European Journal of Medicinal Chemistry xxx (2011) 1–17

Contents lists available at ScienceDirect

European Journal of Medicinal Chemistry

journal homepage: http://www.elsevier.com/locate/ejmech

Original article

Discovery of new nanomolar peroxisome proliferator-activated receptor γ activators via elaborate ligand-based modeling

Belal O. Al-Najjar [a], Habibah A. Wahab [a,b,**], Tengku Sifzizul Tengku Muhammad [b,c], Alexander Chong Shu-Chien [b,d], Nur Adelina Ahmad Noruddin [b], Mutasem O. Taha [e,*]

[a] Pharmaceutical Design and Simulation (PhDS) Laboratory, School of Pharmaceutical Sciences, Universiti Sains Malaysia, 11800 Minden, Pulau Pinang, Malaysia
[b] Malaysian Institute of Pharmaceuticals and Nutraceuticals, Ministry of Science, Technology and Innovation, SAINS@USM, No. 10, 11900 Persiaran Bukit Jambul, Pulau Pinang, Malaysia
[c] Department of Biological Sciences, Universiti Malaysia Terengganu, 21030 Kuala Terengganu, Terengganu, Malaysia
[d] School of Biological Sciences, Universiti Sains Malaysia, 11800, Penang, Malaysia
[e] Drug Discovery Unit, Department of Pharmaceutical Sciences, Faculty of Pharmacy, University of Jordan, Queen Rania Street, Amman, Jordan

ARTICLE INFO

Article history:
Received 4 January 2011
Received in revised form
12 March 2011
Accepted 16 March 2011
Available online xxx

Keywords:
Peroxisome Proliferator-activated receptor γ
Pharmacophore modeling
Quantitative structure–activity relationship
In silico screening
Type 2 diabetes mellitus

ABSTRACT

Peroxisome Proliferator-Activated Receptor γ (PPARγ) activators have drawn great recent attention in the clinical management of type 2 diabetes mellitus, prompting several attempts to discover and optimize new PPARγ activators. With this in mind, we explored the pharmacophoric space of PPARγ using seven diverse sets of activators. Subsequently, genetic algorithm and multiple linear regression analysis were employed to select an optimal combination of pharmacophoric models and 2D physicochemical descriptors capable of accessing self-consistent and predictive quantitative structure–activity relationship (QSAR) ($r^2_{71} = 0.80$, $F = 270.3$, $r^2_{LOO} = 0.73$, r^2_{PRESS} against 17 external test inhibitors $= 0.67$). Three orthogonal pharmacophores emerged in the QSAR equation and were validated by receiver operating characteristic (ROC) curves analysis. The models were then used to screen the national cancer institute (NCI) list of compounds. The highest-ranking hits were tested *in vitro*. The most potent hits illustrated EC_{50} values of 15 and 224 nM.

1. Introduction

Type II diabetes is a progressive disease characterized by insulin resistance in peripheral tissues and/or impaired insulin secretion by the pancreas. The resultant high blood glucose level generally leads to several serious complications. The World Health Organization recently warned that type II diabetes has become global pandemic [1]. Type II diabetes is strongly associated with obesity. The common link between obesity and type II diabetes is insulin resistance [1,2]. At the molecular level, the mechanism of insulin resistance in type II diabetes appears to involve defects in post-receptor signal transduction [3–5].

PPARs are type II nuclear hormone receptors that participate in the regulation of fatty acids, carbohydrates and glucose metabolism [6]. PPARs are ligand-dependent transcriptional regulators heterodimerized with retinoid X receptor. They bind to peroxisome proliferator-response element (PPRE) [7] and regulate the activities of nuclear factors essential in immunomodulatory and inflammatory responses [8,9]. Three PPAR isoforms have been identified: α, β (or δ), and γ [10].

PPARγ agonists have drawn great attention in the clinical management of cardiovascular risk factors associated with metabolic syndrome and type 2 diabetes [11]. In the nineties, several classes of PPARγ agonists were reported to have anti-diabetic actions, including: thiazolidindiones, dihydrobenzofurans, dihydrobenzopyrans, benzofurans benzoxazoles and α-amino-β-phenylpropanoic acid derivatives [12]. Recently, more diverse PPARγ agonists were developed as potential hypoglycemic agents, including: β-carboxyethyl-rhodanines [13], indene *N*-oxides [14], indoles [15], 7-azaindoles [16], 2-aryl-4-oxazolylmethoxy-benzylglycines and 2-aryl-4-thiazolyl-methoxy-benzylglycines [17], 5-arylthiazolidine-2,4-diones [18], aryloxazolidinediones [19] and phenylacetic acid derivatives [20].

Clearly, the main focus of recent efforts towards the development of new PPARγ agonists concentrate on structure-based ligand design [13,21–33] with few ligand-based exceptions [34–37]. To date, many PPARγ X-ray complexes are documented in the Protein Data Bank [38–47].

* Corresponding author. Tel.: +962 6 5355000x23305; fax: +962 6 5339649.
** Corresponding author. Tel.: +60164410188.
E-mail addresses: habibahwahab@yahoo.co.uk (H.A. Wahab), mutasem@ju.edu.jo (M.O. Taha).

0223-5234/$ – see front matter © 2011 Elsevier Masson SAS. All rights reserved.
doi:10.1016/j.ejmech.2011.03.040

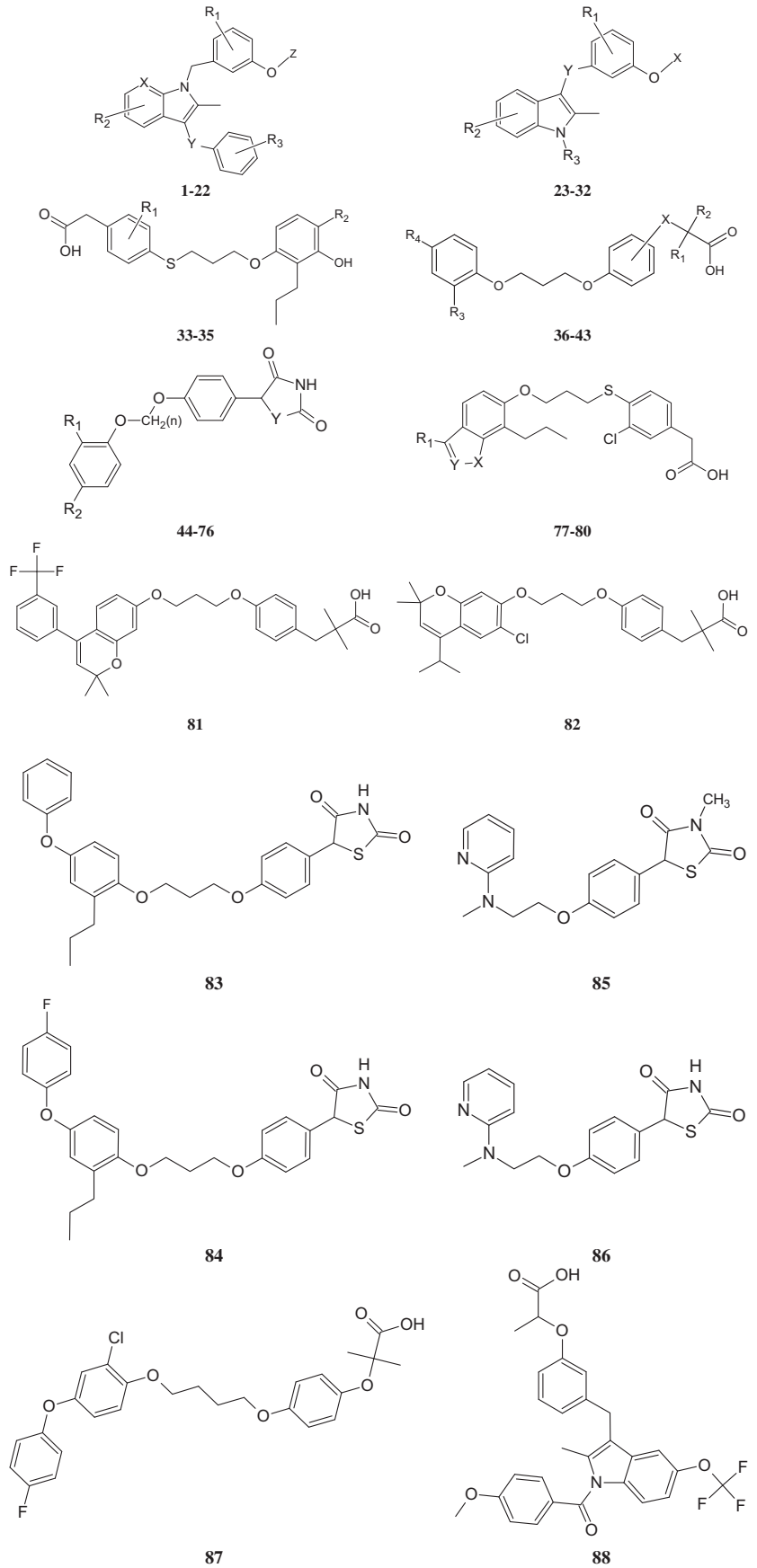

Fig. 1. The chemical scaffolds of training compounds, the detailed structures are as in Table A under Supplementary Materials.

Please cite this article in press as: B.O. Al-Najjar, et al., Discovery of new nanomolar peroxisome proliferator-activated receptor γ activators via elaborate ligand-based modeling, European Journal of Medicinal Chemistry (2011), doi:10.1016/j.ejmech.2011.03.040

Table 1
Performances of best representatives of clustered pharmacophore hypotheses generated for PPARγ activators.

Training set[a]	Run[b]	Hypotheses	Pharmacophoric features in generated hypotheses	Total cost	Cost of null hypothesis	Residual cost[c]	R^2[d]	*F*-Statistic[e]	Cat-scramble (%)
I	1	7	HBA, 2×Hbic, NegIon, RingArom	97.7	111.4	13.6	0.83	23.9	90
		10	HBA, 2×Hbic, NegIon, RingArom	98.1	111.4	13.2	0.83	65.4	90
	2	1	2×HBA, NegIons, RingArom	89.8	111.4	21.6	0.94	52.3	90
		4	2×HBA, 2×Hbic, NegIons	95.4	111.4	15.9	0.86	61.6	90
	3	**1**[f]	**HBA, 3×Hbic, NegIons**	**89.4**	**111.4**	**22**	**0.92**	**83.6**	**90**
		4	HBA, 2×Hbic, NegIons, RingArom	91	111.4	20.4	0.89	72.4	90
	4	8	HBA, 2×Hbic, NegIons, RingArom	96.5	111.4	14.9	0.84	49.7	90
		10	HBA, NegIons, 2×RingArom	97.1	111.4	14.3	0.83	49.2	90
	5	2	HBA, 3×Hbic, RingArom	91.2	111.4	20.1	0.93	60.8	90
		8[f]	**2×HBA, Hbic, RingArom**	**84.1**	**111.4**	**27.2**	**0.89**	**97.3**	**90**
	6	2	HBA, 3×Hbic, RingArom	100.5	111.4	10.9	0.8	43.2	90
		10	HBA, 3×Hbic, RingArom	82.8	111.4	28.5	0.77	35.1	90
	7	3	HBA, 3×Hbic, RingArom	94.3	111.4	17.1	0.90	58.8	90
		10	HBA, 3×Hbic, RingArom	97.2	111.4	14.2	0.86	32.5	90
	8	3	HBA, 3×Hbic, RingArom	92.6	111.4	18.8	0.901	63.6	90
		10	HBA, 3×Hbic, RingArom	94.9	111.4	16.5	0.88	63.8	90
II	1	2	2×HBA, 2×Hbic, NegIons	135	166.1	31.2	0.87	28.4	90
		8	2×HBA, 2×Hbic, NegIons	139.8	166.1	26.3	0.83	54.9	90
	2	1	2×HBA, 2×Hbic, NegIons	130.1	166.1	36	0.92	61.7	90
		3	3×HBA, NegIons	132.6	166.1	33.5	0.90	60.5	90
	3	1	2×HBA, 2×Hbic, NegIons	128.9	166.1	37.3	0.91	46.4	90
		9	2×HBA, 2×Hbic, NegIons	138.5	166.1	27.6	0.83	46.0	90
	4	2	2×HBA, 2×Hbic, NegIons	131.7	166.1	34.4	0.90	42.9	90
		5	3×HBA, NegIons	136.4	166.1	29.7	0.85	44.0	90
	5	7	2×HBA, 3×Hbic	140.8	166.1	25.4	0.84	59.0	90
		10	HBA, 3×Hbic, NegIons	141.2	166.1	25	0.84	76.1	90
	6	3	2×HBA, 2×Hbic, NegIons	138.8	166.1	27.4	0.85	32.8	90
		7	3×HBA, NegIons	139.6	166.1	26.5	0.84	30.2	90
	7	5	2×HBA, 2×Hbic, NegIons	140.8	166.1	25.3	0.82	43.9	90
		8	2×HBA, 3×Hbic	141.5	166.1	24.6	0.82	67.1	90
	8	4	3×HBA, Hbic	139	166.1	27.1	0.86	91.4	90
		6	3×HBA, Hbic	140.3	166.1	25.8	0.84	54.5	90
III	1	8	2×HBA, NegIons, RingArom	94.6	124.4	29.8	0.93	68.0	90
		10	HBA, 2×Hbic, NegIons, RingArom	94.8	124.4	29.6	0.93	72.4	90
	2	**2**[f]	**2×HBA, 2×Hbic, NegIons**	**90.7**	124.4	**33.7**	**0.97**	**75.1**	**90**
		7	HBA, 2×Hbic, NegIons, RingArom	93.0	124.4	31.4	0.94	62.9	90
	3	1	HBA, 2×Hbic, NegIons, RingArom	90.1	124.4	34.4	0.95	32.9	90
		9	HBA, 2×Hbic, NegIons, RingArom	91.7	124.4	32.8	0.95	57.5	90
	4	3	HBA, 2×Hbic, NegIons, RingArom	93.7	124.4	30.8	0.93	73.5	90
		6	2×HBA, NegIons, RingArom	94.3	124.4	30.4	0.92	50.1	90
	5	8	HBA, 2×Hbic, NegIons, RingArom	95.2	124.4	29.3	0.92	33.4	90
		10	2×HBA, 3Hbic	95.6	124.4	28.9	0.92	48.5	90
	6	2	HBA, 2×Hbic, NegIons, RingArom	95.9	124.4	28.6	0.93	19.3	90
		5	HBA, 2×Hbic, NegIons, RingArom	96.0	124.4	28.4	0.92	70.6	90
	7	6	HBA, 2×Hbic, NegIons, RingArom	91.7	124.4	32.8	0.97	56.6	90
		10	HBA, 2×Hbic, NegIons, RingArom	95.6	124.4	28.8	0.93	27.9	90
	8	3	HBA, 3×Hbic, NegIons	95.5	124.4	29.0	0.94	66.6	90
		4	HBA, 3×Hbic, NegIons	95.6	124.4	28.8	0.94	73.1	90
IV	1	2	HBA, 2×Hbic, NegIons, RingArom	87.5	113.2	25.7	0.93	60.7	90
		6	HBA, 3×Hbic, NegIons	89.7	113.2	23.5	0.91	55.9	90
	2	1	HBA, 2×Hbic, NegIons, RingArom	87.3	113.2	25.9	0.94	46.2	90
		10	HBA, 2×Hbic, NegIons, RingArom	94.4	113.2	18.8	0.86	19.8	90
	3	2	HBA, 2×Hbic, NegIons, RingArom	90.2	113.2	23.0	0.92	54.2	90
		8	HBA, 2×Hbic, NegIons, RingArom	91.2	113.2	22.0	0.91	63.1	90
	4	3	2×HBA, 2×Hbic, NegIons	93.3	113.2	19.9	0.89	40.0	90
		9	HBA, 3×Hbic, Negions	95.9	113.2	17.3	0.86	41.9	90
	5	3	HBA, 2×Hbic, NegIons, RingArom	92.5	113.2	20.8	0.92	83.1	90
		4	HBA, 4xHbic	92.6	113.2	20.6	0.92	42.9	90
	6	5	HBA, 2×Hbic, NegIons, RingArom	92.1	113.2	21.1	0.91	28.8	90
		7	HBA, 3×Hbic, RingArom	92.6	113.2	20.6	0.90	53.3	90
	7	1	HBA, 3×Hbic, RingArom	92.1	113.2	21.2	0.92	38.5	90
		7	HBA, Hbic, 2×RingArom	94.9	113.2	18.3	0.91	36.6	90
	8	2	HBA, 3×Hbic, RingArom	92.1	113.2	21.1	0.75	82.0	90
		6	2×HBA, 3×Hbic	95.2	113.2	18.0	0.90	35.7	90
V	1	5	2×Hbic, NegIons, 2×RingArom	77.8	104.6	26.8	0.97	66.9	90
		7	HBA, 3×Hbic, RingArom	78.2	104.6	26.4	0.97	57.4	90
	2	3	HBA, 4xHydrophobic	78	104.6	26.6	0.96	37.6	90
		9	3×Hbic, 2×RingArom	79.4	104.6	25.3	0.94	33.3	90
	3	2	2×Hbic, NegIons, 2×RingArom	76.7	104.6	27.9	0.97	55.0	90
		10	2×Hbic, NegIons, 2×RingArom	78.1	104.6	26.5	0.96	64.6	90

