy conformation. Visual inspection of the conformers showed that they provided adequate coverage of the low-energy conformational space for each compound.

CB1 Pharmacophore Model. The CB1 antagonist pharmacophore model was derived by using the Catalyst HipHop approach. This approach compares the conformational space of each molecule in the training set and identifies the common 3D configurations of functional features. On the basis of the structural characteristics of the training set molecules, the following chemical features were preselected for model generation: HBA, AR, and HPO. Rimonabant and SLV319 were selected as the principal molecules for their good CB1 potency and selectivity. A total of 10 pharmacophore models were generated by the HipHop hypothesis generation process, all of which had five features. Seven of the hypotheses have one HBA, two AR, and two HPO features, and the other three all have one HBA, one AR, and three HPO features. Given that AR is more selective than HPO, the hypotheses with only one AR feature were not further considered. Rimonabant is the most advanced compound in the training set; therefore, the remaining seven hypotheses were evaluated on the basis of the fitting (FAST approach) and the energy of the mapped conformation of this molecule. The model that provided a high fit value (fit = 4.61) with the lowest conformational energy (2.2 kJ/mol above its global energy minimum) for rimonabant was selected for virtual screening. Figure 2A shows the selected CB1 pharmacophore model. The distances among the centers of some features are labeled. Figure 2B shows the mapping of Rimonabant on this pharmacophore model. The carbonyl group matches the HBA feature, and the two phenyl groups match the two aromatic features. The carbon atoms of the piperidyl moiety and the chlorine atom on the monosubstituted phenyl ring match the two hydrophobic features.

This pharmacophore model highlights the interactions that are important for receptor binding and agrees with most of the findings from recently published CB1 homology studies. The carbonyl oxygen of rimonabant was suggested to be important for its activity.[45] A hydrogen bond between this carbonyl oxygen and the K192 residue of the CB1 receptor was proposed to be able to stabilize a K192−D366 salt bridge at the intracellular end of transmembrane helices 3 and 6. Formation of this hydrogen bond increases the affinity and shifts the receptor equilibrium toward the inactive state. The same receptor model shows that the binding of rimonabant is further enhanced by putative favorable aromatic stacking interactions between its 2,4-dichlorophenyl ring and the F200/W279/W356 residues on one side and the *para*-chlorophenyl ring and the W255/Y275/F278 residues on the other side. The lipophilic piperidinyl moiety fits nicely in a cavity formed by several hydrophobic residues Val196/Phe170/Met 384/Leu387.

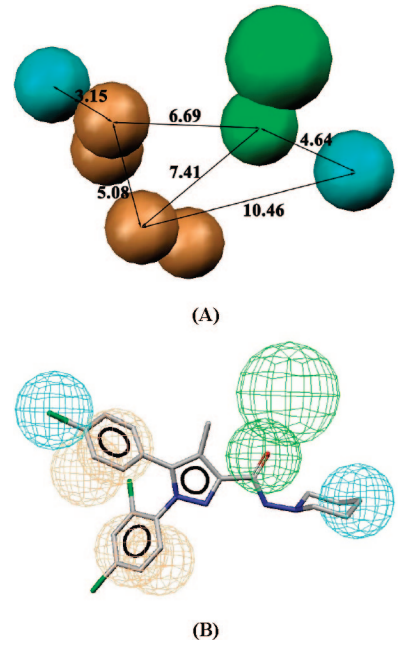
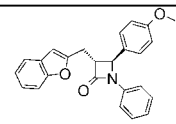
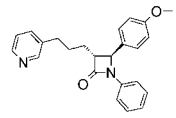
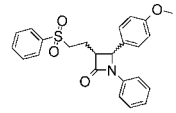
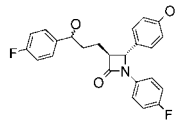
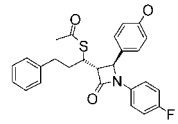
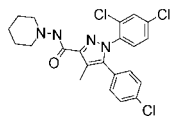


Figure 2. CB1 antagonist pharmacophore model used for virtual screening. (A) Geometrical relationships among the pharmacophore features. AR features are represented by pairs of solid brown spheres; HPO features are represented by cyan spheres; HBA are represented by a pair of green spheres (the smaller sphere represents the location of the HBA atom on the ligand and the larger one the location of an HB donor on the receptor). The distances (in Å) among the centers of the features are labeled. (B) Mapping of rimonabant to the pharmacophore model. The pharmacophore features are represented by mashed spheres. Carbon atoms are colored in gray, nitrogen atoms in blue, oxygen atoms in red, and chlorine atoms in light green. Hydrogen atoms are not shown for clarity.

