J Comput Aided Mol Des (2014) 28:1143–1151
DOI 10.1007/s10822-014-9791-6

Identification of novel peroxisome proliferator-activated receptor-gamma (PPARγ) agonists using molecular modeling method

Veronica M. W. Gee · Fiona S. L. Wong · Lalitha Ramachandran · Gautam Sethi · Alan Prem Kumar · Chun Wei Yap

Received: 28 March 2014 / Accepted: 23 August 2014 / Published online: 29 August 2014

Abstract Peroxisome proliferator-activated receptor-gamma (PPARγ) plays a critical role in lipid and glucose homeostasis. It is the target of many drug discovery studies, because of its role in various disease states including diabetes and cancer. Thiazolidinediones, a synthetic class of agents that work by activation of PPARγ, have been used extensively as insulin-sensitizers for the management of type 2 diabetes. In this study, a combination of QSAR and docking methods were utilised to perform virtual screening of more than 25 million compounds in the ZINC library. The QSAR model was developed using 1,517 compounds and it identified 42,378 potential PPARγ agonists from the ZINC library, and 10,000 of these were selected for docking with PPARγ based on their diversity. Several steps were used to refine the docking results, and finally 30 potentially highly active ligands were identified. Four compounds were subsequently tested for their in vitro activity, and one compound was found to have a K_i values of <5 μM.

Keywords PPARγ · QSAR · Docking

Introduction

Millions of people worldwide are diagnosed with non-insulin dependent type 2 diabetes mellitus (NIDDM), a progressive metabolic disorder characterised by elevated blood glucose levels. Often the cause is a combination of peripheral insulin resistance and impaired compensatory insulin secretion [1]. Untreated or poorly controlled diabetes can lead to numerous long-term complications, which are largely responsible for morbidity and mortality associated with NIDDM. These

Veronica M. W. Gee and Fiona S. L. Wong have contributed equally to this study.

Electronic supplementary material The online version of this article (doi:10.1007/s10822-014-9791-6) contains supplementary material, which is available to authorized users.

V. M. W. Gee · C. W. Yap (✉)
Department of Pharmacy, Faculty of Science, National University of Singapore, Singapore, Singapore
e-mail: phayapc@nus.edu.sg

F. S. L. Wong · G. Sethi · A. P. Kumar (✉)
Cancer Science Institute of Singapore, National University of Singapore, Singapore, Singapore
e-mail: csiapk@nus.edu.sg

F. S. L. Wong
Department of Biological Sciences, Faculty of Science, National University of Singapore, Singapore, Singapore

L. Ramachandran · G. Sethi · A. P. Kumar
Department of Pharmacology, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore

A. P. Kumar
School of Biomedical Sciences, Faculty of Health Sciences, Curtin University, Bentley, WA, Australia

A. P. Kumar
Department of Biological Sciences, University of North Texas, Denton, TX, USA

A. P. Kumar
National University Cancer Institute, Singapore, Singapore

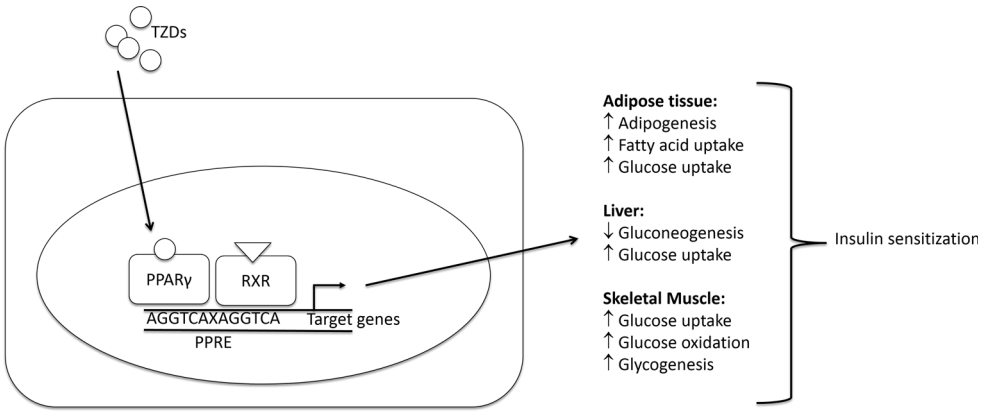


Fig. 1 A schematic diagram depicting the mechanism of action of thiazolidinedione (TZD). PPARγ, peroxisome proliferator-activated receptor-gamma; *RXR*, retinoid X receptor

complications include microvascular and macrovascular disorders such as neuropathy, nephropathy, retinopathy, stroke and cardiovascular diseases, as well as an increased risk for certain cancers [2, 3]. While diet and exercise regimes remain a cornerstone of diabetes management, the lack of patient adherence may necessitate the use of pharmacological interventions in order to achieve adequate glycemic control. Unfortunately, most commonly prescribed oral hypoglycemic agents, such as sulfonylurea and biguanides put the patient at risk for potentially fatal episodes of iatrogenic hypoglycaemia. The search for new drug candidates which are safer and more efficacious is therefore of current interest [1, 2, 4–16].

