Find novel dual-agonist drugs for treating type 2 diabetes by means of cheminformatics

Lei Liu1
Ying Ma2
Run-Ling Wang2
Wei-Ren Xu3
Shu-Qing Wang2,5
Kuo-Chen Chou1,5

1PET/CT Center, General Hospital of Tianjin Medical University, Tianjin, People’s Republic of China; 2Tianjin Key Laboratory on Technologies Enabling Development of Clinical Therapeutics and Diagnostics (Theranostics), School of Pharmacy, Tianjin Medical University, Tianjin, People’s Republic of China; 3Tianjin Institute of Pharmaceutical Research (TIPR), Tianjin, People’s Republic of China; 4Center of Excellence in Genomic Medicine Research (CEGMR), King Abdulaziz University, Jeddah, Saudi Arabia; 5Gordon Life Science Institute, Belmont, MA, USA

Abstract: The high prevalence of type 2 diabetes mellitus in the world as well as the increasing reports about the adverse side effects of the existing diabetes treatment drugs have made developing new and effective drugs against the disease a very high priority. In this study, we report ten novel compounds found by targeting peroxisome proliferator-activated receptors (PPARs) using virtual screening and core hopping approaches. PPARs have drawn increasing attention for developing novel drugs to treat diabetes due to their unique functions in regulating glucose, lipid, and cholesterol metabolism. The reported compounds are featured with dual functions, and hence belong to the category of dual agonists. Compared with the single PPAR agonists, the dual PPAR agonists, formed by combining the lipid benefit of PPARα agonists (such as fibrates) and the glycemic advantages of the PPARγ agonists (such as thiazolidinediones), are much more powerful in treating diabetes because they can enhance metabolic effects while minimizing the side effects. This was observed in the studies on molecular dynamics simulations, as well as on absorption, distribution, metabolism, and excretion, that these novel dual agonists not only possessed the same function as ragaglitazar (an investigational drug developed by Novo Nordisk for treating type 2 diabetes) did in activating PPARα and PPARγ, but they also had more favorable conformation for binding to the two receptors. Moreover, the residues involved in forming the binding pockets of PPARα and PPARγ among the top ten compounds are explicitly presented, and this will be very useful for the in-depth conduction of mutagenesis experiments. It is anticipated that the ten compounds may become potential drug candidates, or at the very least, the findings reported here may stimulate new strategies or provide useful insights for designing new and more powerful dual-agonist drugs for treating type 2 diabetes.

Keywords: diabetes, PPAR-alpha, PPAR-gamma, dual-agonist drug, core hopping, molecular docking, ADME, binding pocket

Introduction

Type 2 diabetes mellitus (T2DM) is the most common type of diabetes mellitus, which is characterized by insulin resistance and combined with relatively reduced insulin secretion.1 It is a chronic metabolic disease that affects the body’s ability to turn food into energy. People developing T2DM may suffer from blindness, renal failure, coronary artery disease, and so forth.2 Because of its dramatic increase worldwide, T2DM has reached an epidemic scale that is anticipated to affect over 360 million people by 2030.3,4

Belonging to the nuclear hormone receptor family, peroxisome proliferator-activated receptors (PPARs) are the ligand-activated transcription factors.5,6 PPARs, particularly the subtypes PPARα and PPARγ, played a central role in regulating the expression of genes involved in the control of lipid and lipoprotein metabolism, glucose
homeostasis, and inflammatory processes. Accordingly, activating these receptors has become a promising therapeutic strategy for designing drugs against T2DM.\(^7\)

The activated PPAR\(\alpha\) receptors have a beneficial effect on lipid metabolism by regulating genes on fatty acid and cholesterol metabolism.\(^8\) PPAR\(\alpha\) agonists belong to the class of lipid-lowering drugs (such as fenofibrate and gemfibrozil), and are widely prescribed to reduce triglycerides.\(^9\) The PPAR\(\gamma\) receptors, the most widely studied subtype of PPARs, are located in adipocytes, muscles, and macrophages, where they have a direct influence on T2DM, dyslipidemia, atherosclerosis, and cardiovascular diseases.\(^10,11\) Thiazolidinediones are a class of PPAR\(\gamma\) agonists used in clinical practice to reduce plasma glucose level in type 2 diabetic patients, or to regulate glucose homeostasis by increasing insulin sensitivity and glucose disposal.\(^12\)