Virtual Screening. The pharmacophore model derived by using the HipHop approach is expected to qualitatively discriminate between active and inactive molecules. To identify new CB1 antagonists, we carried out virtual screening on our compound collection by using this pharmacophore model. The pregenerated Catalyst 3D database contains about a half million molecules. The database search using the protocol specified in the Methods Section resulted in 22 794 hits. The large number of hits was expected, and they represented only about 5% of the compounds in the whole database. But this was much more than the number of compounds we planned to test; therefore, post pharmacophore filtering was implemented.

To narrow down the pharmacophore model hits to the desired number, we followed the stepwise filtering protocol outlined in the Methods section. First, we eliminated compounds with high molecular weight from further consideration. Filtering with $300 < MW < 550$ reduced the number of hits to 18 693. Second, we filtered the compounds on the basis of their availability in our compound distribution center at the time of screening. Only those compounds with an available amount larger than 2 mg were retained, and this further cut the number of hits down to 10 581. At this stage, it appeared we could afford to eliminate compounds with potential problematic pharmacokinetic properties. Accordingly, a modified Lipinski's rule-of-five was applied. As a result, more than two-thirds of the original hits were eliminated after these three filtering steps. A total number of 7247 compounds satisfying all the criteria so far was taken to the next step.

Table 2. CB1 Antagonists Identified by Virtual Screening

Compound Name	Structure	CB1 Ki (nM)	CB2 Ki (nM)	MW	ClogP	No. HBD[α]	No. HBA[α]	No. Rotatable Bonds[α]
Compd A		839.0	194.2	383	5.67	0	2	5
Compd B		243.8	1090.5	372	4.52	0	3	7
Compd C		451.5	1242.8	422	3.88	0	4	7
Compd D		301.4	>1818.2	409	3.96	2	3	6
Compd E		52.8	282.0	450	5.47	1	4	8
Rimonabant		2.4[β]	560[β]	464	6.47	1	4	4

[α] Number of HB donors, number of HB acceptors and number of rotatable bonds are calculated by using Pipeline Pilot software. [β] Obtained under the same assay conditions as the SPRI compounds.

In the last selection step, we employed a combination of Bayesian ranking and clustering to select the final set of 420 compounds for biological testing. For this purpose, we needed to develop a Bayesian model for CB1 antagonists. The MDDR database is a good source for our purpose because accurate potency data are not necessary for Bayesian modeling as long as the compounds can be classified categorically. This database collects biologically active compounds from published literature and covers structurally diverse drug candidates under various development stages. We found 78 CB1 antagonists in this database that could be used as the active set for our Bayesian modeling (the structures and the MDDR registration numbers of these compounds are provided as Supporting Information). To gather a negative control data set, we filtered the rest of the MDDR collection by using the criteria set forth in the Methods section. This resulted in a data set containing ~130 000 druglike molecules. Once we had the two data sets in place, it was straightforward to derive a CB1 antagonist Bayesian model by using Pipeline Pilot software's learn-good-molecule component. By using this binary QSAR model, a score representing the likelihood of being a CB1 antagonist was calculated for each remaining compound from the previous filtering steps.