The earliest discovery of ciglitazone, a thiazolidinedione (TZD), sheds new light on the treatment of NIDDM. Unlike other classes of anti-diabetic agents, TZDs work as insulin sensitizers and do not induce hypoglycemic events. Figure 1 shows the mechanism of action of TZD. Currently, two TZD derivatives, rosiglitazone and pioglitazone have been approved for use in Europe [1, 17]. Mechanistic studies revealed that TZDs are potent agonists of the peroxisome proliferator-activated receptor-gamma (PPARγ), a nuclear receptor highly expressed in adipose tissues [18, 19]. Upon ligand-binding, PPARγ forms heterodimers with the retinoid X receptor (RXR). The complex subsequently binds the PPAR response element (PPREs) to regulate target gene expression involving both transactivation and transrepression [18, 20]. Several genes are activated by TZDs, many of which are involved in insulin actions, such as stimulating lipid and glucose uptake, increasing glucose oxidation and adipogenesis, and suppressing gluconeogenesis. There is growing evidence that PPARγ agonists also display anti-inflammatory activities, potentially reducing the risk of atherosclerotic cardiovascular diseases associated with NIDDM. While unsaturated fatty acids and eicosanoids are natural ligands of PPARγ, other synthetic non-TZD PPARγ ligands (i.e. tyrosine derivatives) have been identified in recent years [4, 21]. Given the pivotal role of PPARγ in lipid and glucose metabolism, it is not surprising that defects in PPARγ signalling (i.e. PPARγ mutations) have been implicated in the pathophysiology of NIDDM [3].

Despite promising clinical efficacy, TZDs are linked to some serious side- and off-target effects like edema, anemia, hepatotoxicity and cardiac hypertrophy [22]. It is however unclear whether toxicity is a result of PPARγ activation or innate to their chemical structure, and this warrants further investigation [1]. Meanwhile, other TZDs can be further explored to identify derivatives that do not present with serious side-effects. As the PPARγ's active site generally has a large cavity in which compounds bind, non-TZDs can also be studied to identify novel and potent agents that may have a more tolerable side-effect profile.

With increasing interest in the discovery of novel PPARγ ligands, there is a growing need for an expeditious and reliable tool that can facilitate the drug discovery process [23]. Recognising that the process of drug discovery and development are time and resource consuming, the use of computational power to accelerate and streamline the process of drug discovery is promising [24, 25]. The power of computer-aided drug design is expected to grow as technology advances and continues to evolve [25]. Computers can effectively use a vast amount of chemical and biological information about ligands or targets to identify and optimise new drugs, speeding up the process by up to 100 times of that of conventional high throughput screening [26].

There are many computational studies that have been conducted on the discovery of more potent, selective and safer PPARγ ligands, both TZDs derivatives and non-TZDs. However, a comparison of these studies' background and methodology reveals two problems:

1. Many QSAR studies on PPARγ were developed using only a small number of compounds. The number of compounds will affect the applicability domain of the model. If the training set is too small, the diversity of the data will be limited. Thus the models may not be sufficiently capable to predict potential compounds outside the chemical space. Similarly, a small test/validation set may not be sufficient to thoroughly validate the accuracy and generalizability of the models.

2. Majority of the studies were conducted using either ligand-based or structure-based approach only. Since both approaches have their own limitations, it will be useful to combine both approaches in a single study so that the two can complement each other.

This study was thus planned to address these shortcomings. The study used more than a thousand compounds to develop QSAR models which were then used to screen a large ZINC database containing several millions of compounds to identify potentially active compounds. These compounds were then further refined by docking them onto PPARγ to obtain scores associated to binding affinity. Four promising hits, ZINC08882146, ZINC08753365, ZINC13030391 and ZINC40251491, were subsequently validated in vitro for PPARγ binding activity by performing a competitive binding assay.

Materials and methods

Data set

A total of 210 compounds were collected from published literature of PPARγ-ligand studies for which experimental data were available [1, 2, 13, 18, 27–29]. Compounds were then classified as either an agonist or a non-agonist using a cut-off pK_i value of six. This threshold is higher than the commonly used value of five because our aim is to identify potent PPARγ agonists. The final agonist set consists of 177 compounds, with pK_i values ranging from 6.01 to 9.16. The final non-agonist set consists of 33 compounds, with pK_i values ranging from 3.72 to 5.98.

To increase the size of the non-agonist set, 1326 FDA-approved drugs were added [30]. In order to prevent false positives, drugs that are known to bind to PPARγ and drugs used in the treatment of diabetes were excluded. For example, glibenclamide, acarbose and metformin were omitted. The remaining 1,307 drugs were subsequently added to the non-agonist set. The resulting data set consists of a total of 1,340 non-agonists and 177 agonists and can be found in Table S1 of supplementary information. CORINA (http://www.molecular-networks.com/products/corina) was used to generate energetically minimised 3D conformations of the compounds.