(*continued on next page*)

Please cite this article in press as: B.O. Al-Najjar, et al., Discovery of new nanomolar peroxisome proliferator-activated receptor γ activators via elaborate ligand-based modeling, European Journal of Medicinal Chemistry (2011), doi:10.1016/j.ejmech.2011.03.040

Table 1 (*continued*)

Training set[a]	Run[b]	Hypotheses	Pharmacophoric features in generated hypotheses	Total cost	Cost of null hypothesis	Residual cost[c]	R^2[d]	*F*-Statistic[e]	Cat-scramble (%)
	4	4	HBA, 3×Hbic, RingArom	77.9	104.6	26.7	0.97	88.4	90
		8	4xHbic, NegIons	78	104.6	26.6	0.97	42.2	90
	5	2	3Hbic, NegIons, RingArom	75.9	104.6	28.8	0.98	13.3	90
		10	2×Hbic, NegIons, 2×RingArom	76.7	104.6	27.9	0.96	61.2	90
	6	2	3×Hbic, NegIons, RingArom	75.9	104.6	28.7	0.97	46.6	90
		4	HBA, 2×Hbic, NegIons, RingArom	76.6	104.6	28	0.96	73.3	90
	7	5	3×Hbic, NegIons, RingArom	77.1	104.6	27.5	0.96	18.1	90
		7	3Hbic, NegIons, RingArom	77.3	104.6	27.3	0.95	25.2	90
	8	6	HBA, 2×Hbic, NegIons, RingArom	76	104.6	28.7	0.96	44.2	90
		9	2×Hbic, NegIons, 2×RingArom	76.3	104.6	28.3	0.96	51.1	90
VI	1	1	2×HBA, Hbic, RingArom	95.9	101	5.1	0.91	10.8	90
		8	HBA, 2×Hbic, RingArom	99.6	101	1.4	0.83	72.7	90
	4	6	2×HBA, Hbic, RingArom	100	101	1	0.83	80.5	90
		10	HBA, Hbic, 2×RingArom	101.1	101	0.1	0.81	6.2	90
	6	1	HBA, HBD, 2×Hbic	99.5	101	1.5	0.87	0.2	90
		10	HBA, HBD, Hbic, RingArom	98.5	101	2.5	0.76	0.2	90
	8	2	2HBA, HBAD, Hbic	97.1	101	3.9	0.84	2.1	90
		10	2×HBA, HBD, Hbic	99.8	101	1.2	0.82	2.5	90
VII	1	2	HBA, 3×Hbic, RingArom	84.5	121.3	36.7	0.96	0.0	90
		4	2×HBA, 3×Hbic	79	121.3	42.3	0.95	0.6	90
	2	3	HBA, Hbic, 2×RingArom	89.5	121.3	31.7	0.93	2.5	90
		6	2×HBA, 2×Hbic	91.7	121.3	29.6	0.91	1.2	90
	3	2	HBA, 2×Hbic, RingArom	90.1	121.3	31.1	0.94	1.0	90
		4	HBA, 3Hbic	91	121.3	30.2	0.93	3.1	90
	4	1	2×HBA, 3×Hbic	87.3	121.3	33.9	0.94	8.1	90
		6	2×HBA, 3×Hbic	94	121.3	27.3	0.88	6.9	90
	5	3	2×HBA, 3×Hbic	88.6	121.3	32.6	0.88	0.4	90
		10	2×HBA, 3×Hbic	95.5	121.3	25.7	0.82	1.8	90
	6	2	2×HBA, HBD, Hbic	92.9	121.3	28.4	0.89	0.0	90
		6	2×HBA, HBD, Hbic	95.3	121.3	25.9	0.87	2.7	90
	7	1	2×HBA, HBD, Hbic	90.1	121.3	31.1	0.92	1.4	90
		5	HBA, HBD, 3×Hbic	93.4	121.3	27.9	0.88	4.4	90
	8	4	HBA, HBD, 3xHbic	88.4	121.3	32.8	0.91	3.7	90
		9	HBA, HBD, 3×Hbic	89.8	121.3	31.5	0.90	3.9	90

[a] Correspond to training sets in Table B under Supplementary Materials.
[b] Correspond to runs in Table C under Supplementary Materials.
[c] Difference between total cost and the cost of the corresponding null hypotheses.
[d] Correlation coefficients between pharmacophore-based bioactivity estimates and bioactivities of corresponding training compounds (subsets in Table B under Supplementary Material).
[e] Fisher statistic calculated based on the linear regression between the fit values of all collected inhibitors (**1**–**88**, Table A under Supplementary Material) against pharmacophore hypothesis (employing the "best fit" option and Eq. (D) under Supplementary Material) and their respective PPARγ activators bioactivities (log ($1/IC_{50}$) values).
[f] Bolded pharmacophores appeared in the best QSAR equations.

However, although considered the most reliable structural information that can be used for drug design, crystallographic structures are limited by inadequate resolution [48] and crystallization-related artifacts of the ligand–protein complex [49–51]. Moreover, crystallographic structures generally ignore structural heterogeneity related to protein anisotropic motion and discrete conformational substrates particularly in cases of pronounced induced-fit protein flexibilities [52].

The continued interest in designing new PPARγ agonits, combined with the drawbacks of structure-based design and the significant induced-fit flexibility observed for PPARγ [53], prompted us to explore the possibility of developing ligand-based three-dimensional (3D) pharmacophore(s) integrated within self-consistent QSAR model(s). This approach avoids the pitfalls of structure-based techniques; furthermore, the pharmacophore model(s) can be used as 3D templates to synthesize new PPARγ agonists scaffolds.

We previously reported the use of this approach towards the discovery of new inhibitory leads against glycogen synthase kinase 3β [54], hormone sensitive lipase [55], bacterial MurF [56], protein tyrosine phosphatase 1B [57], influenza neuraminidase [58], estrogen receptor beta ligands [59], cholesteryl ester transfer protein inhibitors [60] and β-secretase inhibitors [61], CDK1 inhibitors [62], and Heat Shock Protein 90α Inhibitors [63].

We employed the CATALYST-HYPOGEN module embedded in Discovery Studio (version 2.5) [64] to construct numerous reasonable binding hypotheses for PPARγ agonists. Subsequently, genetic function algorithm (GFA) and multiple linear regression (MLR) analyses were employed to search for an optimal QSAR that combine high-quality binding pharmacophores with other molecular descriptors and capable of explaining bioactivity variation across a collection of diverse PPARγ agonists. Optimal pharmacophores were validated by evaluating their abilities to successfully classify a list of compounds as actives or inactives by assessing their receiver operating characteristic (ROC) curves.

2. Results and discussion

CATALYST-HYPOGEN models drug–receptor interactions using information derived only from the drug structure [54–63]. It requires a collection of training molecules with bioactivities ranging from 3 to 4 orders of magnitude to attempt explain bioactivity variation with respect to geometric localization of chemical features in training molecules. To do this, it identifies a 3D array of a maximum of five chemical features common to active training molecules, which provides a relative alignment for each input molecule consistent with their binding to a proposed common receptor site. The chemical features considered can be hydrogen bond donors and acceptors (HBD and HBA), aliphatic and aromatic hydrophobes (Hbic), positive and negative ionizable (PosIon and NegIon) groups and aromatic planes (RingArom). The

Please cite this article in press as: B.O. Al-Najjar, et al., Discovery of new nanomolar peroxisome proliferator-activated receptor γ activators via elaborate ligand-based modeling, European Journal of Medicinal Chemistry (2011), doi:10.1016/j.ejmech.2011.03.040

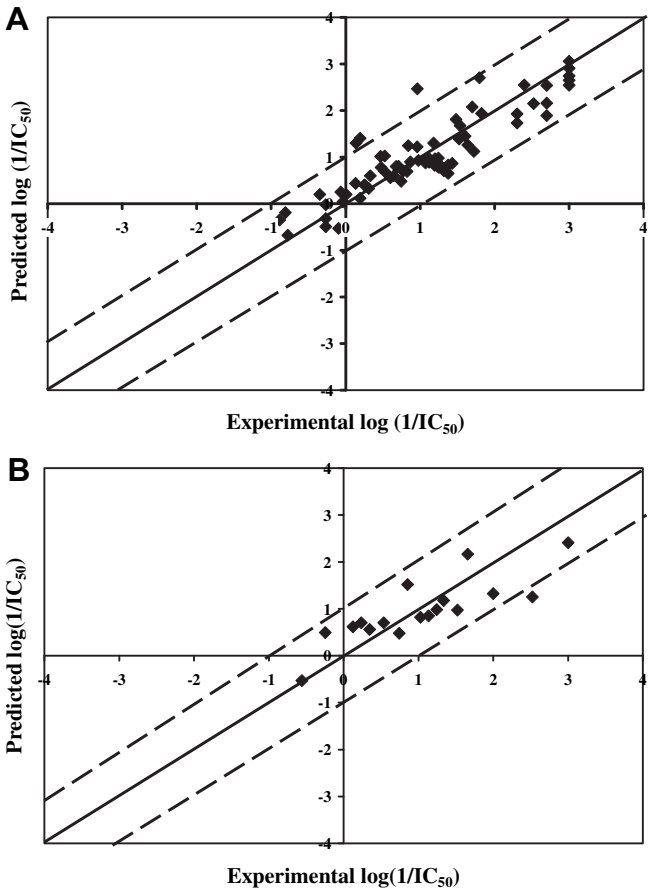

Fig. 2. Experimental versus (A) fitted (71 compounds, $r_{LOO}^2 = 0.73$), and (B) predicted (17 compounds, $r_{PRESS}^2 = 0.67$) bioactivities calculated from the best QSAR model Eq. (1). The solid lines are the regression lines for the fitted and predicted bioactivities of training and test compounds, respectively, whereas the dotted lines indicate 1.0 log point error margins.

conformational flexibility of training ligands is modeled by creating multiple conformers, judiciously prepared to emphasize representative coverage over a specified energy range. CATALYST pharmacophores have been used as 3D queries for mining structural databases for new active leads [54–63,65].

In the present project, we generated diverse hypotheses for a series of PPARγ activators. A total of 88 compounds were used in this study (Fig. 1 and Table A under Supplementary Material) [16,18–20,66,67]. Seven training subsets were selected from the collection (Table B under Supplementary Material). Each subset consisted of inhibitors of wide structural diversity.

2.1. Exploration of PPARγ pharmacophoric space

The literature was surveyed to collect as many reported structurally diverse PPARγ activators as possible (**1**–**88**, see Fig. 1 and Table A under Supplementary Material) [16,18–20,66,67]. The 2D structures of the agonists were imported into Discovery Studio 2.5 and converted automatically into 3D single conformer representations. The structures were used as starting points for conformational analyses and in the determination of various molecular descriptors for QSAR modeling.

The conformational space of each agonist was extensively sampled utilizing the poling algorithm employed within Discovery Studio 2.5 [64]. Conformational coverage was performed employing the "Best" module to ensure extensive sampling of conformational space.

Subsequently, CATALYST-HYPOGEN module in Discovery Studio was employed to identify as many pharmacophoric binding modes assumed by PPARγ agonists as possible. HYPOGEN implements an optimization algorithm that evaluates large number of potential binding models for a particular target through fine perturbations to hypotheses that survived the constructive and subtractive phases of the modeling algorithm (see Section 4.1.4 Pharmacophoric Hypotheses Generation in Experimental and Supplementary Materials) [68]. The extent of the evaluated pharmacophoric space is reflected by the configuration (Config.) cost calculated for each modeling run. It is generally recommended that the Config. cost of any HYPOGEN run not to exceed 17 (corresponding to 2 [48] hypotheses to be assessed by CATALYST) to guarantee thorough analysis of all models [69]. The size of the investigated pharmacophoric space is a function of training compounds, selected input chemical features and other CATALYST control parameters [69].

Restricting the size of explored pharmacophoric space should improve the efficiency of optimization via allowing efficient assessment of limited number of pharmacophoric models. On the other hand, extreme restrictions imposed on the evaluated pharmacophoric space might reduce the possibility of discovering optimal binding hypotheses, as they might occur outside the "boundaries" of the evaluated space.