Two considerations prompted us to use a combination of Bayesian modeling and clustering techniques in selecting our final test set. First, a Bayesian model provides only a crude estimate of the possibility of a compound being active; we did not want to rely too heavily on its ranking for compound selection. Second, our compound collection, like those of many other pharmaceutical companies, contains many molecules with minor structural changes as a result of historical drug discovery projects. The high structural redundancy observed among the hits did not best serve our purpose of finding novel classes of CB1 antagonists; we could use the limited biological screening resources more efficiently by exploring a larger structural space. On the basis of these considerations, we first selected the top ranking 2100 compounds by using the CB1 Bayesian model from the 7247 remaining hits. We then clustered the 2100 compounds into 420 groups by using the Pipeline Pilot maximum dissimilarity clustering algorithm (on average, five compounds per cluster). The compound with the highest Bayesian score in each cluster was selected as its representative. The molecular structures of these representatives were visually inspected, and they all seemed to be reasonably druglike. These 420 representatives constituted the final selection for biological tests.

Cannabinoid Receptor Binding Assays. The 420 compounds were first screened at a single concentration of 0.1 μM in a CB1 receptor competitive binding assay. Five compounds were found to have more than 50% inhibition at this concentration and were further characterized in follow-up assays to measure their affinities for both CB1 and CB2 receptors. Table 2 lists the assay results and their structures. All five compounds show good affinity for the CB1 receptor; even the weakest one, Compd A, has a $K_i < 1$ μM. (Compd A has a higher affinity for CB2 receptor; therefore, it is a potent CB2 selective agent.) The other four compounds all have CB1 $K_i < 0.5$ μM and show

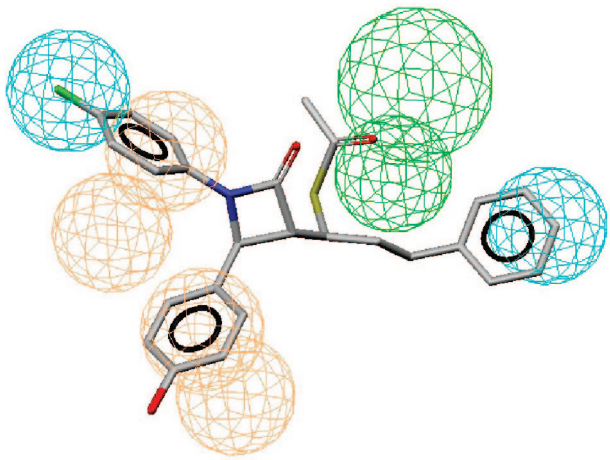


Figure 3. Mapping of the most potent hit, Compd E, to the CB1 antagonist pharmacophore model. Carbon atoms are colored in gray, oxygen atoms in red, nitrogen atoms in blue, sulfur atoms in yellow, and fluorine atoms in light green. Hydrogen atoms are not shown for clarity.

varied selectivities against the CB2 receptor. The most potent compound, Compd E, has a CB1 K_i of 53 nM and is more than 5-fold selective against CB2. CP-55,940, a nonselective cannabinoid agonist, was run as a standard in every assay performed, and generally, the K_i value of CP-55,940 was less than 3-fold different from our historical control. Moreover, rimonabant was tested under the same condition, and its K_i values for the CB1 and CB2 receptors were 2.4 nM and 560 nM, respectively, in line with previously reported values.[30]

Figure 3 shows the mapping of Compd E to the CB1 pharmacophore model. The thioester oxygen atom provides the HBA function; the three phenyl groups match two AR and one HPO features in the pharmacophore model; the fluorine atom and part of the fluorophenyl group fit the other HPO feature. The azetidinone core provides a nice scaffold for the crucial functional groups to reach their desired 3D locations. These compounds represent a novel class of CB1 antagonists that has not been reported before.

Discussion

The effectiveness of cannabinoid CB1 receptor antagonists in treating obesity has been demonstrated by the clinical studies of rimonabant. Even though the marketing of rimonabant in the United States was recently postponed by the FDA while waiting for more data on its psychiatric side effects, CB1 antagonists remain an attractive approach for the treatment of obesity and its related disorders. The search for new chemotypes that can reduce or eliminate these side effects becomes even more urgent. We developed a CB1 antagonist pharmacophore model and carried out virtual screening of our compound collection with this model. In the end, we were able to identify a new class of selective CB1 antagonists. These hits can serve as a starting point for lead optimization to further enhance the potency, selectivity, pharmacokinetic, and metabolic profiles and to reduce undesirable side effects.