Calculation of descriptors

Chemical descriptors for the 1,340 non-agonists and 177 agonists were calculated using Dragon Plus v5.5 (http://www.talete.mi.it/dragon.htm). The number of descriptors calculated for all the compounds was 1,648 and can be divided into 19 groups as shown in Table S2 of supplementary information.

QSAR model development

The dataset containing both agonists and non-agonists were split into training and validation sets in a 4:1 ratio using stratified random sampling. The training set was used to develop the models. The validation set was used to test the prediction power of the model obtained and was not used for developing a model.

For each Dragon descriptor group, a random subset of descriptors were first selected and forward selection was then used to optimize the descriptor subset. A Naive Bayes model was developed using the training set and the optimized descriptor subset. This process was repeated 15 times, with each cycle starting with the selection of a random subset of descriptors for optimisation. Fifteen models were developed for each of the 19 descriptor group and their pessimistic area under the receiver-operating characteristic curve (AUC) was calculated using fivefold cross-validation. The models for each descriptor group were ranked according to their AUC values. Models with the highest cross-validated pessimistic AUC for each descriptor groups were used subsequently to develop a simple majority vote consensus model if they possess a pessimistic AUC value that is >0.5. The consensus model was validated and its prediction power was determined by using the validation set. The software, RapidMiner 5.0.10, was used to develop the QSAR models.

Screening of ZINC library

The consensus model was then used to screen the ZINC library, which contained more than 25 million compounds, for potential novel PPARγ agonists [31]. 10,000 diverse compounds were then selected from the ZINC screening hits using the diverse subset routine in Molecular Operating Environment (MOE) (http://www.chemcomp.com/software.htm). Typed graph triangles fingerprint was calculated for all the predicted actives and Tanimoto coefficient was used to determine structural similarity. This was carried out for two reasons: (1) as some compounds in the ZINC library may be related to each other through classes, a diverse subset was obtained to ensure that compounds of similar structures or classes were not selected; (2) there is a practical limit on the number of compounds that can be confirmed using the subsequent molecular docking method.

Docking template selection

2Q59 from Protein Data Bank [32] was chosen as the docking template as it has a relatively good resolution and R factor of 2.2Å and 0.185 respectively. The downloaded complex was first processed by removing solvent

molecules and omitting redundant chains. Since the positions of hydrogen atoms are not included in the PDB files, these were added using MOE, and energy-minimized with MMFF94x force field.

Optimisation of docking parameters

Different docking parameters were expected to give different docking scores for a compound and the best set will be the one giving a score that relates most to the pKi values of the compound. In order to identify the optimal docking parameters, 100 agonists were randomly selected from Rucker et al. [13], which contributed the majority of the agonists in our dataset, to form the docking training set. This set of known agonists was docked using MOE and the calculated binding affinities, expressed as S scores, were determined. The docking parameters were systematically adjusted to determine the set of parameters that gives the best correlation between pK_i and S scores. The optimised parameters were then used to dock the 10,000 compounds to identify potential active hits.

Refining potential PPARγ agonists

The docked ZINC compounds were passed through a series of filters to select those compounds that are most likely to be active for biological testing. The first filtering process involved removing compounds with a docking S value that is larger than that of the agonists in the docking training set. The second filter involved the selection of compounds that have interactions with residues reported to be important for activity as well as with other residues found to be essential from the dockings performed on the docking training set. The third filtering process involved selecting those compounds that have the highest number of polar interactions (hydrogen bonds or ionic interactions) with the binding residues.

PPARγ competitive binding assay

The LanthaScreen® time-resolved fluorescence resonance energy transfer (TR-FRET) PPARγ competitive binding assay (Invitrogen) was performed according to the manufacturer's protocol in a white 384-well polypropylene assay plate. Briefly, the test compounds were purchased from MolPort Ltd (Riga, Latvia), dissolved in DMSO and incubated for 1 h with human PPAR-LBD tagged with GST, terbium-tagged anti-GST antibody and fluorescently labeled pan-PPARγ ligand (Fluormone™ Pan-PPAR Green). The pan-PPARγ ligand serves as a tracer, and it is in close spatial proximity with the terbium-tagged antibody upon binding to the hPPARγ-LBD. Excitation of terbium at 340 nm results in FRET and partial excitation of the tracer, with emissions at 520 nm. In the presence of competing ligands, the tracer is displaced, leading to a loss in TR-FRET signal. Radiometric emissions at 520 nm were normalized against terbium emissions at 495 nm and subsequently plotted against the indicated concentrations of test compounds to assess their PPARy binding ability. The sigmoidal dose–response curve was obtained using the Prism software (GraphPad Software, Inc., San Diego) and K_i values were calculated with the Cheng-Prusoff equation: $K_i = IC_{50}/(1 + [Tracer]/K_d)$, fixing Ymin values at <25 % (average Ymin values obtained for ZINC40251491 and Pioglitazone) for all compounds [33]. Positive and negative controls were performed under the same assay condition using pioglitazone and DMSO (1 %) respectively. All measurements were performed with the Tecan Ultra plate reader (Tecan, Durham, NC, USA).