Therefore, it was decided to explore the pharmacophoric space of PPARγ agonists under reasonably imposed "boundaries" through 56 HYPOGEN automatic runs and employing seven carefully selected training subsets: subsets 1–7 in Table B under Supplementary Material. The training compounds in these subsets were selected in such away to guarantee maximal 3D diversity and continuous bioactivity spread over more than 3.5 logarithmic cycles. We gave special emphasis to the 3D diversity of the most active compounds in each training subset (Table B under Supplementary Material) because of their significant influence on the extent of the evaluated pharmacophoric space during the constructive phase of HYPOGEN algorithm (see Section 4.1.4 Pharmacophoric Hypotheses Generation in Experimental and Supplementary Materials) [68].

Guided by our rationally restricted pharmacophoric exploration concept, we restricted the software to explore pharmacophoric models incorporating from zero to one Neglon feature, from zero to three HBA, Hbic, and RingArom features instead of the default range of 0–5, as shown in Table C under Supplementary Material. Furthermore, we instructed HYPOGEN to explore only 4- and 5-featured pharmacophores, i.e., ignore models of lesser number of features in order to further narrow the investigated pharmacophoric space and to better represent the diverse interactions between known ligands and PPARγ binding pocket (as shown in Table C under Supplementary Material).

In each run, the resulting binding hypotheses were automatically ranked according to their corresponding "total cost" value, which is defined as the sum of error cost, weight cost and configuration cost (see Section 4.1.5) [68–72]. Error cost provides the highest contribution to total cost and it is directly related to the capacity of the particular pharmacophore as 3D-QSAR model, i.e., in correlating the molecular structures to the corresponding biological responses [68–72]. HYPOGEN also calculates the cost of the null hypothesis, which presumes that there is no relationship in the data and that experimental activities are normally distributed about their mean. Accordingly, the greater the difference from the null hypothesis cost (residual cost, Table 1) the more likely that the hypothesis does not reflect a chance correlation. An additional validation technique based on Fisher's randomization test [73] was recently introduced

Please cite this article in press as: B.O. Al-Najjar, et al., Discovery of new nanomolar peroxisome proliferator-activated receptor γ activators via elaborate ligand-based modeling, European Journal of Medicinal Chemistry (2011), doi:10.1016/j.ejmech.2011.03.040

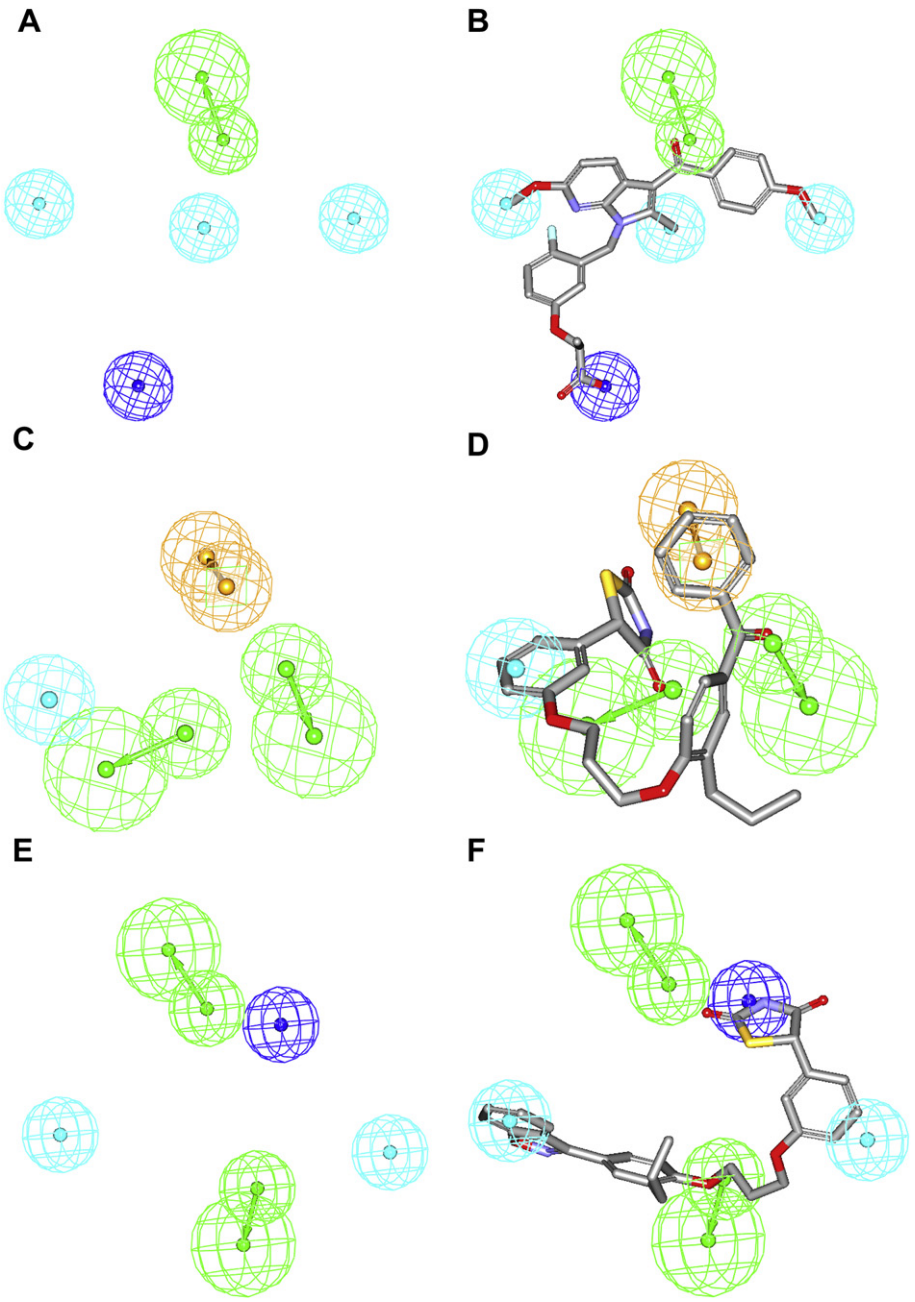

Fig. 3. Pharmacophoric features of (A) $S_1R_3H_1$, (C) $S_1R_5H_8$, (E) $S_3R_2H_2$, (B) $S_1R_3H_1$ fitted against **10** ($IC_{50} = 1$ nM) (D) $S_1R_5H_8$ fitted against **49** ($IC_{50} = 19$ nM), (F) $S_3R_2H_8$ fitted against **53** ($IC_{50} = 30$ nM). HBA shown as green vectored spheres, Hbic as light blue spheres, RingArom as vectored orange spheres and Neglon as dark blue sphere. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

into CATALYST: Cat. Scramble [64]. In this test the biological data and the corresponding structures are scrambled several times and the software is challenged to generate pharmacophoric models from the randomized data. The confidence in the parent hypotheses (i.e., generated from unscrambled data) is lowered proportional to the number of times the software succeeds in generating binding hypotheses from scrambled data of apparently better cost criteria than the parent hypotheses (see Section 4.1.5) [68–72].

Eventually, 519 pharmacophore models emerged from 56 automatic HYPOGEN runs, out of which only 500 models illustrated Cat. scramble confidence levels $\geq 90\%$. These successful models were clustered and the best representatives (104 models) were used in subsequent QSAR modeling (see Section 4.1.6). Table 1 shows the statistical criteria of representative cluster centers. Clearly, from the table, representative models shared comparable features and acceptable statistical success criteria.

The fact that many pharmacophore models were optimal and statistically comparable suggests the ability of PPARγ ligands to assume multiple pharmacophoric binding modes within the binding pocket. Therefore, it is quite challenging to select any

Table 2
The cross-correlation r^2 between the successful pharmacophores hypotheses in Eq. (1).

	$S_1R_3H_1$	$S_3R_2H_2$	$S_1R_5H_8$
$S_1R_3H_1$	1		
$S_3R_2H_2$	0.74	1	
$S_1R_5H_8$	0.25	0.20	1

Please cite this article in press as: B.O. Al-Najjar, et al., Discovery of new nanomolar peroxisome proliferator-activated receptor γ activators via elaborate ligand-based modeling, European Journal of Medicinal Chemistry (2011), doi:10.1016/j.ejmech.2011.03.040

particular pharmacophore hypothesis as a sole representative of the binding process.

2.2. QSAR modeling

Despite that pharmacophoric hypotheses provide excellent insights into ligand-macromolecule recognition and can be used to mine for new biologically interesting scaffolds, their predictive value as 3D-QSAR models is limited by steric shielding and bioactivity-enhancing or reducing auxiliary groups [54–58,60,61,74]. This point combined with the fact that pharmacophore modeling of PPARγ activators furnished several binding hypotheses of comparable success criteria prompted us to employ classical QSAR analysis to search for the best combination of pharmacophore(s) and other 2D descriptors capable of explaining bioactivity variation across the whole list of collected inhibitors (**1**–**88**, Table A in Supplementary Materials and Fig. 1). We employed genetic function approximation and multiple linear regression QSAR (GFA-MLR-QSAR) analysis to search for an optimal QSAR equation(s).

The fit values obtained by mapping representative hypotheses (104 models) against collected PPARγ activators (**1**–**88**) were enrolled, together with nearly 100 other physicochemical descriptors, as independent variables (genes) in GFA-MLR-QSAR analysis (see Section 4.1.7) [54–58,60,61,74,75]. However, since it is essential to access the predictive power of the resulting QSAR models on an external set of inhibitors, we randomly selected 17 molecules (marked with double asterisks in Table A, see Section 4.1.7) and employed them as external test molecules for validating the QSAR models (r_{PRESS}^2). Moreover, all QSAR models were cross-validated automatically using the leave-one-out cross-validation in Discovery Studio 2.5 [64].

Eq. (1) shows the details of the optimal QSAR model. Fig. 2 shows the corresponding scatter plots of experimental versus estimated bioactivities for the training and testing inhibitors.

where, r_{71}^2 is the correlation coefficient against 71 training compounds, r_{LOO}^2 is the leave-one-out correlation coefficient, r_{adj}^2 is r^2 adjusted for the number of terms in the model and r_{PRESS}^2 is the predictive r^2 determined for the 17 test compounds [64,75]. $S_1R_3H_1$ and $S_1R_5H_8$ represent the fit values of the training compounds against the first and eighth pharmacophoric Hypotheses generated in third and fifth HYPOGEN Runs, respectively, using the first training Subset, while $S_3R_2H_2$ represents the fit values against the second pharmacophoric Hypothesis generated during from the second HYPOGEN Run performed on the third training Subset. Table C (under Supplementary Materials) and Table 1 show the pharmacophore modeling runs and the statistical criteria of output models. Bolded runs in Table 1 correspond to the three QSAR-selected pharmacophores (i.e., $S_1R_3H_1$, $S_1R_5H_8$ and $S_3R_2H_2$). The fit values were calculated based on Eq. (D) under Supplementary Materials.

Fig. 3 shows $S_1R_3H_1$, $S_1R_5H_8$ and $S_3R_2H_2$ and how they map three potent training compounds, namely, **10** ($IC_{50} = 1$ nM), **49** ($IC_{50} = 19$ nM), **53** ($IC_{50} = 30$ nM), while Table 3 shows the *X*,*Y*, and *Z* coordinates of the three pharmacophores.

HBD is the number of hydrogen bond donors in a particular molecule, RotableBonds is the number of rotable bonds, FPSA is the molecular fractional polar surface area (calculated as the ratio of the polar surface area divided by the total surface area), and LUMO is the energy of the lowest unoccupied molecular orbital calculated employing the density functional theory method implemented in DS 2.5 [64].

Emergence of three orthogonal pharmacophoric models, i.e., $S_1R_3H_1$, $S_1R_5H_8$ and $S_3R_2H_2$ (of cross-correlation $r^2 \leq 0.74$, Table 2) in Eq. (1) suggests that they represent three complementary binding modes accessible to ligands within the binding pocket of PPARγ, which means that one of the pharmacophores can optimally explain the bioactivities of some training inhibitors, while the others explain the remaining inhibitors. Similar conclusions were

$$\log(1/IC_{50}) = -2.4 + 0.30(S_1R_5H_8) + 5.63 \times 10^{-2}(S_3R_2H_2) + 9.76 \times 10^{-2}(S_1R_3H_1) + 5.1\mathrm{FPSA} - 9.3\mathrm{LUMO} - 0.05(\mathrm{Rotable\ Bonds}) - 0.28\mathrm{HBD}$$
$$r_{71}^2 = 0.80, F - \mathrm{statistic} = 270.3, r_{adj}^2 = 0.77, r_{LOO}^2 = 0.73, r_{PRESS(17)}^2 = 0.67 \qquad (1)$$

Table 3
Pharmacophoric features and corresponding weights, tolerances and 3D coordinates of $S_1R_3H_1$, $S_1R_5H_8$ and $S_3R_2H_2$.

Model	Definitions		Chemical features						
			HBA			Hbic	Hbic	Hbic	NegIons
$S_1R_3H_1$[a]	Weights		2.29089			2.29089	2.29089	2.29089	2.29089
	Tolerances		1.60	2.20		1.60	1.60	1.60	1.60
	Coordinates	*X*	2.79	3.17		1.18	−0.38	−1.61	−2.71
		Y	−4.10	−6.41		−3.57	0.77	−1.65	6.3
		Z	0.46	−1.42		7.38	−6.1	1.43	1.11
			HBA		HBA		Hbic	RingArom	
$S_1R_5H_8$[b]	Weights		2.57353		2.57353		2.57353	2.57353	
	Tolerances		1.60	2.20	1.60	2.20	1.60	1.60	1.60
	Coordinates	*X*	3.09	1.69	1.19	−0.55	0.33	4.84	6.41
		Y	−4.19	−5.19	−0.53	1.90	3.62	−1.90	−0.44
		Z	0.33	−2.13	−3.23	−3.57	0.13	2.62	0.52
			Hbic	Hbic	NegIons	HBA		HBA	
$S_3R_2H_2$[c]	Weights		2.42026	2.42026	2.42026	2.42026		2.42026	
	Tolerances		1.60	1.60	1.60	1.6	2.20	1.6	2.20
	Coordinates	*X*	−0.27	12.92	9.44	7.65	6.30	5.83	4.16
		Y	−5.46	−0.84	−2.87	−4.41	−4.25	−0.06	1.15
		Z	−1.94	−4.23	1.57	3.25	5.92	−4.92	−7.10

[a] $S_1R_3H_1$: the code refer to training Subset I (see Table B in Supplementary Materials), third Run and first pharmacophore Hypothesis, as in Table 1 and Table C under Supplementary Materials.
[b] The 8th Hypothesis from 5th Run on 1st training Subset.
[c] The 2nd Hypothesis from 2nd Run on 3rd Subset.