Virtual screening has been established as a powerful alternative complementary to high-throughput screening. When virtual screening is performed optimally, impressive hit rates have been reported. Despite the fact that homology models based on the X-ray structure of bovine rhodopsin have been proposed for the CB1 receptor, they have not been used as virtual screening tools to identify new CB1 antagonists. Virtual screening with a receptor model is hindered by at least three factors: the model reliability, the absence of a good docking score, and the low throughput. Even though the development of homology models has been guided by ever increasing information from mutagenesis studies, many uncertainties, for example, the location and orientation of the transmembrane helices, the conformation of the sidechains, and the activation state of the receptor, still hamper the model accuracy because of the low sequence identity with the available GPCR templates. Furthermore, protein flexibility is rarely considered in structure-based virtual screening. Many scoring functions for ligand−protein docking have been developed in the past decade, and several successful examples have been reported, but they are still far from perfect and have yet to be proven adequate for predicting binding affinities. Docking is relatively slow, and currently, virtual screening of a half million compounds takes extensive computational resources.

On the other hand, pharmacophore models based solely on the ligand structures can be employed, as shown in this study, to identify new hits. To the best of our knowledge, this paper presents the first 3D CB1 antagonist pharmacophore model and demonstrates another success story of using a pharmacophore model for virtual screening. This model identifies the crucial chemical features for ligand−receptor interactions that are in agreement with the findings of the published receptor models. Both types of model showed that hydrogen bonding and aromatic interactions were very important for binding to the CB1 receptor. Pharmacophore models also offer a strong alternative to docking even when crystal structures of the target are available. On many occasions, protein-structure-based pharmacophore models have been used to prefilter compounds for docking because of their speed or to constrain structures with required interactions for reliable docking results.[25]

High-quality conformational models are crucial for the development of pharmacophore models. In this study, we employed MacroModel conformations instead of those generated by Catalyst for their better quality in covering the low-energy conformational space. For example, we have seen considerable differences between the two approaches in generating conformations for saturated six-member rings such as cyclohexane, piperidine, and piperazine. These groups appear frequently in drug molecules and are widely used as structural scaffolds. In crystal structures of druglike molecules or ligand−protein complexes, these saturated rings overwhelmingly adopt low-energy chair conformations. Catalyst generates predominantly twisted conformations and, therefore, leads to many incorrect mappings of functionally important groups. In comparison, the MacroModel conformational search produces a fair number of representative chair conformations for these groups. Of course, it takes a much longer time to generate the conformational models this way, but for model development, it is worthwhile to use better conformational models. We did use Catalyst conformations in 3D database generation where speed is more of a concern.

One primary issue with virtual screening by using pharmacophore models is the large number of hits retrieved. Postscreening filtering or ranking is usually needed to reduce the number of hits to a manageable scale. A practical consideration is the availability of the samples. Because the number of hits is plentiful, we can limit ourselves to only the readily available compounds, therefore, speeding up the experimental procedures. In many cases, the abundance of hits also affords the opportunity to retain only compounds with a better chance of having desirable pharmacokinetic properties. We applied a modified

rule-of-five filter in this study. Several modeling methods can be used to further narrow down the hits to a desired subset, for example, docking with a protein structure,[46] filtering with a QSAR model,[47,48] clustering, and similarity search. Bayesian modeling represents a useful approach when accurate biological activity data are not available.[49] This method provides unsupervised learning for large data collections with a simple two-class relationship, active or inactive. The resulting QSAR model can be used to estimate the likelihood of a compound being active and, therefore, a rough but rapid ranking of the virtual screening hits.