Results

QSAR model

For each descriptor groups, the performance of the top ranking models selected is shown in Table S2 of supplementary information. The consensus model comprised of 19 models as all the models with the highest cross-validated pessimistic AUC for each descriptor groups had a pessimistic AUC value that is >0.5. The number of descriptors used in each of the 19 models ranged from 2 to 19. The consensus model had a pessimistic AUC of 0.993, sensitivity of 95.7 % and specificity of 96.9 % on the training set. Its corresponding performance on the validation set are 0.963, 80.0 and 91.8 % respectively.

The screening of the ZINC library using the validated consensus model identified 42,378 potential PPARγ agonists. 10,000 compounds out of these hits were selected as the diverse subset for molecular docking.

Docking

The docking training set was docked using various parameters and the S scores were compared with the corresponding pK_i values to calculate the correlation coefficient. Only 55 compounds were docked successfully and the set of docking parameter that gave the most favorable correlation coefficient of −0.44 between pK_i and S score is (Ligand: Rotate bonds; Placement: Triangle matcher; Rescoring 1: London dG; Refinement: Forcefield with side chains free; Rescoring 2: London dG). The S scores for the docking training set ranged from −11.50 to −17.22 kcal/mol. Self-docking of the compound MRL20 in the 2Q59 crystal structure using the optimized docking parameters gave a RMSD of 1.260 Å and a S score of −16.68 kcal/mol. The

docking operation performed on the 10,000 diverse compounds using the selected parameters produces S scores ranging from −4.55 to −20.76 kcal/mol.

Refining potential PPARγ agonists

The first filter removed compounds with docking S value larger than −11.5 kcal/mol since these compounds had poorer S scores compared to known agonists and thus are expected to be weaker agonists. This retained 8,076 compounds.

For the second filter, compounds which did not interact with key residues were removed. The key residues were determined by analysing the interaction profiles of the docking training set and 14 PPARγ-agonists complexes with x-ray crystallographic structures. It has been reported previously that five key residues (His 323, His 449, Cys 285, Tyr 473 and Ser 289) are essential for favorable binding with PPARγ agonists [14, 34]. Inspection of the crystal structures of 14 PPARγ-agonists complexes with EC_{50} from 4 to 570 nM verified the relevance of these residues. Analysis of the interaction profiles of the 55 compounds in the docking training set that were docked successfully found two key residues in common with the literature residues (Cys 285 and Ser 289), and two additional key residues (Arg 288 and Ile 341). Based on these results, only compounds that interacted with all four of these residues were retained, which produced 61 compounds.

The last filter sorts the 61 potential PPARγ agonists into two groups based on their potential activity relative to each other. Compounds having more polar interactions (hydrogen bonds or ionic interactions) with the binding residues were considered to be relatively more stable and potentially more active than those having only one such interaction. This is supported by the inspection of the 14 PPARγ-agonists complexes which showed that they have multiple polar interactions. These binding residues can include the five reported key residues as well as the two key residues found to be essential from our docking studies. Ligands with no such interactions at all were removed. The final list of 52 compounds with 30 potentially highly active compounds is shown in Table 1. Unfortunately, the top five compounds were not available for purchase. So after consideration of the availability of the compounds and budgetary constraints, four compounds, bolded in Table 1 and shown in Table 2, were selected for PPARγ ligand binding assay.

Table 1 Classification and ranking of the final list of ligands based on their interaction profile and their S score (in ascending order)

Potentially highly active (having two or more hydrogen and ionic interactions)	Potentially moderately active (having one hydrogen or ionic interaction)
ZINC33761955	ZINC40053156
ZINC12743866	ZINC23541835
ZINC33529285	ZINC36219340
ZINC09617094	ZINC16176464
ZINC40026159	ZINC22836695
ZINC35421105	ZINC40914814
ZINC29758787	ZINC03614699
ZINC36219754	ZINC35712099
ZINC40109114	ZINC31175348
ZINC40251491	ZINC14298747
ZINC18064587	ZINC09775773
ZINC40251726	ZINC27653456
ZINC08882146	ZINC11159196
ZINC40252210	ZINC39750909
ZINC27067575	ZINC14440896
ZINC15184616	ZINC36219267
ZINC13030391	ZINC35836164
ZINC21833194	ZINC24763348
ZINC36219582	ZINC35349121
ZINC12990264	ZINC09010722
ZINC26990849	ZINC36015087
ZINC22069813	ZINC27169769
ZINC35836269	
ZINC38362986	
ZINC10691448	
ZINC21900959	
ZINC26716415	
ZINC08753365	
ZINC16917459	
ZINC23599509	

In bold are the 4 compounds selected for PPARγ ligand binding assay

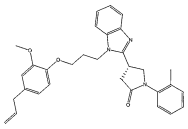
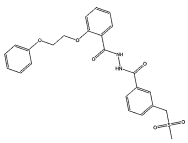
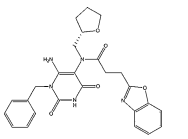
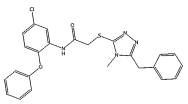
PPARγ ligand binding assay