Unfortunately, if the drug were used alone, the undesirable side effects such as weight gain, edema, and anemia would be caused;\(^13\) therefore, it is an urgent and challenging task to develop new dual agonists. In this regard, the novel PPAR\(\alpha/\gamma\) dual agonist, formed by combining the lipid benefit of PPAR\(\alpha\) agonists (such as fibrates) with the glycemic advantages of the PPAR\(\gamma\) agonists (such as thiazolidinediones), has drawn considerable attention.\(^14\) In developing the dual agonists, a critical problem surrounds how to identify the selectivity ratio of the receptor subtype,\(^15\) because it may provide useful insights for finding the new drug candidates. There is more information available about designing potential dual treatments for treating T2DM.\(^16\)

Many studies have indicated that computational approaches, such as structural bioinformatics, molecular docking, pharmacophore modeling, Quantitative Structure−Activity Relationship (QSAR) techniques, and a series of user-friendly web server predictors developed recently, such as G-protein-coupled receptors-grey incidence degree analysis (GPCR-GIA) and G-protein-coupled receptors -2-layer predictor (GPCR-2L) for identifying G protein-coupled receptors and their types,\(^25,26\) iLoc-Euk and iLoc-Hum for predicting subcellular localization of eukaryotic and human proteins,\(^27,28\) NR-2L for identifying nuclear receptors and their subfamilies,\(^29\) ProtIdent for identifying proteases and their types,\(^30\) and HIVcleave for predicting human immunodeficiency virus protease cleavage sites in proteins,\(^31,32\) can provide timely and very useful information, as well as insights for drug development. In addition to the aforementioned approaches, the Virtual Screening technique and Core Hopping technique are also very useful.\(^33–36\) The former can be used to screen for novel molecular scaffolds, while the latter can be used to generate new lead compounds with improved core properties.\(^37,38\)

Encouraged by the aforementioned successful studies, the present study was initiated in an attempt to screen the fragment database in hopes to find a new antidiabetic compound. Meanwhile, the techniques of the core hopping associated with the glide docking and molecular dynamic simulation were utilized to analyze the binding interactions between the agonist and PPARs.\(^39,40\) The absorption, distribution, metabolism, and excretion (ADME) predictions were also used to evaluate whether the new agonist found to date possesses great potential to become a promising drug candidate for treating diabetes mellitus.\(^41,42\)

### Materials and methods

The representative complex crystal structures of PPAR\(\alpha\) (Protein Data Bank [PDB] ID 1k7l) and PPAR\(\gamma\) (PDB ID 1k74) were downloaded from the PDB,\(^43,44\) and were to be used for the molecular modeling studies.

Our calculations are operated with the Schrödinger software package (version 9.0.111; Schrödinger, LLC, New York, NY, USA), which runs on a Dell Precision™ T5500 computer (Dell, Round Rock, TX, USA). The molecular dynamic simulation was performed using the Gromacs 4.0 package (www.gromacs.com) for the Linux system.

### Protein structures and databases

We downloaded the crystal structures of PPAR\(\alpha\) (1k7l.pdb) and PPAR\(\gamma\) (1k74.pdb) from the Protein Data Bank,\(^43,44\) and used them as receptors for the current modeling. The reasons behind the selection of the two proteins are: (1) the two proteins contain the same ligand, GW409544, which is used to easily identify the binding site for comparison; (2) the resolutions of the two crystal structures are quite similar with one (1k7l.pdb) having 2.50 Å and the other (1k74.pdb) having 2.30 Å; (3) the source organism of both structures was humans.\(^43\)

The protein-ligand binding site was identified by the SiteMap tool embedded in the Schrödinger Suite 2009 (Schrödinger, LLC), as described previously.\(^45–47\) The binding site encompassed the ligand GW409544, which was observed in the crystal structures of both PPAR\(\alpha\) (1k7l) and PPAR\(\gamma\) (1k74). The binding site of PPARs is composed of three arms: arm 1, arm 2, and arm 3, as marked in the panels A and B of Figure 1. These three segments play an important role in the interactions between the receptor and ligand.