One way to reduce the number of hits from pharmacophore screening is to introduce additional constraints, such as shape and excluded volumes, to the pharmacophore model itself. The shape constraint can be derived from a mapped active compound. We did have some successful experiences with the shape-based approach in the past, especially in cases where pharmacophore features crowded into a limited space; however, we chose to simply use a molecular weight filter in this study. The dimension of our CB1 pharmacophore model is quite large, the most distant features (the two HPO's) being 14 Å apart. When a compound with MW < 550 maps on this pharmacophore model, its volume overlap with rimonabant is quite high and often leads to a reasonable shape similarity. Therefore, for the sake of speed in database search, we did not use the shape constraint in our virtual screening. Another modification to make a pharmacophore model more specific is to use excluded volumes. The excluded volumes can be derived from structurally similar but inactive compounds with the assumption that the inactive compounds protrude into the receptor because of their bulky size. Alternatively, excluded volumes can be derived from the locations of the receptor atoms constituting the binding site.

Our objective was focused on identifying a new chemical class rather than finding hits from the same structural family. To maximize the chance of finding novel chemotypes, we needed to balance highly ranked compounds in the Bayesian model and the structural diversity among the compounds selected for biological assays. Consequently, we applied a combination of Bayesian ranking and structural clustering to choose the final screening subset. The CB1 Bayesian model developed by using known CB1 antagonists in the MDDR database was by no means a perfect one; hence, the ranking by this model was not used as an exclusive criterion for compound selection. The predictive power of Bayesian modeling, like 2D similarity methods, is usually limited to the scope of the chemotypes in the training set. On the other hand, in the absence of other reliable means, this model did provide us with a way to rank the hits and may enrich true actives among the top ranking compounds compared to random selection. Another approach to reduce the number of hits to a preset assay capacity is to simply cluster them and choose the cluster centers. Although this approach can cover more structural diversity, it favors compounds from very small clusters and, in some cases, even singletons. In our situation, where the selection ratio is about 17:1, the compounds in the same cluster are not quite similar in a chemist's eyes. The bias against hits from large structural clusters under such circumstance may not be desirable. On the basis of the limited biological resources, it seems that using a combination strategy instead of relying solely on either approach can provide us with an optimal balance between structural diversity and the chance of finding actives. On the basis of these considerations, we decided to select a subset of the top ranking 2100 compounds from the 7247 hits by using the Bayesian model. They were then clustered into 420 groups to reduce the structural redundancy, and the Bayesian score was used again to select one member from each cluster to give the final set of 420 compounds.

This study demonstrated the power of virtual screening with a pharmacophore model, and using the modeling techniques in concert was essential for our success. The bar was set high for this exercise; only compounds with CB1 inhibition larger than 50% at 0.1 μM were considered hits. This is rarely seen in the literature, where most of the studies consider compounds with IC_{50} or K_i lower than 10 μM as hits. The most potent hit had a double digit nanomolar affinity for the CB1 receptor and more than 5-fold selectivity against the CB2 receptor. The CB1 antagonists identified belong to a new chemical class that has never been reported. The hits represent promising candidates for further lead optimization: very good potency and selectivity and proper molecular weight and log *P*, as well as feasible chemistry. Indeed, we believe that the following factors helped ensure the drug likeness of our hits: the screening database was geared toward drug discovery programs; most undesirable compounds were filtered out before Catalyst database generation; and a rule-of-five filter was applied in hits selection. For future lead optimization, we can combine the pharmacophore model with 3D-QSAR methods like CoMFA to generate quantitative predictions. Pharmacophore models like this one can provide a good alignment template for 3D-QSAR studies. Additionally, a CB2 pharmacophore model may be used to improve the selectivity.

Conclusion

In summary, we have developed a pharmacophore model by using representative CB1 receptor antagonists. We demonstrated that we can screen a large chemical database with this ligand-based pharmacophore model to identify new CB1 antagonists. Usually, the number of hits from such virtual screening exceeds the number of compounds one intends to test in the biological assays. In this study, we applied a stepwise filtering protocol to funnel down the number of initial hits to a manageable scale. We then developed a Bayesian model and applied clustering techniques to maximize the chance of finding new chemotypes while retaining the top scoring compounds. In the end, we were able to identify a novel structural class of CB1 antagonists with good potency and selectivity. These hits can be subjected to further chemical optimization for the treatment of obesity and its related diseases.