To validate the predicted hits from the virtual screening approach, we performed a TR-FRET competitive binding assay to assess the ability of the test compounds to bind to purified human PPARγ LBD. ZINC08882146 (1), ZINC08753365 (2) and ZINC40251491 (4) bound to hPPARγ LBD in vitro (Fig. 2), whereas ZINC13030391 (3) did not bind the receptor at concentrations up to 400μM (Fig. 2). The percentage ligand displacement at the 10 μM concentrations for the compounds are 1.1, 20.8, 0, 33.0 % respectively and the K_i value for ZINC40251491 is 4.87 μM.

Discussion

In the present study, the potential efficiency of 2D QSAR modelling in the identification of active PPARγ agonists was demonstrated. The QSAR models were developed using 177 PPARγ agonists as well as using weak PPARγ

Table 2 Structures of the four compounds selected for PPARγ ligand binding assay

Compound	S score (kcal/mol)	Top concentration tested (μM)	Ligand displacement at top concentrations tested (%)	Estimated in vitro K_i (μM)	Structure
ZINC40251491	−13.70	1,000	74.94	4.87	
ZINC08882146	−13.38	400	64.69	131.63	
ZINC13030391	−12.91	400	4.80	>100	
ZINC08753365	−11.92	50	32.05	14.26	

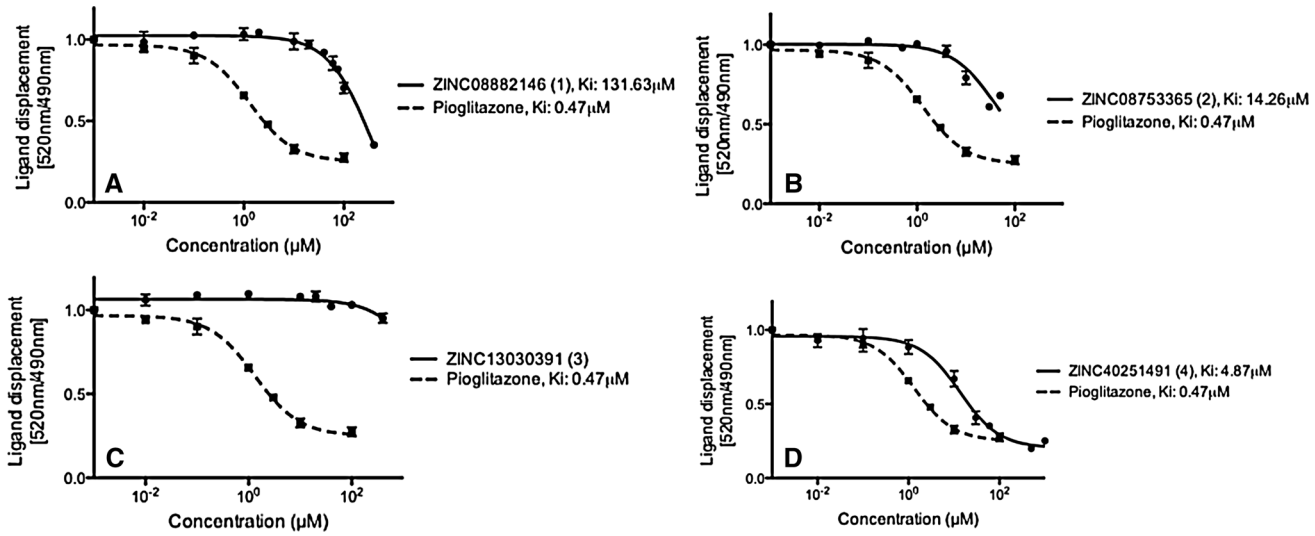


Fig. 2 PPARγ receptor binding of test compounds **a** ZINC08882146 (1), **b** ZINC08753365 (2), **c** ZINC13030391 (3) and **d** ZINC40251491 (4). Serial dilutions of the compounds were performed in DMSO and subsequently incubated for 1 hour with hPPARγ-LBD tagged with GST, terbium-tagged anti-GST antibody and fluorescently labelled pan-PPARγ ligand. The ability of the test compounds to bind to the PPARγ-LBD and thus displace the pan-PPARγ ligand was estimated from the decrease in the 520 nm/490 nm emission ratio upon excitation at 340 nm. The data shown are mean (SEM) from three independent experiments performed in duplicates. *Ki* values were calculated with the Cheng-Prusoff equation: $Ki = IC_{50}/(1 + [\text{Tracer}]/Kd)$, fixing Ymin values at <25 % (average Ymin values obtained for ZINC40251491 and Pioglitazone) for all compounds