Please cite this article in press as: B.O. Al-Najjar, et al., Discovery of new nanomolar peroxisome proliferator-activated receptor γ activators via elaborate ligand-based modeling, European Journal of Medicinal Chemistry (2011), doi:10.1016/j.ejmech.2011.03.040

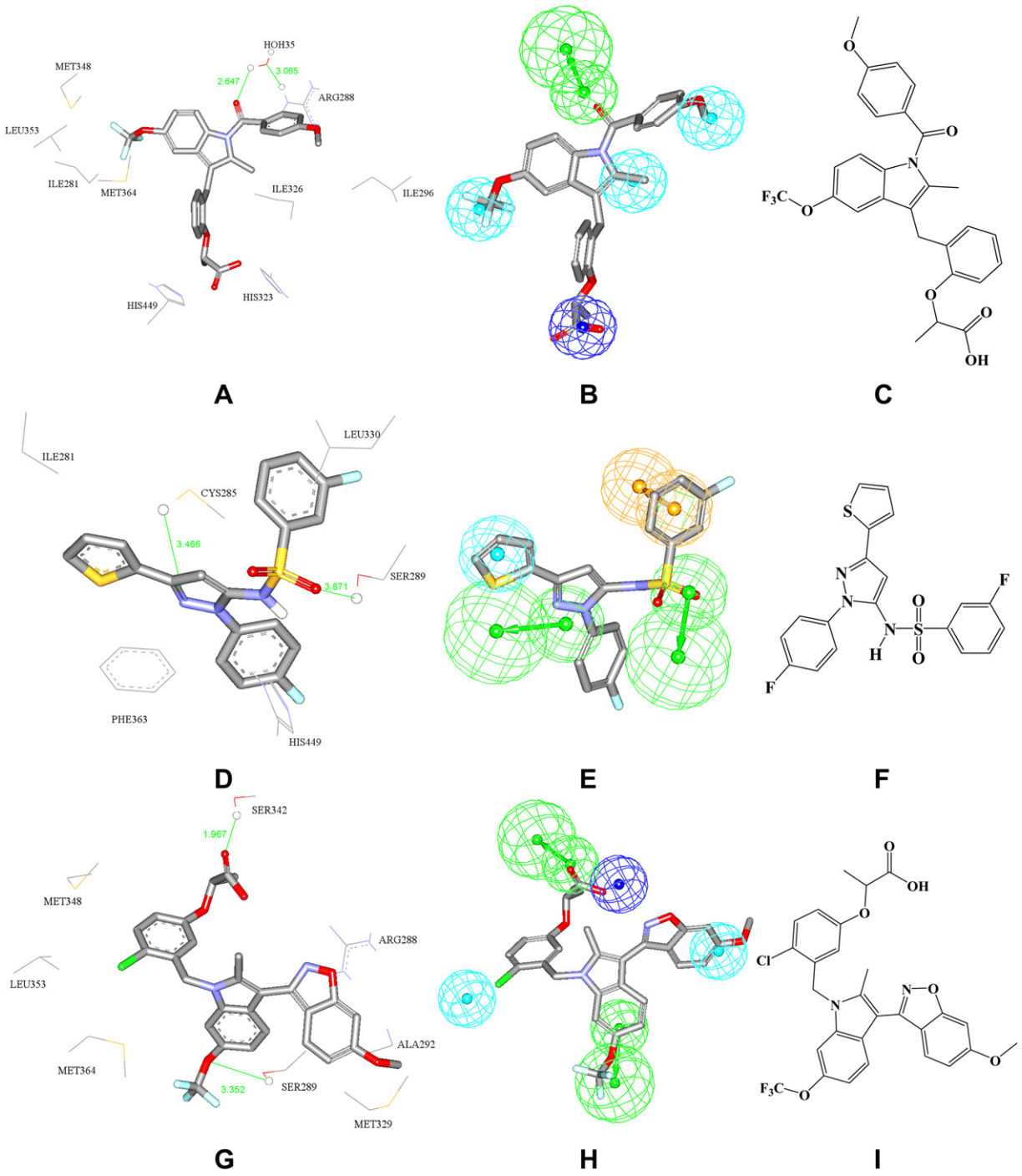

Fig. 4. X-ray structures of three ligands co-crystallized within PPARγ binding pocket: (A) 2Q59 (resolution = 2.20 Å), (D) 2G0 G (resolution = 2.54 Å) and (G) 2P4Y (resolution = 2.25 Å). (B), (E) and (H) mapping pharmacophores $S_1R_3H_1$, $S_1R_5H_8$, and $S_3R_2H_2$ against the co-crystallized ligands of 2Q59, 2G0G, and 2P4Y (rigid mapping), respectively; (C), (F) and (I) the chemical structures of the co-crystallized ligands of 2Q59, 2G0G, and 2P4Y, respectively.

reached about the binding pockets of other targets based on QSAR analysis [54–58,60,61,74].

Emergence of FPSA in Eq. (1) in association with positive slope suggests a direct relationship between ligand/PPARγ affinity and ligands' hydrophilicity. We believe this trend is explainable by the fact that hydrophilic ligands favor docking into the binding site due to the presence of good number of hydrophilic amino acid residues in binding site (e.g., Lys367, Ser289, Arg288, Glu286, His449, His323, Tyr473, Gln283, and Asp362). However, this seems to contradict with the appearance of HBD in Eq. (1) combined with negative slope, which suggest that ligands of more hydrogen bond donors disfavor binding. Nevertheless, careful evaluation of training compounds reveals that excess HBDs in poorly potent ligands are always negative ionizable, thus they render their corresponding ligands extremely hydrophilic and favor hydration instead of binding within bindng pocket.

On the other hand, emergence of LUMO in Eq. (1) combined with negative slope suggests that ligand/PPARγ affinity favors

Please cite this article in press as: B.O. Al-Najjar, et al., Discovery of new nanomolar peroxisome proliferator-activated receptor γ activators via elaborate ligand-based modeling, European Journal of Medicinal Chemistry (2011), doi:10.1016/j.ejmech.2011.03.040

Table 4
ROC curve analysis criteria for QSAR-selected pharmacophores and their sterically refined versions.

Pharmacophore model	ROC[a]–AUC[b]	ACC[c]	SPC[d]	TPR[e]	FNR[f]
$S_1R_3H_1$	0.99	0.96	0.97	0.56	0.026
$S_1R_5H_8$	0.82	0.96	0.95	1.00	0.047
$S_3R_2H_2$	0.97	0.96	0.99	0.28	0.013

[a] ROC: receiver operating characteristic curve.
[b] AUC: area under the curve.
[c] ACC: overall accuracy.
[d] SPC: overall specificity.
[e] TPR: overall true positive rate.
[f] FNR: overall false negative rate.

electrophilic ligands (i.e., of more negative LUMO values) probably due to π-stacking with certain electron-rich aromatic centers in the binding pocket (e.g., Phe363, Phe287, Tyr327, and Tyr473).

Finally, appearance of RotableBonds in Eq. (1) combined with negative regression coefficient suggests that flexible molecules disfavor binding with PPARγ, which is not unexpected since the entropic cost of protein binding with flexible ligands is much higher that that required for rigid molecules. Therefore, a significant fraction of energy gains resulting from binding enthalpy will be wasted on entropic costs required for fixing flexible molecules within the binding pocket.

2.3. Comparing pharmacophore models with crystallographic complexes

To further emphasize the validity of our QSAR-selected pharmacophores, we compared the crystallographic structures of three PPARγ/ligand complexes (PDB codes: 2Q59, 2G0G, and 2P4Y) with $S_1R_3H_1$, $S_1R_5H_8$, and $S_3R_2H_2$. Fig. 4 shows the chemical structures of the ligands and compares their PPARγ complexes with the ways they map $S_1R_3H_1$, $S_1R_5H_8$, and $S_3R_2H_2$ employing rigid mapping, i.e., fitting the ligands' bound states against corresponding pharmacophores without conformational adjustments.

Fitting the carboxylate of the 2Q59 ligand against NegIon in $S_1R_3H_1$ (Fig. 4b) corresponds to electrostatic interactions connecting this fragment with the imidazole side chains of His449 and His323, as in Fig. 4a. Similarly, mapping the ligand's amidic carbonyl with

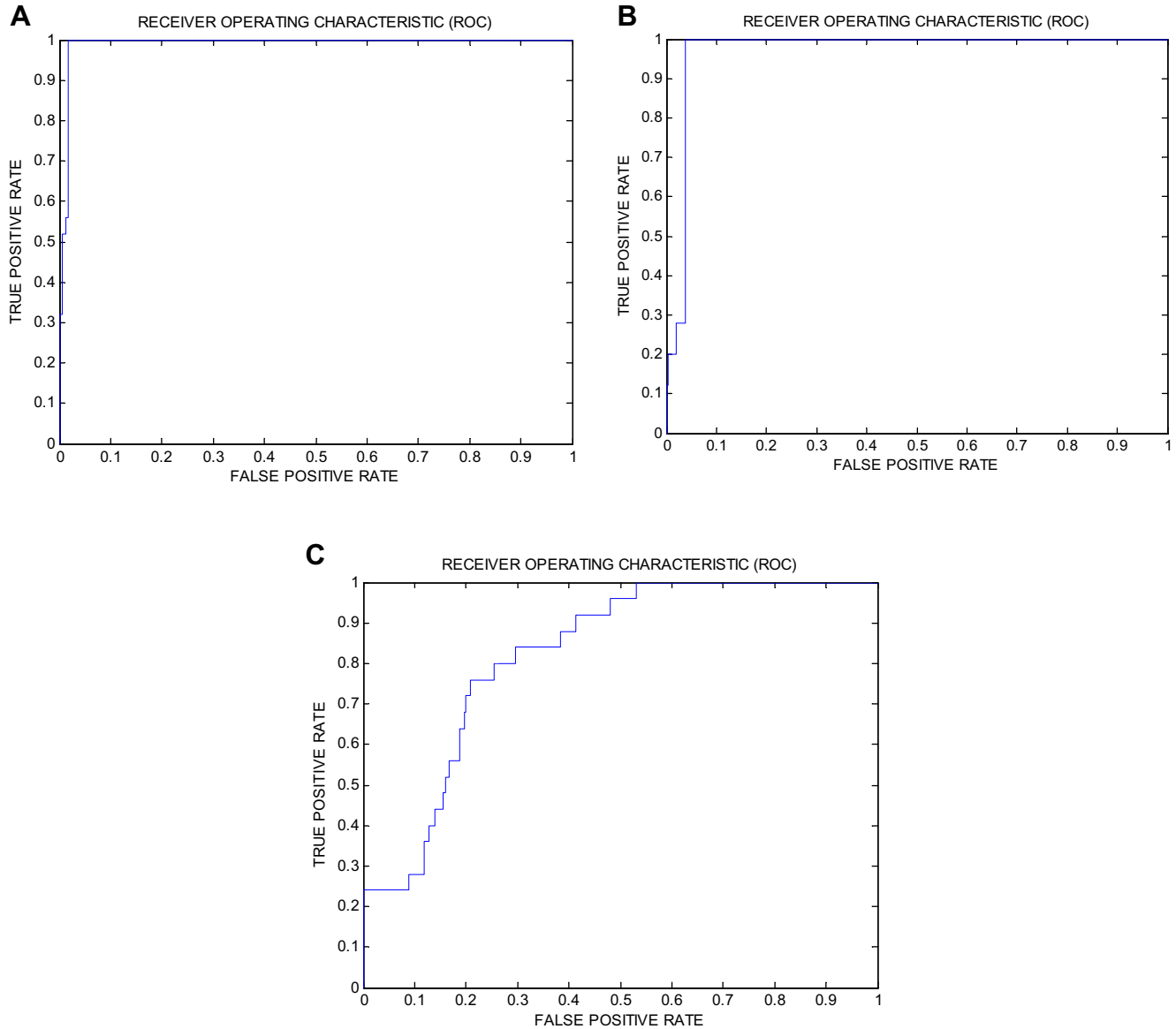

Fig. 5. ROC curves of: (A) $S_1R_3H_1$, (B) $S_3R_2H_2$, (C) $S_1R_5H_8$.

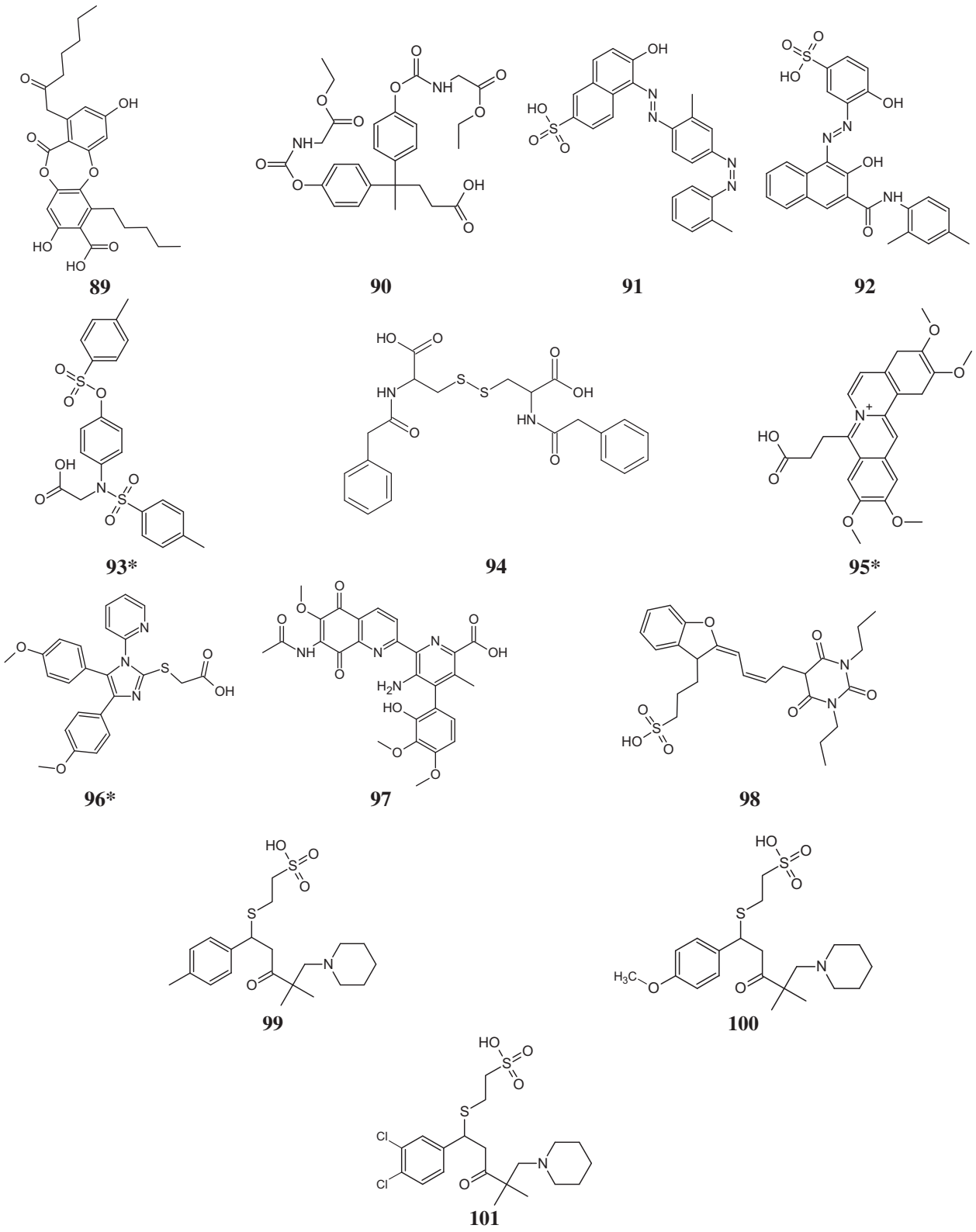

Fig. 6. Chemical structure of captured hits. Active hits are marked with asterisks. Corresponding bioactivities are shown in Table 5.

a HBA feature in $S_1R_3H_1$ (Fig. 4b) corresponds to hydrogen bonding interaction between this carbonyl and the explicit water molecule H_2O35. Apparently, this water molecule is fixed via hydrogen bonding with guanidine of Arg288 (Fig. 4a). Similarly, mapping the trifluoromethoxy of 2Q59 ligand against Hbic feature in $S_1R_3H_1$ (Fig. 4b) correlates with fitting this group into a hydrophobic pocket composed of the hydrophobic side chains of Met364, Ile281, Le353 and Met348, as in Fig. 4a. Finally, mapping the indole methyl and

Please cite this article in press as: B.O. Al-Najjar, et al., Discovery of new nanomolar peroxisome proliferator-activated receptor γ activators via elaborate ligand-based modeling, European Journal of Medicinal Chemistry (2011), doi:10.1016/j.ejmech.2011.03.040

Table 5
Predicted and experimental bioactivities of high-ranking hit molecules.