Several hydrophobic interactions between the three “arms” and the Y-shaped ligand were deemed as the key points for designing the new PPARs agonist.\(^48\)
Arm 1 has a predominately polar character and includes the AF2-helix (shown as the purple ribbon in Figure 1), which is responsible for the function of transcriptional activation. The vital conformation of AF2-helix could generate a charge clamp formed by the network hydrogen bonds, thus significantly reducing the mobility of AF2 and regulating gene expression. The hydrophobic tail of the agonist interacts with either arm 2 or arm 3 to improve binding affinity.

The lead compounds and the fragment databases derived from ZINC (Shoichet Laboratory, Department of Pharmaceutical Chemistry, University of California, San Francisco, CA, USA) were used for virtual screening and core hopping searching, as described below, respectively.

Virtual screening

ZINC (Shoichet Library, Department of Pharmaceutical Chemistry, University of California) was a free database of commercially available compounds and contained over 13 million purchasable compounds in three-dimensional formats. The lead compound (or “lead-now”) database, one subdatabase of the ZINC, was chosen for virtual screening to find the optimal molecular scaffold by the Glide5 docking program (Schrödinger, LLC) which was interfaced with Schrödinger Suite 2009 (Schrödinger, LLC). The preparation and refinement protocols for the protein receptor and all compound structures were docked into the receptor pocket through the rigid docking model with a high throughput virtual screening scoring function to estimate the binding affinities.

Molecular docking with core hopping method

Many useful insights for drug design could be gained by conducting molecular docking studies. In this study, a new docking algorithm called “core hopping” was employed. Core hopping offers function of being able to perform both the fragment-based replacing and molecular docking, which makes it a very powerful and cutting-edge technique for de novo drug design because it allows for rapidly screening novel cores to help overcome unwanted properties by generating new lead compounds with improved core properties, as demonstrated by a series of previous studies.

The procedures of the core hopping computation can be briefly described as follows. The first step was to define the possible points to which the cores were attached; this could be done by using the program “Define Combinations” from the Combinatorial Screening panel in Schrodinger 2009 (Schrödinger, LLC). The second step was to define the “receptor grid file,” which could be done in the “Receptor Preparation” panel. The third step was the cores preparation; this could be done by operating the “Protocore Preparation” module to find the cores attached to the scaffold and using the fragment database derived from ZINC.

The fourth step was to align and dock the entire molecular structure built up by the core and scaffold. After all the above four steps were accomplished, all the investigated compounds were redocked into the receptor pocket via the rigid protein
docking model with the Stand-Precision scoring function to estimate the binding affinities.39,40

Molecular dynamics simulations

Many marvelous biological functions in proteins and deoxyribonucleic acid (DNA) and their profound dynamic mechanisms, such as switching between active and inactive states,54,55 cooperative effects,56 allosteric transition,57–59 intercalation of drugs into DNA,60 and assembly of microtubules,61 can be revealed by studying their internal motions.62 Likewise, to really understand the interaction mechanism of a receptor with its ligand, we should consider not only the static structures concerned, but also the dynamic information obtained by simulating their internal motions or dynamic processes.63,64

To examine whether the designed agonist remains bound by an explicit solvent from a dynamic point of view, the molecular dynamic simulation was performed with the periodic boundary conditions using the GROMACS 4.0 package under GROMACS 96-53a6 force fields for proteins.65 The topology files and charges for the ligand atoms were generated by the Dundee PRODRG2.5 Server (beta).66 Before starting the simulations, the systems were solvated in explicit water with a flexible simple point charge model in a cubic box, under the condition that the surface of the protein was covered with a water shell of 1.0 nm (nanometer or 10\(^{-9}\) of a meter). The minimum number of sodium and chloride ions needed to balance the system charge was placed randomly in the solvated system. Subsequently, the energy minimization was performed for the complex system by using the steepest descent method.67,68 Then, the 10 ns (or 10\(^{-9}\) of a second) MD simulations were carried out with a time step of 2 fs (femtosecond, or 10\(^{-15}\) of a second). The corresponding coordinates are stored every 2 ps (picosecond, or 10\(^{-12}\) of a second). The Particle Mesh Ewald algorithm was used to calculate the electrostatic interactions.69 All simulations were run under the periodic boundary condition with the Normal Pressure and Temperature (NPT) ensemble by using Berenson’s coupling algorithm for keeping the temperature at 310 K and the pressure at 1 atm. All bonds are constrained by using the Linear Constraint Solver algorithm.70