agonists and approved drugs not known to bind to PPARγ as non-agonists. Although the total number of compounds used to develop the QSAR models was larger than in previous QSAR studies, the number of PPARγ agonists was still relatively small. This created an unbalanced dataset and caused the QSAR model to be biased towards the prediction of non-agonists. This is actually desirable as it had been shown that such unbalanced dataset could reduce the false positive rate [35]. This is why we added approved drugs not known to bind to PPARγ to the non-agonists pool. One concern with the use of approved drugs which are structurally different from the PPARγ agonists is that the QSAR models may learn to differentiate between structurally similar and structural dissimilar compounds

instead of learning the active difference between actives and inactives. The use of weak PPARγ agonists in this study is to prevent the development of such QSAR models. From the results of the consensus model on the validation set, it appears that this strategy of including weak PPARγ agonists is working as intended. There are six false positives prediction for the validation set. Four of the compounds are weak PPARγ agonists, which suggests that the models are not blindly separating structurally similar and structurally dissimilar compounds.

In QSAR modeling, the selection of relevant descriptors from a large pool of descriptors could result in over training, especially if the training set is relatively small compared to the pool of descriptors. In this study, we tried to reduce this risk of over training by using fivefold cross-validation to measure the performance of the models.

Simple majority vote was used for the consensus model instead of weighting based on cross-validated pessimistic AUC because the AUCs of the models in the consensus model are similar. Thus any weighting influence will be small and for practical reasons, they can all be assumed to have equal weights.

The screening of ZINC library using the consensus model identified a relatively large number of hits, which may contain a number of false positives. False positives may arise because the domain of applicability for the model was not defined. Thus some compounds in the ZINC library may be out of the PPARγ-ligands domain of the model and hence wrongly predicted.

In view of this problem, a docking study was used to further narrow down the hits to a manageable list of potentially potent, novel compounds by removing false positives. By using the filtering criteria based on two different sources of interaction profile (PDB and literature's agonists), ligands selected via the criteria were more likely to have good activity. There are some limitations of the docking study. Firstly, it is to be noted that the MOE diverse subset selection, which is used to reduce the number of compounds for docking, have a number of known limitations. The selection method tends to select outliers and the results are dependent on the order of the compounds in the database. Hence, some potential actives which are more interesting may be excluded by the use of this selection methods. Future studies should explore the use of other subset selection methods. The second limitation is that only 55 out of the 100 docking training set were docked successfully. A likely reason for the unsuccessful docking of the other 45 compounds is that the crystal structure of PPARγ in 2Q59 is induced-fitted to the ligand that is present inside its binding site. Hence, compounds which are structurally very different from that ligand may have difficulties docking into the binding site. A possible way to overcome this problem is to use PPARγ from multiple crystal structures with different ligands for docking. An implication of the unsuccessful docking of 45 compounds is that the docking study may miss out on potential hits.

It was unfortunate that the top five potentially highly active compounds were not commercially available despite their inclusion in the ZINC library. In the end, we only managed to purchase and test the in vitro activity of four compounds further down the potentially highly active list using PPARγ competitive binding assay. Although one of the compound had an in vitro K_i *value* of <5μM, it is important to note that the in vitro assay only shows that the compound can bind to PPARγ. It does not imply that the compound is a functional PPARγ agonist. Additional in vivo tests are required to verify that the compound is a PPARγ agonist.

In this study, QSAR and molecular docking were combined to identify a list of potentially active PPARγ agonists. Although we had a reasonably good rate of success (three active compounds out of four screened and one had in vitro K_i *value* of <5μM), it may be possible to further refine the computational methods to improve prioritising compounds for in vitro screening. We re-analyzed the docked structures of the four compounds in order to further understand the link between binding interactions and in vitro binding activity. The binding interactions of the four compounds are given in Fig. 3. Comparing ZINC40251491 and ZINC13030391, interactions with Leu 228, Gly 284, Gln 286, Tyr 327, Ser 342 and Glu 343 were identified to be present in the binder, ZINC40251491 but not in the non-binder, ZINC13030391. Among these six residues, Gln 286 seemed to be associated with binding activity as all the 14 PPARγ-agonists complexes with x-ray crystallographic structures also interact with this residue. One possible explanation for the importance of this residue is that it is closely connected to the key residues Cys 285, Arg 288 and Ser 289. Hence a compound that interacts with Gln 286 will help bring these three key residues closer to the compound for interaction. We can also view this from another angle. A compound which interacts strongly with the three key residues will tend to be closer to Gln 286 also since it is connected to these three residues, and thus have interactions with it. Regardless of the mechanism, for drug discovery purpose, it appears that interaction with Gln 286 is also an important determinant for good binding activity. It is to be noted that ZINC08753365, the compound with moderate binding affinity, also interacts with Gln 286, whereas ZINC08882146, the compound with weak binding affinity, does not.