Hits[a]	Name[b]	Best fit values[c]			Predicted affinity IC_{50} (nM)[d]	Experimental EC_{50}(nM)[e,g]
		$S_1R_3H_1$	$S_1R_5H_8$	$S_3R_2H_2$		
93	NCI144248	3.59542	8.42089	7.34686	2.4	15.3 ($r^2 = 0.99$)[f]
95	NCI197178	2.11611	7.40627	0	3.6	220×10^3 ($r^2 = 0.98$)[f]
96	NCI289920	0.475151	4.59272	0	224.8	224 ($r^2 = 0.99$)[f]
Rosiglitazone	N/A	0	5.332	0	3.5	10 ($r^2 = 0.95$)[f]

[a] Chemical structures shown in Fig. 6.
[b] NCI number.
[c] Fit values calculated against respective hypotheses using Eq. (D) in Supplementary Materials.
[d] Predicted IC_{50} nM according to QSAR Eq. (1).
[e] Experimental bioactivation (EC_{50} values) determined in triplicates.
[f] Values in brackets represent the correlation coefficients of the corresponding dose-response line.
[g] These values represent average results of triplicate measurements.

aromatic methoxy of the co-crystallized ligand against two Hbic features in $S_1R_3H_1$ (Fig. 4b) agrees with positioning these groups at close proximity with the hydrophobic side chains of Ile326 and Ile296, respectively, in the bound crystallographic structure of the ligand (Fig. 4a).

A similar trend can be seen by comparing the crystallographic structure of bound ligand in 2G0 G with the way it maps $S_1R_5H_8$ (Fig. 4d and e). Mapping the ligand's pyrazole nitrogen and sulfonamidic oxygen against two HBAs in $S_1R_5H_8$ corresponds to hydrogen bonding interactions connecting these atoms with SH and OH side chains of Cys285 and Ser289, respectively, as in Fig. 4d and e. On the other hand, mapping the fluorobenzene and thiophene rings of the bound ligand against RingArom and Hbic features in $S_1R_5H_8$, respectively (Fig. 4e), correlates with stacking these rings against the hydrophobic side chains of Leu330 and Ile281 in 2G0G (Fig. 4d).

Finally, $S_3R_2H_2$ seems to match the co-crystallized complex 2P4Y (Fig. 4g and h): Mapping the ligand's carboxylate against HBA and NegIon features in $S_3R_2H_2$ corresponds to hydrogen bonding and electrostatic attraction connecting this carboxylate with the OH and guanidine side chains of Ser342 and Arg288 in 2P4Y. Similarly, mapping the methoxy benzene ring against Hbic feature in $S_3R_2H_2$ correlates with projecting this group close to the hydrophobic side chains of Met329 and Ala292 in 2P4Y. The hydrogen bonding interaction between the OH of Ser289 and the ether oxygen of trifluoromethoxy of 2P4Y ligand agrees with mapping this group against HBA feature in $S_3R_2H_2$. Finally, the close proximity between the chloro- substituent of the ligand's central aromatic linker correlates with placing this group in a hydrophobic pocket comprised of the side chains of Met348, Leu353 and Met364.

Clearly from the above discussion, the three pharmacophores $S_1R_3H_1$, $S_1R_5H_8$, and $S_3R_2H_2$ represent three valid binding modes assumed by ligands within PPARγ. Incidentally, the three pharmacophores point to limited number of critical interactions required for high ligand-PPARγ affinity in each of the binding modes. In contrast, crystallographic complexes reveal many bonding interactions without highlighting critical ones. Fig. 4a, d and g shows only interactions corresponding to pharmacophoric features while other binding interactions were hidden for clarity.

2.4. Receiver operating characteristic (ROC) curve analysis

To further validate the resulting models (both QSAR and pharmacophores), we subjected our QSAR-selected pharmacophores to receiver operating curve (ROC) analysis. In ROC analysis, the ability of a particular pharmacophore model to correctly classify a list of compounds as actives or inactives is indicated by the area under the curve (AUC) of the corresponding ROC as well as other parameters, namely, overall accuracy, overall specificity, overall true positive rate and overall false negative rate (see Section 4.1.8 for more details) [76–79].

Table 4 and Fig. 5 show the ROC results of our QSAR-selected pharmacophores. $S_1R_3H_1$ and $S_3R_2H_2$ illustrated excellent overall performances with ROC-AUC values exceeding 95%. On the other hand, $S_1R_5H_8$ exhibited moderate performance with AUC value of 82%. This is not unexpected, as both $S_1R_3H_1$ and $S_3R_2H_2$ are 5-featured pharmacophores, while $S_1R_5H_8$ exhibit 4 features only. Higher-featured pharmacophores are expected to be more selective as 3D search queries since additional features impose more provisions on captured hits.

2.5. In silico screening and subsequent in vitro evaluation

$S_1R_3H_1$, $S_1R_5H_8$ and $S_3R_2H_2$ were employed as 3D search queries against the National Cancer Institute list of compounds (NCI, 238,819 structures) using the "Best Flexible Database Search" option implemented within CATALYST. Compounds that have their chemical groups spatially overlap (map) with corresponding features of each particular pharmacophoric model were captured as hits. Captured hits were filtered based on Lipinski's and Veber's rules [80,81]. Surviving hits were fitted against $S_1R_3H_1$, $S_1R_5H_8$, $S_3R_2H_2$ and their fit values, together with other relevant molecular descriptors, were substituted in QSAR Eq. (1) to predict their affinity IC_{50} values. The highest-ranking available hits (13 compounds, Fig. 6) were evaluated *in vitro* for potential PPARγ ligand activity using a PPRE-luciferase reporter system transfected in HepG2 cells. Hits were initially screened at 40 μM concentrations, subsequently; compounds that significantly activated PPARγ were further assessed to determine their EC_{50} values. Table 5 lists hits that illustrated significant PPARγ ligand activities; estimated affinities to PPARγ and *in vitro* experimental bioactivation EC_{50} values.

To validate our bioassay settings we determined the EC_{50} value of the standard PPARγ activator rosiglitazone under the same conditions. Our conditions determined the EC_{50} value of rosiglitazone to be 10 nM, which is within the reported literature range of 2 to 16 nM [82,83].

Clearly from Table 5, *in vitro* testing showed that 3 NCI high-ranking hits activated PPARγ at nanomolar to micromolar EC_{50} values. Fig. 6 shows the chemical structures of the tested hits including active ones, while Fig. 7 shows how the most potent hit **93** fits the three successful pharmacophore models.

However, it must be mentioned that QSAR and pharmacophore modeling were based on affinity values (IC_{50}), and therefore, the corresponding predictions were in IC_{50} format. On the other hand, since we implemented a functional bioassay that detects agonistic bioactivities of captured hits, i.e., EC_{50} values, explains the limited number of active hits and the apparent differences between predicted and experimental bioactivities.

Please cite this article in press as: B.O. Al-Najjar, et al., Discovery of new nanomolar peroxisome proliferator-activated receptor γ activators via elaborate ligand-based modeling, European Journal of Medicinal Chemistry (2011), doi:10.1016/j.ejmech.2011.03.040

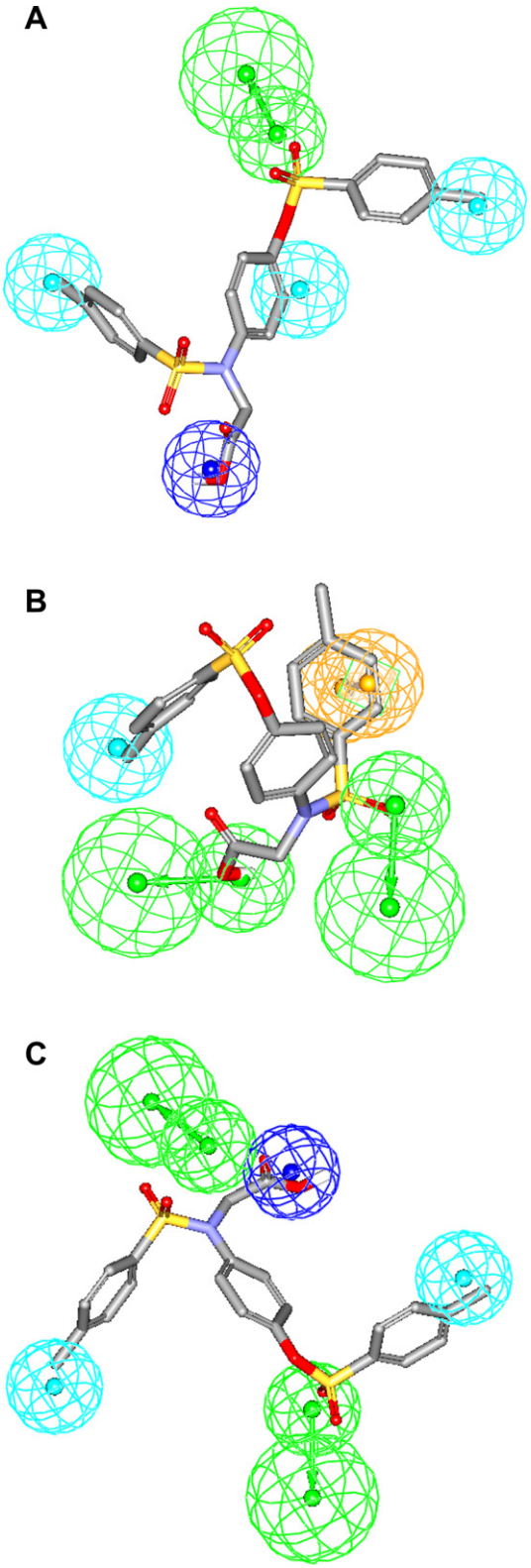

Fig. 7. Mapping active hit 93 (NCI144248, $EC_{50} = 15.3$ nM) against (A) $S_1R_3H_1$, (B) $S_1R_5H_8$, and (C) $S_3R_2H_8$ (see Table 5). HBA shown as green vectored spheres, Hbic as light blue spheres, RingArom as vectored orange spheres and Neglon as dark blue sphere. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3. Conclusions

PPARγ activators are currently considered as potential treatments for diabetes and hyperglycemia. The pharmacophoric space of PPARγ activators was explored via eight diverse sets of activators and using CATALYST-HYPOGEN module of Discovery Studio to identify high-quality binding model(s). Subsequently, genetic algorithm and multiple linear regression analysis were employed to access optimal QSAR model capable of explaining PPARγ activation variation across 88 collected PPARγ activators. Three orthogonal pharmacophoric models emerged in the QSAR equation suggesting the existence of at least three distinct binding modes accessible to activating ligands within PPARγ binding pocket. The QSAR equation and the associated pharmacophoric models were experimentally validated by the identification of three nanomolar to micromolar PPARγ activators retrieved via *in silico* screening. Our results suggest that the combination of pharmacophoric exploration and QSAR analyses can be useful tool for finding new diverse PPARγ activators.

4. Experimental

4.1. Molecular modeling

4.1.1. Software and hardware

The following software packages were utilized in the present research.

- CS ChemDraw Ultra 6.0, Cambridge Soft Corp. (www.cambridgesoft.com), USA.
- Discovery Studio 2.5, Accelrys Inc. (www.accelrys.com), USA.

Pharmacophore and QSAR modeling studies were performed using Discovery Studio 2.5 suite from Accelrys Inc. (San Diego, California, www.accelrys.com) installed on installed on a Core 2 Duo Pentium PC.

4.1.2. Dataset

The structures of 88 PPARγ ligands (Table A under Supplementary Materials) were collected from published literature [16,18–20,66,67]. The *in-vitro* bioactivities of the collected agonists were expressed as the concentration of the test compound that causes 50% displacement of a radio-labeled full agonist bound to the receptor (IC_{50}). The logarithm of measured IC_{50} (nM) values were used in pharmacophore modeling and QSAR analysis, thus correlating the data linear to the free energy change. In one instance the IC_{50} was reported to be below 1 nM (compound **13**, Table A under Supplementary Materials), however, to allow proper QSAR modeling we assumed that IC_{50} equals 1 nM. The logarithmic transformation of IC_{50} values should minimize any potential errors resulting from this assumption.

The two-dimensional (2D) chemical structures of the inhibitors were sketched using ChemDraw Ultra and saved in MDL-molfile format. Subsequently, they were imported into Discovery Studio, converted into corresponding standard 3D structures and energy minimized to the closest local minimum using the molecular mechanics CHARMm force field implemented in CATALYST module of Discover Studio. The resulting 3D structures were utilized as starting conformers for CATALYST-based conformational analysis.

4.1.3. Conformational analysis

The conformational space of each agonist (**1**–**88**, Table 1) was explored adopting the "best conformer generation" option within CATALYST module of Discovery Studio, which is based on the generalized CHARMm force field implemented in the program. Conformational ensembles were generated with an energy

Please cite this article in press as: B.O. Al-Najjar, et al., Discovery of new nanomolar peroxisome proliferator-activated receptor γ activators via elaborate ligand-based modeling, European Journal of Medicinal Chemistry (2011), doi:10.1016/j.ejmech.2011.03.040

threshold of 20 kcal/mol from the local minimized structure and a maximum limit of 250 conformers per molecule.