Acknowledgment. The authors thank Dr. Johannes Voigt and Dr. Eduardo Zaborowski for generating and maintaining a Catalyst database of Schering-Plough compounds. This database was used in the virtual screening.

Supporting Information Available: Supplementary Table 1 lists the CB1 antagonists/inverse agonists from the MDDR database that were used to derive the Bayesian model. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Pertwee, R. G. The pharmacology of cannabinoid receptors and their ligands: An overview. *Int. J. Obesity* **2006**, *30*, S13–18.
(2) Mokdad, A. H.; Marks, J. S.; Stroup, D. F.; Gerberding, J. L. Actual causes of death in the United States, 2000. *J. Am. Med. Assoc.* **2005**, *291*, 1238–1245.
(3) Mokdad, A. H.; Ford, E. S.; Bowman, B. A.; Dietz, W. H.; Vinicor, F.; et al. Prevalence of obesity, diabetes, and obesity-related health risk factors, 2001. *J. Am. Med. Assoc.* **2003**, *289*, 76–79.
(4) Li, C.; Ford, E. S.; Mokdad, A. H.; Cook, S. Recent trends in waist circumference and waist-height ratio among US children and adolescents. *Pediatrics* **2006**, *118*, 1390.

(5) Engelgau, M. M.; Geiss, L. S.; Saaddine, J. B.; Boyle, J. P.; Benjamin, S. M.; et al. The evolving diabetes burden in the United States. *Ann. Intern. Med.* **2004**, *140*, 945–950.
(6) Carai, M. A.; Colombo, G.; Gessa, G. L. Rimonabant: The first therapeutically relevant cannabinoid antagonist. *Life Sci.* **2005**, *77*, 2339–2350.
(7) Gelfand, E. V.; Cannon, C. P. Rimonabant: A cannabinoid receptor type 1 blocker for management of multiple cardiometabolic risk factors. *J. Am. Coll. Cardiol.* **2006**, *47*, 1919–1926.
(8) Woods, S. C. Role of the endocannabinoid system in regulating cardiovascular and metabolic risk factors. *Am. J. Med.* **2007**, *120*, S19–25.
(9) Boyd, S. T. The endocannabinoid system. *Pharmacotherapy* **2006**, *26*, S218–221.
(10) Piomelli, D. The endocannabinoid system: A drug discovery perspective. *Curr. Opin. Invest. Drugs* **2005**, *6*, 672–679.
(11) Fride, E.; Foox, A.; Rosenberg, E.; Faigenboim, M.; Cohen, V.; et al. Milk intake and survival in newborn cannabinoid CB1 receptor knockout mice: Evidence for a CB3 receptor. *Eur. J. Pharmacol.* **2003**, *461*, 27–34.
(12) Palczewski, K.; Kumasaka, T.; Hori, T.; Behnke, C. A.; Motoshima, H.; et al. Crystal structure of rhodopsin: A G protein-coupled receptor. *Science* **2000**, *289*, 739–745.
(13) Tuccinardi, T.; Ferrarini, P. L.; Manera, C.; Ortore, G.; Saccomanni, G.; et al. Cannabinoid CB2/CB1 selectivity. Receptor modeling and automated docking analysis. *J. Med. Chem.* **2006**, *49*, 984–994.
(14) Shim, J.-Y.; Howlett, A. C. WIN55212–2 docking to the CB1 cannabinoid receptor and multiple pathways for conformational induction. *J. Chem. Inf. Model.* **2006**, *46*, 1286–1300.
(15) Montero, C.; Campillo, N. E.; Goya, P.; Paez, J. A. Homology models of the cannabinoid CB1 and CB2 receptors. A docking analysis study. *Eur. J. Med. Chem.* **2005**, *40*, 75–83.
(16) Salo, O. M. H.; Lahtela-Kakkonen, M.; Gynther, J.; Jarvinen, T.; Poso, A. Development of a 3D model for the human cannabinoid CB1 receptor. *J. Med. Chem.* **2004**, *47*, 3048–3057.