Comparing the number of polar interactions for the four compounds, it can be seen that there is no correlation between the number of polar interactions and binding activity of the compounds. This appears to be different

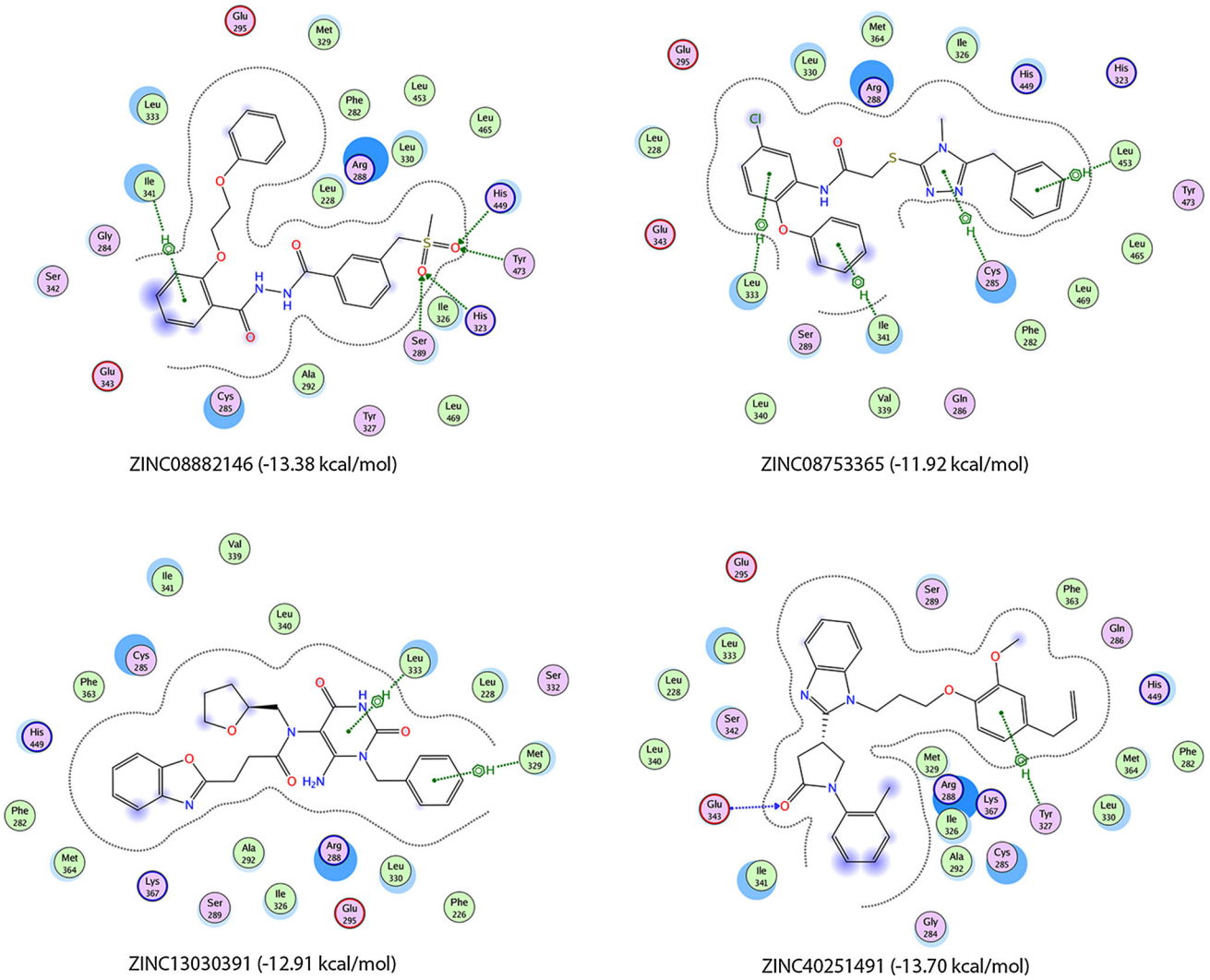


Fig. 3 Binding interaction profiles of the four selected compounds for in vitro assay

from the interactions seen in the 14 PPARγ-agonists complexes. One possible reason is that polar interactions are necessary but not the sole determinant of binding activity. In addition, there are limitations to the analysis of binding interactions using molecular docking. Molecular docking only shows a static picture of how a compound will interact with the active site of PPARγ. However, the interactions are not expected to be static in solution. One possible solution to overcome this limitation is to incorporate molecular dynamics simulations into the workflow. Molecular dynamics can capture the dynamics of the interactions and thus may potentially be useful to shed more light on the true interactions between a compound and PPARγ. Molecular dynamics simulations could be done for the final list of 52 compounds and the results used to prioritize the compounds instead of relying solely on analysis of number of polar interactions.

Conclusion

In this study, we identified 52 compounds that are potentially active against PPARγ from a total of more than 25 million compounds in the ZINC library using a combination of QSAR and molecular docking methods. Four compounds were subsequently selected for in vitro assay and one compound had K_i values of <5μM. These results are encouraging and suggests that these compounds are potentially useful for developing novel classes of PPARγ ligands.