ADME prediction

ADME is a very useful program to eliminate unfavorable compounds with undesired properties in the early stage of drug development via evaluating the proposed structural refinements prior to expensive synthesis.27,42,71 Incorporating ADME predictions as a part of the drug development process can generate lead compounds that are more likely to exhibit satisfactory ADME performances during clinical trials.

The program of QikProp embedded in Schrödinger 2009 (Schrödinger, LLC) is a quick, accurate, and easy-to-use program for predicting the ADME properties of the compounds.72,73 The QikProp program adopted here was featured by having the functions predict a total of 44 properties of compounds, including the principal descriptors and physiochemical properties.

All the investigated compounds need not be neutralized before using QikProp because it will be automatically done in QikProp. The normal mode of prediction of QikProp was used in this study. The property analysis for the molecular weight – partition coefficient (QP logP o/w) – predicted aqueous solubility (QP logS); in addition, apparent PMDCK permeability (QPP MDCK) was computed in the QikProp program to evaluate whether the compounds hold the potential to become drug candidates.74

Results and discussion

Virtual screening and core hopping

The detailed procedure of discovering the desired dual agonists is shown in Figure 2. The lead-now database from ZINC was screened by utilizing the Glide5 software for its optimal performance targeting the PPAR receptors (PPAR\(\alpha\):1k71 and PPAR\(\gamma\):1k74).33,36,49 The top hit, ZINC36728034 (see Figure 2) was selected as the optimal lead compound for further modification according to the well known structures of ragaglitazar and GW409544, as shown in Figure 2.43,75 Consequently, the phenyl group rendered by red color in ZINC36728034 (see Figure 2) is modified into an acidic group colored with blue, as shown in Comp#0 (Figure 2). The functional acidic group also exits in the following structures: muraglitazar,76 faraglitazar,77 GW6471,78 GW7845,77 CHEMBL166572,79 and so on. The new molecule Comp#0 was designed with the intention so it has stronger affinity than ZINC36728034. Subsequently, the core hopping method was used to search the fragment database for the desired chemical groups to replace the amide group (ie, the R\(_0\) group) (Figure 2). Doing this was to help lengthen the Comp#0 structure so as to make it better fit the roomy space in arm 2 (see Figure 1). As a result, the amide group (ie, R\(_0\) group) of Comp#0 in Figure 2 was replaced by each of the best ten R groups (ie, the R\(_0\) to R\(_{10}\), as marked on the right side of Figure 2). Subsequently, they were used to generate the whole molecules (ie, Comp#1 to Comp#10) through the core hopping approach (see Figure 2). The ten compounds displayed the same chiral center (S) of the
The best ten potential lead compounds or candidates obtained via the above step were redocked into the two receptors, PPARα (1k71) and PPARγ (1k74), respectively. From the docking simulation results, it can be seen that all the compounds form good steric complementarity with the hydrophobic pocket of each receptor and produce strong van der Waals contacts and hydrophobic interactions. As shown in the black in panels C and D of Figure 1, ragaglitazar (a novel dual agonist for PPARα and PPARγ that also includes the important polar acidic head group in its structure) was chosen as a positive control for its high affinity with the PPARα and PPARγ receptors in the previous research.\textsuperscript{75,81–83}