(17) Shim, J. Y.; Welsh, W. J.; Cartier, E.; Edwards, J. L.; Howlett, A. C. Molecular interaction of the antagonist *N*-(piperidin-1-yl)-5-(4-chlorophenyl)-1-(2,4-dichlorophenyl)-4-methyl-1*H*-pyrazole-3-carboxamide with the CB1 cannabinoid receptor. *J. Med. Chem.* **2002**, *45*, 1447–1459.
(18) Hurst, D. P.; Lynch, D. L.; Barnett-Norris, J.; Hyatt, S. M.; Seltzman, H. H.; et al. *N*-(piperidin-1-yl)-5-(4-chlorophenyl)-1-(2,4-dichlorophenyl)-4-methyl-1*H*-pyrazole-3-carboxamide (SR141716A) interaction with LYS 3.28(192) is crucial for its inverse agonism at the cannabinoid CB1 receptor. *Mol. Pharmacol.* **2002**, *62*, 1274–1287.
(19) Lundstrom, K. Latest development in drug discovery on G protein-coupled receptors. *Curr. Protein Pept. Sci.* **2006**, *7*, 465–470.
(20) Chen, J.-Z.; Wang, J.; Xie, X.-Q. GPCR structure-based virtual screening approach for CB2 antagonist search. *J. Chem. Inf. Model.* **2007**, *47*, 1626–1637.
(21) Lopez-Rodriguez, M. L.; Benhamu, B.; de la Fuente, T.; Sanz, A.; Pardo, L.; et al. A three-dimensional pharmacophore model for 5-hydroxytryptamine6 (5-HT6) receptor antagonists. *J. Med. Chem.* **2005**, *48*, 4216–4219.
(22) Guner, O.; Clement, O.; Kurogi, Y. Pharmacophore modeling and three dimensional database searching for drug design using catalyst: Recent advances. *Curr. Med. Chem.* **2004**, *11*, 2991–3005.
(23) Catalyst *Catalyst Software*; Accelrys Inc.: San Diego.
(24) Pandit, D.; So, S. S.; Sun, H. Enhancing specificity and sensitivity of pharmacophore-based virtual screening by incorporating chemical and shape features: A case study of HIV protease inhibitors. *J. Chem. Inf. Model.* **1236**, *46*, 1236–1244.
(25) Toba, S.; Srinivasan, J.; Maynard, A. J.; Sutter, J. Using pharmacophore models to gain insight into structural binding and virtual screening: An application study with CDK2 and human DHFR. *J. Chem. Inf. Model.* **2006**, *46*, 728–735.
(26) Laggner, C.; Schieferer, C.; Fiechtner, B.; Poles, G.; Hoffmann, R. D.; et al. Discovery of high-affinity ligands of sigma1 receptor, ERG2, and emopamil binding protein by pharmacophore modeling and virtual screening. *J. Med. Chem.* **2005**, *48*, 4754–4764.
(27) Vadivelan, S.; Sinha, B. N.; Irudayam, S. J.; Jagarlapudi, S. A. R. P. Virtual screening studies to design potent CDK2-cyclin A inhibitors. *J. Chem. Inf. Model.* **2007**, *47*, 1526–1535.
(28) Schuster, D.; Maurer, E. M.; Laggner, C.; Nashev, L. G.; Wilckens, T.; et al. The discovery of new 11 beta-hydroxysteroid dehydrogenase type 1 inhibitors by common feature pharmacophore modeling and virtual screening. *J. Med. Chem.* **2006**, *49*, 3454–3466.
(29) Langer, T.; Krovat, E. M. Chemical feature-based pharmacophores and virtual library screening for discovery of new leads. *Curr. Opin. Drug Discovery Dev.* **2003**, *6*, 370–376.
(30) Rinaldi-Carmona, M.; Barth, F.; Heaulme, M.; Alonso, R.; Shire, D.; et al. Biochemical and pharmacological characterisation of SR141716A, the first potent and selective brain cannabinoid receptor antagonist. *Life Sci.* **1995**, *56*, 1941–1947.