Acknowledgments This work was supported by the Ministry of Education Academic Research Fund Tier 1 grant [R-148-000-165-112] to CWY. APK is supported by grants from the Singapore Ministry of Education Tier 2 [MOE2012-T2-2-139], Academic Research Fund Tier 1 [R-184-000-228-112], Cancer Science Institute of Singapore, Experimental Therapeutics I Program [Grant R-713-

001-011-271], NUHS Bench-to-Beside-Product Grant [184-000-243-515] and John Nott Cancer Fellowship from Cancer Council, Western Australia. GS was supported by grants from National Medical Research Council of Singapore [R-184-000-211-213]. We would like to thank the two reviewers for their insightful and encouraging comments, which helped to improve the quality of our manuscript greatly.

References

1. Dixit A, Saxena AK (2008) Eur J Med Chem 43(1):73
2. Rathi L, Kashaw SK, Dixit A, Pandey G, Saxena AK (2004) Bioorg Med Chem 12(1):63
3. Reaven G (2005) Circulation 112(20):3030
4. Giaginis C, Theocharis S, Tsantili-Kakoulidou A (2008) Chem Biol Drug Des 72(4):257
5. Giaginis C, Theocharis S, Tsantili-Kakoulidou A (2009) Mini Rev Med Chem 9(9):1075
6. Huang HJ, Lee KJ, Yu HW, Chen HY, Tsai FJ, Chen CY (2010) J Biomol Struct Dyn 28(2):187
7. Iwata Y, Miyamoto S, Takamura M, Yanagisawa H, Kasuya A (2001) J Mol Graph Model 19(6):536
8. Khanna S, Sobhia ME, Bharatam PV (2005) J Med Chem 48(8):3015
9. Kumar PM, Hemalatha R, Mahajan SC, Karthikeyan C, Moorthy NS, Trivedi P (2008) Med Chem 4(3):273
10. Kurogi Y (1999) Drug Des Discov 16(2):109
11. Lather V, Kairys V, Fernandes MX (2009) Chem Biol Drug Des 73(4):428
12. Liao C, Xie A, Zhou J, Shi L, Li Z, Lu XP (2004) J Mol Model 10(3):165
13. Rucker C, Scarsi M, Meringer M (2006) Bioorg Med Chem 14(15):5178
14. Salam NK, Huang TH, Kota BP, Kim MS, Li Y, Hibbs DE (2008) Chem Biol Drug Des 71(1):57
15. Sundriyal S, Bharatam PV (2009) Eur J Med Chem 44(1):42
16. Sundriyal S, Viswanad B, Bharathy E, Ramarao P, Chakraborti AK, Bharatam PV (2008) Bioorg Med Chem Lett 18(11):3192
17. Issemann I, Green S (1990) Nature 347(6294):645
18. Vedani A, Descloux AV, Spreafico M, Ernst B (2007) Toxicol Lett 173(1):17
19. Barish GD, Narkar VA, Evans RM (2006) J Clin Invest 116(3):590
20. Brown JD, Plutzky J (2007) Circulation 115(4):518
21. Youssef J, Badr M (2004) J Biomed Biotechnol 2004(3):156
22. Simo R, Rodriguez A, Caveda E (2010) Curr Drug Saf 5(3):234
23. Stahl M, Guba W, Kansy M (2006) Drug Discov Today 11(7–8):326
24. Huang CM, Elmets CA, Tang DC, Li F, Yusuf N (2004) Genomics Proteomics Bioinform 2(3):143
25. Desvergne B, Wahli W (1999) Endocr Rev 20(5):649
26. Gasteiger J, Teckentrup A, Terfloth L, Spycher S (2003) J Phys Org Chem 16(4):232
27. Suh YG, Kim NJ, Koo BW, Lee KO, Moon SH, Shin DH, Jung JW, Paek SM, Chang DJ, Li F, Kang HJ, Le TV, Chae YN, Shin CY, Kim MK, Lim JI, Ryu JS, Park HJ (2008) J Med Chem 51(20):6318
28. Giaginis C, Theocharis S, Tsantili-Kakoulidou A (2007) J Chromatogr B Analyt Technol Biomed Life Sci 857(2):181
29. Scarsi M, Podvinec M, Roth A, Hug H, Kersten S, Albrecht H, Schwede T, Meyer UA, Rucker C (2007) Mol Pharmacol 71(2):398
30. Wishart DS (2008) Pharmacogenomics 9(8):1155
31. Irwin JJ, Shoichet BK (2005) J Chem Inf Model 45(1):177
32. Kirchmair J, Markt P, Distinto S, Schuster D, Spitzer GM, Liedl KR, Langer T, Wolber G (2008) J Med Chem 51(22):7021
33. Cheng Y, Prusoff WH (1973) Biochem Pharmacol 22(23):3099
34. Nolte RT, Wisely GB, Westin S, Cobb JE, Lambert MH, Kurokawa R, Rosenfeld MG, Willson TM, Glass CK, Milburn MV (1998) Nature 395(6698):137
35. Han LY, Ma XH, Lin HH, Jia J, Zhu F, Xue Y, Li ZR, Cao ZW, Ji ZL, Chen YZ (2008) J Mol Graph Model 26(8):1276