Listed in Table 1 are the top ten compounds (ie, Comp#1 to Comp#10) that are much stronger than ragaglitazar and Comp#0 in binding to the two PPARα (1k71) and PPARγ (1k74) receptors, and these are ranked roughly according to their docking scores to the receptors. The best docked conformations of Comp#1 (carbon atoms colored in gray) and ragaglitazar (carbon atoms colored in black) are aligned together in panels C and D of Figure 1. The results of receptor–ligand interactions obtained from the docking simulation had proved that the key residues for the binding interactions between Comp#1 and the receptor were fully consistent with the previous reports.\textsuperscript{45,71,84} The conservative H-bonding network formed by the polar acidic head group in both ragaglitazar and Comp#1 to the four key residues of PPARα (or PPARγ), such as Ser280 (or Ser289), Tyr314 (or His323), Tyr464 (or Tyr473), and His440 (or His449) were observed in our docking simulation. It is noteworthy that Comp#1 has three oxygen atoms to serve as the polar acidic head, which allows Comp#1 to have more opportunities to interact with the AF2-helix. In addition, one more H-bond formed between the oxygen atom of the R1 group and Ala333 residue only in the PP ARγ receptors in arm 11, which is extraordinarily important for the regulation of gene expression.\textsuperscript{7} The hydrophobic part of Comp#1 derived from ZINC36728034 played an important role in complementing the hydrophobic arm 1 formed by the hydrophobic residues, as shown by the light green surface (see panels C and D of Figure 1). Compared with ragaglitazar, the Comp#1 compound has an extra hydrophobic part, making it more fitted to the hydrophobic arm 1, and hence resulting in a much better binding affinity than ragaglitazar, as clearly indicated by the docking scores shown in Table 1. Among the compounds listed in
Table 1 The compound ragaglitazar was used as a positive control, and the ten compounds (Comp#1–#10) were ranked roughly according to their docking scores to the receptors PPARα and PPARγ.

<table>
<thead>
<tr>
<th>Compound</th>
<th>Docking scores (Kcal/mol)</th>
<th>ADME properties predicted</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PPARα (1k71)</td>
<td>PPARγ (1k74)</td>
</tr>
<tr>
<td>Ragaglitazar</td>
<td>−11.49</td>
<td>−12.29</td>
</tr>
<tr>
<td>Comp#0</td>
<td>−10.33</td>
<td>−11.79</td>
</tr>
<tr>
<td>Comp#1</td>
<td>−13.65</td>
<td>−14.98</td>
</tr>
<tr>
<td>Comp#2</td>
<td>−13.64</td>
<td>−14.58</td>
</tr>
<tr>
<td>Comp#3</td>
<td>−13.31</td>
<td>−14.55</td>
</tr>
<tr>
<td>Comp#4</td>
<td>−13.22</td>
<td>−14.41</td>
</tr>
<tr>
<td>Comp#5</td>
<td>−13.18</td>
<td>−13.85</td>
</tr>
<tr>
<td>Comp#6</td>
<td>−13.01</td>
<td>−13.68</td>
</tr>
<tr>
<td>Comp#7</td>
<td>−12.95</td>
<td>−14.11</td>
</tr>
<tr>
<td>Comp#8</td>
<td>−13.85</td>
<td>−13.07</td>
</tr>
<tr>
<td>Comp#9</td>
<td>−13.36</td>
<td>−13.45</td>
</tr>
<tr>
<td>Comp#10</td>
<td>−13.04</td>
<td>−14.89</td>
</tr>
</tbody>
</table>

Notes: Listed are also the corresponding physiochemical descriptors calculated with QP simulations. aThe van der Waals surface area of the polar nitrogen and oxygen atoms; the accepted region is (7.0 to 200.0); bthe predicted octanol/water partition coefficient; the accepted region is (−2.0 to 6.5); cthe predicted aqueous solubility, where 5 (mol/dm³) is the concentration of the solute in a saturated solution that is in equilibrium with the crystalline solid; the accepted region is (−6.5 to 0.5); dpredicted apparent Caco-2 cell permeability in mm/second. Caco-2 cells are a model for the gut-blood barrier. QikProp predictions are for nonactive transport. The result of <25 is poor; epredicted percent of human oral absorption on a scale from 0% to 100%. The prediction is based on a quantitative multiple linear regression model. This property usually correlates well with human oral absorption. The result of <25 is poor; fComp#0 was an initial structure for core hopping that was designed with the intention to make it have stronger affinity than ZINC36728034. See the bottom left of Figure 2 for its structure.