(31) Ruiu, S.; Pinna, G. A.; Marchese, G.; Mussinu, J. M.; Saba, P.; et al. Synthesis and characterization of NESS 0327: A novel putative antagonist of the CB1 cannabinoid receptor. *J. Pharmacol. Exp. Ther.* **2003**, *306*, 363–370.
(32) Lange, J. H.; Coolen, H. K.; van Stuivenberg, H. H.; Dijksman, J. A.; Herremans, A. H.; et al. Synthesis, biological properties, and molecular modeling investigations of novel 3,4-diarylpyrazolines as potent and selective CB1 cannabinoid receptor antagonists. *J. Med. Chem.* **2004**, *47*, 627–643.
(33) Felder, C. C.; Joyce, K. E.; Briley, E. M.; Glass, M.; Mackie, K. P.; et al. LY320135, a novel cannabinoid CB1 receptor antagonist, unmasks coupling of the CB1 receptor to stimulation of cAMP accumulation. *J. Pharmacol. Exp. Ther.* **1998**, *284*, 291–297.
(34) Achard, D. B. H.; Bouquerel, J.; Capet, M.; Grisoni, S.; Malleron, J. L.; Mignani, S.; Hittinger, A. Acetindine derivatives, their preparation and medicaments containing them; Aventis Pharma SA: US, 2003.
(35) Hagmann, W. K.; Lin, L. S.; Shah, S. K. Substituted aryl amides. Merck & Co., Inc., 2003.
(36) Ooms, F.; Wouters, J.; Oscari, O.; Happaerts, T.; Bouchard, G.; et al. Exploration of the pharmacophore of 3-alkyl-5-arylimidazolidinediones as new CB1 cannabinoid receptor ligands and potential antagonists: Synthesis, lipophilicity, affinity, and molecular modeling. *J. Med. Chem.* **2002**, *45*, 1748–1756.
(37) Makriyannis, A.; Liu, Q.; Thotapally, R. Novel pyrazole analogs acting on cannabinoid receptors. *University of Connecticut*; University of Connecticut, 2004.
(38) Pearlman, R. S. *Concord*; distributed by Tripos Inc.: St. Louis, MO.
(39) MacroModel *MacroModel Software*; Scherodinger, LLC.: New York.
(40) Veber, D. F.; Johnson, S. R.; Cheng, H. Y.; Smith, B. R.; Ward, K. W.; et al. Molecular properties that influence the oral bioavailability of drug candidates. *J. Med. Chem.* **2002**, *45*, 2615–2623.
(41) 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.
(42) PipelinePilot *Pipeline Pilot*; SciTegic, Inc.: San Diego.
(43) MDDR *MDDR database*; MDL Information Systems, Inc.: San Leandro, CA.
(44) Prism *Prism software*; GraphPad Software, Inc.: San Diego, CA.
(45) Hurst, D.; Umejiego, U.; Lynch, D.; Seltzman, H.; Hyatt, S.; et al. Biarylpyrazole inverse agonists at the cannabinoid CB1 receptor: Importance of the C-3 carboxamide oxygen/lysine3.28(192) interaction. *J. Med. Chem.* **2006**, *49*, 5969–5987.
(46) Warren, G. L.; Andrews, C. W.; Capelli, A. M.; Clarke, B.; LaLonde, J.; et al. A critical assessment of docking programs and scoring functions. *J. Med. Chem.* **2006**, *49*, 5912–5931.
(47) Chen, J. Z.; Han, X. W.; Liu, Q.; Makriyannis, A.; Wang, J.; et al. 3D-QSAR studies of arylpyrazole antagonists of cannabinoid receptor subtypes CB1 and CB2. A combined NMR and CoMFA approach. *J. Med. Chem.* **2006**, *49*, 625–636.
(48) Fichera, M.; Cruciani, G.; Bianchi, A.; Musumarra, G. A 3D-QSAR study on the structural requirements for binding to CB1 and CB2 cannabinoid receptors. *J. Med. Chem.* **2000**, *43*, 2300–2309.
(49) Hassan, M.; Brown, R. D.; Varma-O'brien, S.; Rogers, D. Cheminformatics analysis and learning in a data pipelining environment. *Mol. Diversity* **2006**, *10*, 283–299.

JM701519H