4.1.4. Pharmacophoric hypotheses generation

All 88 molecules with their associated conformational models were regrouped into a spreadsheet. The biological data of the inhibitors were reported with uncertainty values of 2 or 3, which means that the actual bioactivity of a particular inhibitor is assumed to be situated somewhere in intervals ranging from 1/2 to 2 or 1/3 to 3 times the reported bioactivity value of that inhibitor, respectively (See Table C under Supplementary Materials) [70,84,85]. The uncertainty value is of great impact on the qualities of the resulting pharmacophores, as it controls the number of training compounds within the "most potent category" (see Eq. (A) under section SM-1 in Supplementary Materials).

Subsequently, seven structurally diverse training subsets were carefully selected from the collection for pharmacophore modeling. Table B under Supplementary Material shows the selected subsets and their member compounds. Typically, CATALYST requires informative training sets that include at least 16 compounds of evenly spread bioactivities over at least three and a half logarithmic cycles. Lesser training lists could lead to chance correlation and thus faulty models.

The selected training sets were utilized to conduct 56 modeling runs to explore the pharmacophoric space of PPARγ (Table C under Supplementary Material). The exploration process included altering interfeature spacing parameter (100 and 300 picometers), the uncertainty value (2 or 3) and the maximum number of allowed features in the resulting pharmacophore hypotheses, i.e., they were allowed to vary from 4 to 5 or from 5 to 5. Furthermore, some features were fixed in some runs, i.e., NegIon was fixed by setting the number of possible NegIon features in the resulting pharmacophore models to one, while other features were allowed to vary during pharmacophore modeling, as shown in Table C under Supplementary Material.

Pharmacophore modeling employing CATALYST proceeds through three successive phases: the constructive phase, subtractive phase and optimization phase (see CATALYST Modeling Algorithm in section SM-1 under Supplementary Materials) [64,68–72].

4.1.5. Assessment of the generated hypotheses

When generating hypotheses, CATALYST attempts to minimize a cost function consisting of three terms: Weight cost, Error cost and Configuration cost (see CATALYST Cost Analysis in Assessment of Generated Binding Hypotheses under Supplementary Materials).

An additional approach to assess the quality of CATALYST-HYPOGEN pharmacophores is to cross-validate them using the Cat-Scramble program implemented in CATALYST. This validation procedure is based on Fischer's randomization test [73]. In this validation tes, we selected a 95% confidence level, which instruct CATALYST to generate 19 random spreadsheets by the Cat-Scramble command. Subsequently, CATALYST-HYPOGEN is challenged to use these random spreadsheets to generate hypotheses using exactly the same features and parameters used in generating the initial unscrambled hypotheses. Success in generating pharmacophores of comparable cost criteria to those produced by the original unscrambled data reduces the confidence in the training compounds and the unscrambled original pharmacophore models [64,73,86]. Based on Fischer randomization criteria; all the pharmacophores exceeded the 85% significance threshold, and therefore were considered fit for subsequent processing (clustering and QSAR analyses).

4.1.6. Clustering of the generated pharmacophore hypotheses

The pharmacophore models were clustered into 104 groups utilizing the hierarchical average linkage method available in CATALYST. Subsequently, the highest-ranking representatives, as judged based on their significance *F*-values, were selected to represent their corresponding clusters in subsequent QSAR modeling. Table 1 shows the pharmacophoric features and statistical criteria of representative pharmacophores including their pharmacophoric features, success criteria and differences from corresponding null hypotheses. The table also shows the corresponding Cat. Scramble confidence levels determined for each representative pharmacophore.

4.1.7. QSAR modeling

A subset of 72 compounds from the total list of inhibitors (**1**–**88**, Table A under Supplementary Materials and Fig. 1) [16,18–20,66,67] was utilized as a training set for QSAR modeling; the remaining 16 molecules (*ca.* 20% of the dataset) were employed as an external test subset for validating the QSAR models. The test molecules were selected as follows: the 88 agonists were ranked according to their IC_{50} values, and then every fifth compound was selected for the test set starting from the high-potency end. This selection considers the fact that the test molecules must represent a range of biological activities similar to that of the training set. The selected test inhibitors are marked with double asterisks in Table A (Supplementary Material).

The logarithm of measured $1/IC_{50}$ (μM) values was used in QSAR, thus correlating the data linear to the free energy change. The chemical structures of the inhibitors were imported into Discovery Studio as standard 3D single conformer representations in SD format. Subsequently, different descriptor groups were calculated for each compound. The calculated descriptors included various simple and valence connectivity indices, electro-topological state indices and other molecular descriptors (e.g., logarithm of partition coefficient, polarizability, dipole moment, molecular volume, molecular weight, molecular surface area, etc.) [64]. The training compounds were fitted (using the best fit option in CATALYST module of Discovery Studio) [64] against the representative pharmacophores, and their fit values were added as additional descriptors. The fit value for any compound is obtained automatically via Eq. (D) (under Section SM-2 in Supplementary Materials) [64].

Genetic function approximation (GFA) was employed to search for the best possible QSAR regression equation capable of correlating the variations in biological activities of the training compounds with variations in the generated descriptors, i.e., multiple linear regression modeling (MLR). GFA techniques rely on the evolutionary operations of "crossover and mutation" to select optimal combinations of descriptors (i.e., chromosomes) capable of explaining bioactivity variation among training compounds from a large pool of possible descriptor combinations, i.e., chromosomes population. However, to avoid overwhelming GFA-MLR with large number of poor descriptor populations, we removed lowest-variance descriptors (20%) prior to QSAR analysis.

Each chromosome is associated with a fitness value that reflects how good it is compared to other solutions. The fitness function employed herein is based on Friedman's 'lack-of-fit' (LOF) [64].

Our preliminary diagnostic trials suggested the following optimal GFA parameters: explore linear, quadratic and spline equations at mating and mutation probabilities of 50%; population size = 500; number of genetic iterations = 30,000 and lack-of-fit (LOF) smoothness parameter = 1.0. However, to determine the optimal number of explanatory terms (QSAR descriptors), it was decided to scan and evaluate all possible QSAR models resulting from 4 to 10 explanatory terms.

All QSAR models were validated employing leave-one-out cross-validation (r_{LOO}^2), bootstrapping (r_{adj}^2) and predictive r^2 (r_{PRESS}^2) calculated from the test subsets. The predictive r_{PRESS}^2 is defined as in Eq. (2):

Please cite this article in press as: B.O. Al-Najjar, et al., Discovery of new nanomolar peroxisome proliferator-activated receptor γ activators via elaborate ligand-based modeling, European Journal of Medicinal Chemistry (2011), doi:10.1016/j.ejmech.2011.03.040

$$r^2_{\mathrm{PRESS}} = \mathrm{SD} - \mathrm{PRESS}/\mathrm{SD} \quad (2)$$

where SD is the sum of the squared deviations between the biological activities of the test set and the mean activity of the training set molecules, PRESS is the squared deviations between predicted and actual activity values for every molecule in the test set.

4.1.8. Receiver operating characteristic (ROC) curve analysis

The optimal pharmacophore models (i.e., $S_1R_3H_1$, $S_1R_5H_8$ and $S_3R_2H_2$) were validated by assessing their abilities to selectively capture diverse PPARγ active compounds from a large testing list of actives and decoys.

The testing list was prepared as described by Verdonk and coworkers [76,77]. Briefly, decoy compounds were selected based on three basic one-dimensional (1D) properties that allow the assessment of distance (D) between two molecules (e.g., i and j): (1) the number of hydrogen-bond donors (NumHBD); (2) number of hydrogen-bond acceptors (NumHBA) and (3) count of nonpolar atoms (NP, defined as the summation of Cl, F, Br, I, S and C atoms in a particular molecule). For each active compound in the test set, the distance to the nearest other active compound is assessed by their Euclidean distance (Eq. (3)):

$$D(i,j) = \sqrt{(\mathrm{NumHBD}_i - \mathrm{NumHBD}_j)^2 + (\mathrm{NumHBA}_i - \mathrm{NumHBA}_j)^2 + (\mathrm{NP}_i - \mathrm{NP}_j)^2} \quad (3)$$

The minimum distances are then averaged over all active compounds (D_{min}). Subsequently, for each active compound in the test set, around 30 decoys were randomly chosen from the ZINC database [78]. The decoys were selected in such a way that they did not exceed D_{min} distance from their corresponding active compound.

To diversify active members in the list, we excluded any active compound having zero distance ($D(i,j)$) from other active compound (s) in the test set. Active testing compounds were defined as those possessing PPARγ affinities ranging from 0.6 nM to 6.0 μM. The test set included 25 active compounds and 776 ZINC decoys.

The test list (801 compounds) was screened by each particular pharmacophore employing the "Best flexible search" option implemented in Discovery Studio, while the conformational spaces of the compounds were generated employing the "Fast conformation generation option" implemented in Discovery Studio. Compounds missing one or more features were discarded from the hit list. *In-silico* hits were scored employing their fit values as calculated by Eq. (D) in Supplementary Materials.

The ROC curve analysis describes the sensitivity (Se or true positive rate, Eq. (4)) for any possible change in the number of selected compounds (n) as a function of (1-Sp). Sp is defined as specificity or true negative rate (Eq. (5)) [77,79].

$$\mathrm{Se} = \frac{\text{Number of Selected Actives}}{\text{Total Number of Actives}} = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}} \quad (4)$$

$$\mathrm{Sp} = \frac{\text{Number of Discarded Inactives}}{\text{Total Number of Inactives}} = \frac{\mathrm{TN}}{\mathrm{TN}+\mathrm{FP}} \quad (5)$$

where, TP is the number of active compounds captured by the virtual screening method (true positives), FN is the number of active compounds discarded by the virtual screening method, TN is the number of discarded decoys (presumably inactives), while FP is the number of captured decoys (presumably inactive) [77,79].

A ROC curve is plotted by setting the score (fit value) of the highest scoring active molecule as the first threshold. Afterwards, the number of decoys within this cutoff is counted and the corresponding Se and Sp pair is calculated [77,79]. This calculation is repeated for the active molecule with the second highest score and so forth, until the scores of all actives are considered as selection thresholds.

The ROC curve representing ideal distributions, where no overlap between the scores of active molecules and decoys exists, proceeds from the origin to the upper-left corner until all the actives are retrieved and Se reaches the value of 1. In contrast to that, the ROC curve for a set of actives and decoys with randomly distributed scores tends towards the Se = 1 − Sp line asymptotically with increasing number of actives and decoys [77,79]. The success of a particular virtual screening workflow can be judged from the following criteria (shown in Table 4):

(1) Area under the ROC curve (AUC) [77,79,87]. In an optimal ROC curve an AUC value of 1 is obtained; however, random distributions cause an AUC value of 0.5. Virtual screening that performs better than a random discrimination of actives and decoys retrieve an AUC value between 0.5 and 1, whereas an AUC value lower than 0.5 represents the unfavorable case of a virtual screening method that has a higher probability to assign the best scores to decoys than to actives [77,79].

(2) Overall Accuracy (ACC): describes the percentage of correctly classified molecules by the screening protocol (Eq. (7)). Testing compounds are assigned a binary score value of zero (compound not captured) or one (compound captured) [77,79,88].

$$\mathrm{ACC} = \frac{\mathrm{TP}+\mathrm{TN}}{N} = \frac{A}{N}\cdot\mathrm{Se} + \left(1 - \frac{A}{N}\right)\cdot\mathrm{Sp} \quad (7)$$

where, N is the total number of compounds in the testing database, A is the number of true actives in the testing database.

(3) Overall specificity (SPC): describes the percentage of discarded inactives by the particular virtual screening workflow. Inactive test compounds are assigned a binary score value of zero (compound not captured) or one (compound captured) regardless to their individual fit values [77,79,88].

(4) Overall true positive rate (TPR or overall sensitivity): describes the fraction percentage of captured actives from the total number of actives. Active test compounds are assigned a binary score value of zero (compound not captured) or one (compound captured) regardless to their individual fit values [77,79,88].

(5) Overall false negative rate (FNR or overall percentage of discarded actives): describes the fraction percentage of active compounds discarded by the virtual screening method. Discarded active test compounds are assigned a binary score value of zero (compound not captured) or one (compound captured) regardless to their individual fit values [77,79,88].

4.1.9. In silico screening for new PPARγ activators

$S_1R_3H_1$, $S_1R_5H_8$ and $S_3R_2H_2$ were employed as 3D search queries to screen the national cancer institute (NCI) 3D structural database. Virtual screening was performed employing the "Best Flexible Database Search" option implemented within the CATALYST module of Discovery Studio. The hits were filtered according to Lipinski's [80] and Veber's [81] rules and the remaining hits were

Please cite this article in press as: B.O. Al-Najjar, et al., Discovery of new nanomolar peroxisome proliferator-activated receptor γ activators via elaborate ligand-based modeling, European Journal of Medicinal Chemistry (2011), doi:10.1016/j.ejmech.2011.03.040

combined together and fitted against the three pharmacophores using the "best fit" option within CATALYST. The fit values together with the relevant molecular descriptors of each hit were substituted in QSAR Eq. (1). The highest-ranking molecules based on QSAR predictions were acquired and tested *in vitro*.

4.2. In vitro experimental studies

NCI hits were dissolved in DMSO in serial dilutions starting from *ca.* 40 μM. The amount of DMSO did not exceed 1% of the final concentration in each well. DMSO solution (1% v/v) was used as negative control, while rosiglitazone was used as positive control at a concentration range of 0.1–1.0 μM. All runs were repeated in triplicates.

Bioassay was performed in a similar way to previously reported methods [89]. Briefly, HepG2 cells (ATCC, Manassas, USA) were cultured in MEM with L-glutamine (Invitrogen Corporation, Carlsbald) supplemented with non-essential amino acids and sodium pyruvate (Sigma Aldrich, Germany), penicillin and streptomycin, and 10% fetal bovine serum at 37 °C in a humidified atmosphere containing 5% CO_2 in air. One day prior to transfection, cells were seeded into culture plates to reach about 60% confluence. Transfection was carried out using lipofectin (Invetrogen Corporation, Carlsbald) according to manufacturer's recommendation. In brief, the transfection reagent was mixed with optimum serum free medium (Invitrogen Corporation, Carlsbald) and incubated for 30–45 min before adding to the plasmid cocktail and incubated further for 15 min before adding another portion of serum free media and then pipetted slowly on the pre-washed cells and incubated overnight. The plasmids used in this study were PPREx3-Tk-Luc (a kind gift by Professor Ronald M. Evans of Salk Institute for Biological Studies, USA), pSV-sport PPARγ2 and pSV-sport RXRα both supplied by Professor Bruce Spiegelman of Harvard Medical school through Addgene website (www.addgene.org) and pRKTK plasmid (from Promega Corporation, USA).