Abbreviations: PPARα, peroxisome proliferator-activated receptor-alpha; PPARγ, peroxisome proliferator-activated receptor-gamma; ADME, absorption, distribution, metabolism, and excretion; PSA, polar surface area; QP, QikProp.

Table 1, Comp#1 has the strongest binding affinity with the two receptors PPARα and PPARγ, and hence it was singled out as the representative compound for further investigation.

Molecular dynamics trajectory analysis

To acquire the relevant information from a dynamic point of view, the 10 ns molecular dynamic simulations were carried out, respectively, for the crystal structures of PPARα (1k71), PPARγ (1k74), as well as their complexes with ragaglitazar and Comp#1: PPARα–ragaglitazar, PPARγ–ragaglitazar, PPARα–Comp#1, and PPARγ–Comp#1. As we can see from Figure 3, all of the characters concerned reached the simulation equilibrium within the 5 ns.

The RMSD from the initial conformation is a major criterion used to evaluate the stability of a protein system. Shown in the panels A and B of Figure 3 are the backbone RMSD curves for PPARα, PPARγ, PPARα–ragaglitazar, PPARγ–ragaglitazar, PPARα–Comp#1, and PPARγ–Comp#1. It is interesting to see that the RMSD values of PPARα–Comp#1 and PPARγ–Comp#1 are remarkably smaller than those of PPARα–ragaglitazar and PPARγ–ragaglitazar, indicating that the complexes of PPARα–Comp#1 and PPARγ–Comp#1 are more stable than PPARα–ragaglitazar and PPARγ–ragaglitazar.

In order to perform an in-depth study of the interactions of the AF2 helix with the agonist (Figure 1), the root mean square fluctuations for all the side-chain atoms of the receptors were also computed, as shown in panels C and D of Figure 3. It can be seen from the figure that the side-chain of the root mean square fluctuations of PPARα–Comp#1 or PPARγ–Comp#1 were very similar to that of PPARα–ragaglitazar or PPARγ–ragaglitazar, respectively. This is particularly true for the residues 459–465 of PPARα–Comp#1 and the residues 469–477 of PPARγ–Comp#1 (see the gray framed box marked with AF2 in panels C and D of Figure 3), indicating that the new designed Comp#1 is very likely to perform exactly the same function in activating the AF2 helix as the compound ragaglitazar did.

ADME prediction

Also listed in Table 1 are the corresponding ADME properties that are directly relevant to pharmaceutical potentials. It can be seen from the table that, compared with ragaglitazar, the predicted ADME properties, such as the polar surface area, logPo/w, logS, PPCaco, and percent of human oral absorption for the newly designed agonists are either much better than or quite close to that of ragaglitazar. Actually, these percentages are all within the acceptable ranges for humans, indicating that the ten compounds found in this study, particularly
The residues involved in forming the binding pocket of PPAR\(\alpha\) and PPAR\(\gamma\) for the ligand Comp\#1 are listed in Table 2, which are identical to the other nine compounds, where those marked with an asterisk are the key residues for forming the H-bonding network, as shown in Figure 1.

### Acknowledgments

The authors thank the three anonymous reviewers and the editor for their valuable comments, which were very helpful for strengthening the presentation of this study.

### Disclosure

The authors report no conflicts of interest in this work.

### References


27. Liu et al. Drug Design, Development and Therapy downloaded from https://www.dovepress.com/ by 54.70.40.11 on 08-Aug-2018


Drug Design, Development and Therapy is an international, peer-reviewed open-access journal that spans the spectrum of drug design and development through to clinical applications. Clinical outcomes, patient safety, and programs for the development and effective, safe, and sustained use of medicines are a feature of the journal, which has also been accepted for indexing on PubMed Central. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit http://www.dovepress.com/testimonials.php to read real quotes from published authors.

Submit your manuscript here: http://www.dovepress.com/drug-design-development-and-therapy-journal