Cells were transfected with 0.5 μg of the PPREx3-Tk-Luc, 0.17 μg pRLTK, 0.5 μg pSV-sport PPARγ2, and 0.5 μg pSV-sport RXRα plasmids. After overnight incubation, the transfected cells were treated with tested compounds' and incubated over 24 h at 37 °C and 5% CO_2.

Bioactivity readings were measured by luminescence quantifications using Dual-Glo Luciferase Assay System (Promega Corporation, USA) according to manufacturer's protocol. Briefly, 70μL Dual-Glo luciferase reagent was added to the media of transfected cells and incubated for 10 min and luminescence signals were read using Glomax® 96 microplate luminometer. Following that, 70 μL Stop and Glo reagent was added into the wells and incubated further for 10 min and luminescence signal were read again using the same protocol.

The fold changes of luciferase ratio caused by the tested hits were calculated using the following formula:

$$\text{Fold Changes of Luciferaseration} = \frac{\text{Luciferase Ratio}_{\text{HIT}} - \text{Luciferase Ration}_{\text{NTC}}}{\text{Luciferase Ratio}_{\text{DMSO}} - \text{Luciferase Ration}_{\text{NTC}}}$$

where, NTC = non-transfected cells. Statistical analysis were conducted using Minitab 15 software and statistical significance were determined using Analysis of Variance (ANOVA) tests with $p < 0.05$ considered significant.

Acknowledgements

This project was partly funded by Malaysian Ministry of Science, Technology and Innovation, Grant Scheme 311/IFN/69230111. The authors also thank the Deanship of Scientific Research and Hamdi-Mango Centre for Scientific Research at the University of Jordan for their generous funds.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ejmech.2011.03.040.

References

[1] L.M. Hall, N. Sattar, J.M. Gill, Risk of metabolic and vascular disease in South Asians: potential mechanisms for increased insulin resistance, Future Lipidology 3 (2008) 411–424.
[2] B.R. Johns, G.M. Reaven, PPAR-γ agonists, insulin resistance and dyslipidemia: not a simple relationship, Clinical Lipidology 5 (2010) 509–525.
[3] P. Arck, B. Toth, A. Pestka, U. Jeschke, Nuclear receptors of the peroxisome proliferator-activated receptor (PPAR) family in gestational diabetes: from animal models to clinical trials, Biology of Reproduction 83 (2010) 168–176.
[4] A. Hiukka, M. Maranghi, N. Matikainen, M.R. Taskinen, PPAR alpha: an emerging therapeutic target in diabetic microvascular damage, Nature Reviews Endocrinology 6 (2010) 454–463.
[5] A. Farce, N. Renault, P. Chavatte, Structural insight into PPAR γ ligands binding, Current Medicinal Chemistry 16 (2009) 1768–1789.
[6] R. Siersbaek, R. Nielsen, S. Mandrup, PPAR γ in adipocyte differentiation and metabolism – novel insights from genome-wide studies, FEBS Letters 584 (2010) 3242–3249.
[7] J.S. Bae, T.H. Kim, M.Y. Kim, J.M. Park, Y.H. Ahn, Transcriptional Regulation of glucose sensors in pancreatic beta-cells and liver: an update, Sensors 10 (2010) 5031–5053.
[8] J. Bassaganya-Riera, R. Song, P.C. Roberts, R. Hontecillas, PPAR-γ activation as an anti-inflammatory therapy for respiratory virus infections, Viral Immunology 23 (2010) 343–352.
[9] S.Z. Duan, M.G. Usher, R.M. Mortensen, PPARs: the vasculature, inflammation and hypertension, Current Opinion in Nephrology and Hypertension 18 (2009) 128–133.
[10] B.D. Abbott, Review of the expression of peroxisome proliferator-activated receptors alpha (PPARα), beta (PPARβ), and gamma (PPARγ) in rodent and human development, Reproductive Toxicology 27 (2009) 246–257.
[11] E.R. Pearson, Pharmacogenetics and future strategies in treating hyperglycaemia in diabetes, Frontiers in Bioscience 14 (2009) 4348–4362.
[12] B.R. Henke, Peroxisome proliferator-activated receptor gamma (PPAR γ) ligands and their therapeutic utility, in: Progress in Medicinal Chemistry. Elsevier, 2004, pp. 1–53.
[13] J. Choi, Y. Ko, H.S. Lee, Y.S. Park, Y. Yang, S. Yoon, Identification of (β-carboxyethyl)-rhodanine derivatives exhibiting peroxisome proliferator-activated receptor γ activity, European Journal of Medicinal Chemistry 45 (2010) 193–202.
[14] J.H. Ahn, M.S. Shin, S.H. Jung, J.A. Kim, H.M. Kim, S.H. Kim, S.K. Kang, K.R. Kim, S.D. Rhee, S.D. Park, J.M. Lee, J.H. Lee, H.G. Cheon, S.S. Kim, Synthesis and structure–activity relationship of novel indene N-oxide derivatives as potent peroxisome proliferator activated receptor(PPARγ) agonists, Bioorganic & Medicinal Chemistry Letters 17 (2007) 5239–5244.
[15] Y. Lamotte, P. Martres, N. Faucher, A. Laroze, D. Grillot, N. Ancellin, Y. Saintillan, V. Beneton, R.T. Gampe Jr., Synthesis and biological activities of novel indole derivatives as potent and selective PPAR γ modulators, Bioorganic & Medicinal Chemistry Letters 20 (2010) 1399–1404.
[16] S.D. Debenham, A. Chan, F.W. Lau, W. Liu, H.B. Wood, K. Lemme, L. Colwell, B. Habulihaz, T.E. Akiyama, M. Einstein, T.W. Doebber, N. Sharma, C.F. Wang, M. Wu, J.P. Berger, P.T. Meinke, Highly functionalized 7-azaindoles as selective PPARγ modulators, Bioorganic & Medicinal Chemistry Letters 18 (2008) 4798–4801.
[17] X.-Y. Ye, S. Chen, H. Zhang, K.T. Locke, K. O'Malley, L. Zhang, R. Srivastava, B. Miao, D. Meyers, H. Monshizadegan, D. Search, D. Grimm, R. Zhang, J. Lippy, C. Twamley, J.K. Muckelbauer, C. Chang, Y. An, V. Hosagrahara, L. Zhang, T.J. Yang, R. Mukherjee, P.T.W. Cheng, J.A. Tino, Synthesis and structure-activity relationships of 2-aryl-4-oxazolylmethoxy benzylglycines and 2-aryl-4-thiazolylmethoxy benzylglycines as novel, potent PPAR α selective activators- PPAR α and PPAR γ selectivity modulation, Bioorganic & Medicinal Chemistry Letters 20 (2010) 2933–2937.
[18] R.C. Desai, W. Han, E.J. Metzger, J.P. Bergman, D.F. Gratale, K.L. MacNaul, J.P. Berger, T.W. Doebber, K. Leung, D.E. Moller, J.V. Heck, S.P. Sahoo, 5-Aryl thiazolidine-2,4-diones: discovery of PPAR dual α/γ agonists as antidiabetic agents, Bioorganic & Medicinal Chemistry Letters 13 (2003) 2795–2798.
[19] R.C. Desai, D.F. Gratale, W. Han, H. Koyama, E. Metzger, V.K. Lombardo, K.L. MacNaul, T.W. Doebber, J.P. Berger, K. Leung, R. Franklin, D.E. Moller, J.V. Heck, S.P. Sahoo, Aryloxazolidinediones: identification of potent orally active PPAR dual α/γ agonists, Bioorganic & Medicinal Chemistry Letters 13 (2003) 3541–3544.
[20] C. Santini, G.D. Berger, W. Han, R. Mosley, K. MacNaul, J. Berger, T. Doebber, M. Wu, D.E. Moller, R.L. Tolman, S.P. Sahoo, Phenylacetic acid derivatives as hPPAR agonists, Bioorganic & Medicinal Chemistry Letters 13 (2003) 1277–1280.

[21] T. Kaya, S.C. Mohr, D.J. Waxman, S. Vajda, Computational screening of phthalate monoesters for binding to PPARγ, Chemical Research in Toxicology 19 (2006) 999–1009.
[22] M. Scarsi, M. Podvinec, A. Roth, Sulfonylureas and glinides exhibit peroxisome proliferator-activated receptor γ activity: a combined virtual screening and biological assay approach, Molecular Pharmacology 71 (2007) 398–406.
[23] J. Choi, Y. Ko, H.S. Lee, Y.S. Park, Y. Yang, S. Yoon, Structure based drug design studies on heteroaryl propanoic acid derivatives as PPARγ agonists, European Journal of Medicinal Chemistry 45 (2010) 193–202.
[24] M. Hieke, J. Ness, R. Steri, M. Dittrich, C. Greiner, O. Werz, K. Baumann, M. Schubert-Zsilavecz, S. Weggen, H. Zettl, Design, synthesis, and biological evaluation of a novel class of γ-secretase modulators with PPARγ activity, Journal of Medicinal Chemistry 53 (2010) 4691–4700.
[25] M. Goebel, M. Clemenz, B. Staels, T. Unger, U. Kintscher, R. Gust, Characterization of new PPARγ agonists: analysis of telmisartan's structural components, ChemMedChem 4 (2009) 445–456.
[26] X. Dong, Z. Zhang, R. Wen, J. Shen, X. Shen, H. Jiang, Structure-based de novo design, synthesis, and biological evaluation of the indole-based PPAR γ ligands (I), Bioorganic & Medicinal Chemistry Letters 16 (2006) 5913–5916.
[27] B. Kuhn, H. Hilpert, J. Benz, A. Binggeli, U. Grether, R. Humm, H.P. Märki, M. Meyer, P. Mohr, Structure-based design of indole propionic acids as novel PPARα/γ co-agonists, Bioorganic & Medicinal Chemistry Letters 16 (2006) 4016–4020.
[28] Y. Iwata, S. Miyamoto, M. Takamura, H. Yanagisawa, A. Kasuya, Interaction between peroxisome proliferator-activated receptor γ and its agonists: docking study of oximes having 5-benzyl-2,4-thiazolidinedione, Journal of Molecular Graphics & Modelling 19 (2001) 536–542.
[29] C. Liao, B. Liu, L. Shi, J. Zhou, X.-P. Lu, Construction of a virtual combinatorial library using SMILES strings to discover potential structure-diverse PPAR modulators, European Journal of Medicinal Chemistry 40 (2005) 632–640.
[30] N. Mahindroo, C.-F. Huang, Y.-H. Peng, C.-C. Wang, C.-C. Liao, T.-W. Lien, S.K. Chittimalla, W.-J. Huang, C.-H. Chai, E. Prakash, C.-P. Chen, T.-A. Hsu, C.-H. Peng, I.L. Lu, L.-H. Lee, Y.-W. Chang, W.-C. Chen, Y.-C. Chou, C.-T. Chen, C.M.V. Goparaju, Y.-S. Chen, S.-J. Lan, M.-C. Yu, X. Chen, Y.-S. Chao, S.-Y. Wu, H.-P. Hsieh, Novel indole-based peroxisome proliferator-activated receptor agonists: design, SAR, structural biology, and biological activities, Journal of Medicinal Chemistry 48 (2005) 8194–8208.
[31] Y. Tanrikulu, O. Rau, O. Schwarz, E. Proschak, K. Siems, L. Müller-Kuhrt, M. Schubert-Zsilavecz, G. Schneider, Structure-based pharmacophore screening for natural-product-derived PPARγ agonists, ChemBioChem 10 (2009) 75–78.
[32] I.L. Lu, C.-F. Huang, Y.-H. Peng, Y.-T. Lin, H.-P. Hsieh, C.-T. Chen, T.-W. Lien, H.-J. Lee, N. Mahindroo, E. Prakash, A. Yueh, H.-Y. Chen, C.M.V. Goparaju, X. Chen, C.-C. Liao, Y.-S. Chao, J.T.A. Hsu, S.-Y. Wu, Structure-based drug design of a novel family of PPARγ partial agonists: virtual screening, X-ray crystallography, and in vitro/in vivo biological activities, Journal of Medicinal Chemistry 49 (2006) 2703–2712.
[33] X.Y. Xu, F. Cheng, J.H. Shen, X.M. Luo, L.L. Chen, L.D. Yue, Y. Du, F. Ye, S.H. Jiang, D.Y. Zhu, H.L. Jiang, K.X. Chen, Agonist-PPAR γ interactions: molecular modeling study with docking approach, International Journal of Quantum Chemistry 93 (2003) 405–410.
[34] S. Sundriyal, P.V. Bharatam, Important pharmacophoric features of pan PPAR agonists: common chemical feature analysis and virtual screening, European Journal of Medicinal Chemistry 44 (2009) 3488–3495.
[35] L. Rathi, S.K. Kashaw, A. Dixit, G. Pandey, A.K. Saxena, Pharmacophore identification and 3D-QSAR studies in N-(2-benzoyl phenyl)-L-tyrosines as PPAR g agonists, Bioorganic & Medicinal Chemistry 12 (2004) 63–69.
[36] C.Z. Liao, A.H. Xie, J.J. Zhou, L.M. Shi, Z.B. Li, X.P. Lu, 3D QSAR studies on peroxisome proliferator-activated receptor gamma agonists using CoMFA and CoMSIA, Journal of Molecular Modeling 10 (2004) 165–177.
[37] N.K. Salam, T.H.-W. Huang, B.P. Kota, M.S. Kim, Y. Li, D.E. Hibbs, Novel PPAR-γ agonists identified from a natural product library: a virtual screening, induced fit docking and biological assay study, Chemical Biology and Drug Design 71 (2008) 57–70.
[38] A. Furukawa, T. Arita, S. Satoh, K. Wakabayashi, S. Hayashi, Y. Matsui, K. Araki, M. Kuroha, J. Ohsumi, Discovery of a novel selective PPAR γ modulator from (-)-cercosporamide derivatives, Bioorganic & Medicinal Chemistry Letters 20 (2010) 2095–2098.
[39] Y. Li, Z. Wang, N. Furukawa, P. Escaron, J. Weiszmann, G. Lee, M. Lindstrom, J. Liu, X. Liu, H. Xu, O. Plotnikova, V. Prasad, N. Walker, R.M. Learned, J.-L. Chen, T2384, a novel antidiabetic agent with unique peroxisome proliferator-activated receptor γ binding properties, Journal of Biological Chemistry 283 (2008) 9168–9176.
[40] G. Fracchiolla, A. Laghezza, L. Piemontese, P. Tortorella, F. Mazza, R. Montanari, G. Pochetti, A. Lavecchia, E. Novellino, S. Pierno, D. Conte Camerino, F. Loiodice, New 2-aryloxy-3-phenyl-propanoic acids as peroxisome proliferator-activated receptors α/γ dual agonists with improved potency and reduced adverse effects on skeletal muscle function, Journal of Medicinal Chemistry 52 (2009) 6382–6393.
[41] A. Motani, Z. Wang, J. Weiszmann, L.R. McGee, G. Lee, Q. Liu, J. Staunton, Z. Fang, H. Fuentes, M. Lindstrom, J. Liu, D.H.T. Biermann, J. Jaen, N.P.C. Walker, R.M. Learned, J.-L. Chen, Y. Li, INT131: a selective modulator of PPAR γ, Journal of Molecular Biology 386 (2009) 1301–1311.
[42] T. Waku, T. Shiraki, T. Oyama, Y. Fujimoto, K. Maebara, N. Kamiya, H. Jingami, K. Morikawa, Structural insight into PPAR γ activation through covalent modification with endogenous fatty acids, Journal of Molecular Biology 385 (2009) 188–199.
[43] D.R. Artis, J.J. Lin, C. Zhang, W. Wang, U. Mehra, M. Perreault, D. Erbe, H.I. Krupka, B.P. England, J. Arnold, A.N. Plotnikov, A. Marimuthu, H. Nguyen, S. Will, M. Signaevsky, J. Kral, J. Cantwell, C. Settachatgull, D.S. Yan, D. Fong, A. Oh, S. Shi, P. Womack, B. Powell, G. Habets, B.L. West, K.Y.J. Zhang, M.V. Milburn, G.P. Vlasuk, K.P. Hirth, K. Nolop, G. Bollag, P.N. Ibrahim, J.F. Tobin, Scaffold-based discovery of indeglitazar, a PPAR pan-active anti-diabetic agent, Proceedings of the National Academy of Sciences 106 (2009) 262–267.
[44] R. Montanari, F. Saccoccia, E. Scotti, M. Crestani, C. Godio, F. Gilardi, F. Loiodice, G. Fracchiolla, A. Laghezza, P. Tortorella, A. Lavecchia, E. Novellino, F. Mazza, M. Aschi, G. Pochetti, Crystal structure of the peroxisome proliferator-activated receptor γ (PPARγ) ligand binding domain complexed with a novel partial agonist: a new region of the hydrophobic pocket could be exploited for drug design, Journal of Medicinal Chemistry 51 (2008) 7768–7776.
[45] V. Chandra, P. Huang, Y. Hamuro, S. Raghuram, Y. Wang, T.P. Burris, F. Rastinejad, Structure of the intact PPAR-γ-RXR- nuclear receptor complex on DNA, Nature (2008) 350–356.
[46] M. Einstein, T.E. Akiyama, G.A. Castriota, C.F. Wang, B. McKeever, R.T. Mosley, J.W. Becker, D.E. Moller, P.T. Meinke, H.B. Wood, J.P. Berger, The differential interactions of peroxisome proliferator-activated receptor γ ligands with Tyr473 is a physical basis for their unique biological activities, Molecular Pharmacology 73 (2008) 62–74.
[47] Y. Li, M. Choi, K. Suino, A. Kovach, J. Daugherty, S.A. Kliewer, H.E. Xu, Structural and biochemical basis for selective repression of the orphan nuclear receptor liver receptor homolog 1 by small heterodimer partner, Proceedings of the National Academy of Sciences of the United States of America 102 (2005) 9505–9510.
[48] N.R.A. Beeley, C. Sage, GPCRs: an update on structural approaches to drug discovery, Targets 2 (2003) 19–25.
[49] G. Klebe, Virtual ligand screening: strategies, perspectives and limitations, Drug Discovery Today 11 (2006) 580–594.
[50] H. Steuber, M. Zentgraf, C. Gerlach, C.A. Sotriffer, A. Heine, G. Klebe, Expect the unexpected or caveat for drug designers: multiple structure determinations using aldose reductase crystals treated under varying soaking and co-crystallisation conditions, Journal of Molecular Biology 363 (2006) 174–187.
[51] M.T. Stubbs, S. Reyda, F. Dullweber, M. Möller, G. Klebe, D. Dorsch, W.W.K.R. Mederski, H. Wurziger, pH-dependent binding modes observed in trypsin crystals: lessons for structure-based drug design, ChemBioChem 3 (2002) 246–249.
[52] M.A. DePristo, P.I.W. de Bakker, T.L. Blundell, Heterogeneity and inaccuracy in protein structures solved by X-ray crystallography, Structure 12 (2004) 831–838.
[53] B.C. Kallenberger, J.D. Love, V.K.K. Chatterjee, J.W.R. Schwabe, A dynamic mechanism of nuclear receptor activation and its perturbation in a human disease, Nature Structural Biology 10 (2003) 136–140.
[54] M.O. Taha, Y. Bustanji, M.A.S. Al-Ghussein, M. Mohammad, H. Zalloum, I.M. Al-Masri, N. Atallah, Pharmacophore modeling, quantitative structure–activity relationship analysis, and in silico screening reveal potent glycogen synthase kinase-3β inhibitory activities for cimetidine, hydroxychloroquine, and gemifloxacin, Journal of Medicinal Chemistry 51 (2008) 2062–2077.
[55] M.O. Taha, L.A. Dahabiyeh, Y. Bustanji, H. Zalloum, S. Saleh, Combining ligand-based pharmacophore modeling, quantitative structure–activity relationship analysis and in silico screening for the discovery of new potent hormone sensitive lipase inhibitors, Journal of Medicinal Chemistry 51 (2008) 6478–6494.
[56] M.O. Taha, N. Atallah, A.G. Al-Bakri, C. Paradis-Bleau, H. Zalloum, K.S. Younis, R.C. Levesque, Discovery of new MurF inhibitors via pharmacophore modeling and QSAR analysis followed by in-silico screening, Bioorganic & Medicinal Chemistry 16 (2008) 1218–1235.
[57] M.O. Taha, Y. Bustanji, A.G. Al-Bakri, A.-M. Yousef, W.A. Zalloum, I.M. Al-Masri, N. Atallah, Discovery of new potent human protein tyrosine phosphatase inhibitors via pharmacophore and QSAR analysis followed by in silico screening, Journal of Molecular Graphics and Modelling 25 (2007) 870–884.
[58] A.M. Abu Hammad, M.O. Taha, Pharmacophore modeling, quantitative structure–activity relationship analysis, and shape-complemented in silico screening allow access to novel influenza neuraminidase inhibitors, Journal of Chemical Information and Modeling 49 (2009) 978–996.
[59] M.O. Taha, M. Tarairah, H. Zalloum, G. Abu-Sheikha, Pharmacophore and QSAR modeling of estrogen receptor beta ligands and subsequent validation and in silico search for new hits, Journal of Molecular Graphics & Modelling 28 (2010) 383–400.
[60] R. Abu Khalaf, G. Abu Sheikha, Y. Bustanji, M.O. Taha, Discovery of new cholesteryl ester transfer protein inhibitors via ligand-based pharmacophore modeling and QSAR analysis followed by synthetic exploration, European Journal of Medicinal Chemistry 45 (2010) 1598–1617.
[61] A. Al-Nadaf, G. Abu Sheikha, M.O. Taha, Elaborate ligand-based pharmacophore exploration and QSAR analysis guide the synthesis of novel pyridinium-based potent beta-secretase inhibitory leads, Bioorganic & Medicinal Chemistry 18 (2010) 3088–3115.
[62] M.A. Al-Sha'er, M.O. Taha, Discovery of novel CDK1 inhibitors by combining pharmacophore modeling, QSAR analysis and in silico screening followed by in vitro bioassay, European Journal of Medicinal Chemistry 45 (2010) 4316–4330.
[63] M.A. Al-Sha'er, M.O. Taha, Elaborate ligand-based modeling reveals new nanomolar heat shock protein 90α inhibitors, Journal of Chemical Information and Modeling 50 (2010) 1706–1723.

Please cite this article in press as: B.O. Al-Najjar, et al., Discovery of new nanomolar peroxisome proliferator-activated receptor γ activators via elaborate ligand-based modeling, European Journal of Medicinal Chemistry (2011), doi:10.1016/j.ejmech.2011.03.040

[64] Discovery Studio Version 2.5 (DS 2.5) User Manual. Accelrys Inc, San Diego, CA, 2009.
[65] M.O. Taha, A.G. Al-Bakri, W.A. Zalloum, Discovery of potent inhibitors of pseudomonal quorum sensing via pharmacophore modeling and in silico screening, Bioorganic & Medicinal Chemistry Letters 16 (2006) 5902–5906.
[66] K. Liu, R.M. Black, J.J. Acton Iii, R. Mosley, S. Debenham, R. Abola, M. Yang, R. Tschirret-Guth, L. Colwell, C. Liu, M. Wu, C.F. Wang, K.L. MacNaul, M.E. McCann, D.E. Moller, J.P. Berger, P.T. Meinke, A.B. Jones, H.B. Wood, Selective PPAR γ modulators with improved pharmacological profiles, Bioorganic & Medicinal Chemistry Letters 15 (2005) 2437–2440.
[67] R.C. Desai, E. Metzger, C. Santini, P.T. Meinke, J.V. Heck, J.P. Berger, K.L. MacNaul, T.-Q. Cai, S.D. Wright, A. Agrawal, D.E. Moller, S.P. Sahoo, Design and synthesis of potent and subtype-selective PPARα agonists, Bioorganic & Medicinal Chemistry Letters 16 (2006) 1673–1678.
[68] H. Li, J. Sutter, R. Hoffman, HypoGen: an automated system for generating 3D predictive pharmacophore models. in: O. Guner (Ed.), Pharmacophore Perception, Development and Use in Drug Design. International University Line, 2000, pp. 173–189.
[69] H. Li, J. Sutter, R. Hoffman, O. Guner, Effect of variable weights and tolerances on predictive model generation. in: O. Guner (Ed.), Pharmacophore Perception, Development and Use in Drug Design. International University Line, 2000, pp. 501–511.
[70] K. Poptodorov, T. Luu, T. Langer, R. Hoffmann, Pharmacophores and Pharmacophores Searches, second ed. Wiley-VCH, Weinheim, Germany, 2006.
[71] Y. Kurogi, O.F. Guner, Pharmacophore modeling and three-dimensional database searching for drug design using catalyst, Current Medicinal Chemistry 8 (2001) 1035–1055.
[72] I. Bersuker, S. Bahçeci, J. Boggs, The electron-conformational method of identification of pharmacophore and anti pharmacophore shielding. in: O. Guner (Ed.), Pharmacophore Perception, Development and Use in Drug Design. International University Line, 2000, pp. 457–473.
[73] R. Fischer, The Principle of Experimentation Illustrated by a Psycho-Physical ExpeHafner Publishing Co, eighth ed. Hafner Publishing, New York, 1966.
[74] I. Al-masri, M. Mohammad, M. Taha, Discovery of DPP IV inhibitors by pharmacophore modeling and QSAR analysis followed by in silico screening, ChemMedChem 3 (2008) 1763–1779.
[75] F.L. Ramsey, D.W. Schafer, The Statistical Sleuth: a Course in Methods of Data Analysis, first ed. Wadsworth Publishing Company, Belmont, CA, 1997.
[76] M.L. Verdonk, V. Berdini, M.J. Hartshorn, W.T.M. Mooij, C.W. Murray, R.D. Taylor, P. Watson, Virtual screening using protein–ligand docking: avoiding artificial enrichment, Journal of Chemical Information and Computer Sciences 44 (2004) 793–806.
[77] J. Kirchmair, P. Markt, S. Distinto, G. Wolber, T. Langer, Evaluation of the performance of 3D virtual screening protocols: RMSD comparisons, enrichment assessments, and decoy selection – what can we learn from earlier mistakes? Journal of Computer-Aided Molecular Design 22 (2008) 213–228.
[78] J.J. Irwin, B.K. Shoichet, ZINC – a free database of commercially available compounds for virtual screening, Journal of Chemical Information and Modeling 45 (2004) 177–182.
[79] N. Triballeau, F. Acher, I. Brabet, J.-P. Pin, H.-O. Bertrand, Virtual screening workflow development guided by the "receiver operating characteristic" curve approach. Application to high-throughput docking on metabotropic glutamate receptor subtype 4, Journal of Medicinal Chemistry 48 (2005) 2534–2547.
[80] C.A. Lipinski, F. Lombardo, B.W. Dominy, P.J. Feeney, Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings, Advanced Drug Delivery Reviews 46 (2001) 3–26.
[81] D.F. Veber, S.R. Johnson, H.-Y. Cheng, B.R. Smith, K.W. Ward, K.D. Kopple, Molecular properties that influence the oral bioavailability of drug candidates, Journal of Medicinal Chemistry 45 (2002) 2615–2623.
[82] D. Debendranath, M. Satya, N. Partha, G. Maya, C. Jin, G. Coleman, D.S. Somesh, M.R. Gerald, N. Bishwajit, A novel peroxisome proliferator-activated gamma (PPARγ) agonist, CLX-0921, has potent antihyperglycemic activity with low adipogenic potential, Metabolism: Clinical and Experimental 52 (2003) 1012–1018.
[83] P. Neogi, F.J. Lakner, S. Medicherla, J. Cheng, D. Dey, M. Gowri, B. Nag, S.D. Sharma, L.B. Pickford, C. Gross, Synthesis and structure–activity relationship studies of cinnamic acid-based novel thiazolidinedione antihyperglycemic agents, Bioorganic & Medicinal Chemistry 11 (2003) 4059–4067.
[84] S.-K. Lin, Pharmacophore perception, development and use in drug design. in: Osman F. Güner (Ed.), Molecules, 5, 2000, pp. 987–989.
[85] G. Osman, C. Omoshile, K. Yasuhisa, Pharmacophore modeling and three dimensional database searching for drug design using catalyst: recent advances, Current Medicinal Chemistry 11 (2004) 2991–3005.
[86] E.M. Krovat, T. Langer, Impact of scoring functions on enrichment in docking-based virtual screening: an application study on renin inhibitors, Journal of Chemical Information and Computer Sciences 44 (2004) 1123–1129.
[87] M. Jacobsson, P. Lidén, E. Stjernschantz, H. Boström, U. Norinder, Improving structure-based virtual screening by multivariate analysis of scoring data, Journal of Medicinal Chemistry 46 (2003) 5781–5789.
[88] H. Gao, C. Williams, P. Labute, J. Bajorath, Binary quantitative structure–activity relationship (QSAR) analysis of estrogen receptor ligands, Journal of Chemical Information and Computer Sciences 39 (1998) 164–168.
[89] S. Park, I.S. Ahn, J.H. Kim, M.R. Lee, J.S. Kim, H.J. Kim, Glyceollins, one of the phytoalexins derived from soybeans under fungal stress, enhance insulin sensitivity and exert insulinotropic actions, Journal of Agricultural and Food Chemistry 58 (2010) 1551–1557.

Please cite this article in press as: B.O. Al-Najjar, et al., Discovery of new nanomolar peroxisome proliferator-activated receptor γ activators via elaborate ligand-based modeling, European Journal of Medicinal Chemistry (2011), doi:10.1016/j.ejmech.2011.03.040