Bioorganic & Medicinal Chemistry 20 (2012) 3523–3532

Contents lists available at SciVerse ScienceDirect

Bioorganic & Medicinal Chemistry

journal homepage: www.elsevier.com/locate/bmc

Activity landscape modeling of PPAR ligands with dual-activity difference maps

Oscar Méndez-Lucio [a], Jaime Pérez-Villanueva [b], Rafael Castillo [a,*], José L. Medina-Franco [c,*]

[a] Facultad de Química, Departamento de Farmacia, Universidad Nacional Autónoma de México, México DF 04510, Mexico
[b] Departamento de Sistemas Biológicos, División de Ciencias Biológicas y de la Salud, UAM-X, México DF 04960, Mexico
[c] Torrey Pines Institute of Molecular Studies, 11350 SW Village Parkway, Port St. Lucie, FL 34987, USA

ARTICLE INFO

Article history:
Received 15 February 2012
Revised 27 March 2012
Accepted 4 April 2012
Available online 19 April 2012

Keywords:
Activity cliffs
Diabetes mellitus
Dual-activity difference maps
PPAR agonist
Structure–activity relationships

ABSTRACT

Activation of peroxisome proliferator-activated receptor (PPAR) subtypes offers a promising strategy for the treatment of diabetes mellitus and metabolic diseases. Selective and dual PPAR agonists have been developed and the systematic characterization of their structure–activity relationships (SAR) is of major significance. Herein, we report a systematic description of the SAR of 168 compounds screened against the three PPAR subtypes using the principles of activity landscape modeling. As part of our effort to develop and apply chemoinformatic tools to navigate through activity landscapes, we employed consensus dual-activity difference maps recently reported. The analysis is based on pairwise relationships of potency difference and structure-similarity which were calculated from the combination of four different 2D and 3D structure representations. Dual-activity difference maps uncovered regions in the landscape with similar SAR for two or three receptor subtypes as well as regions with inverse SAR, that is, changes in structure that increase activity for one subtype but decrease activity for the other subtype. Analysis of pairs of compounds with high structure similarity revealed the presence of single-, dual-, and 'pan-receptor' activity cliffs, that is, small changes in structure with high changes in potency for one, two, or three receptor subtypes, respectively. Single-, dual-, and pan-receptor scaffold hops are also discussed. The analysis of the chemical structures of selected data points reported in this paper points to specific structural features that are helpful for the design of new PPAR agonists. The approach presented in this work is general and can be extended to analyze larger data sets.

1. Introduction

Diabetes mellitus (DM) is a global health problem. Estimates indicate that in 2010 there were 285 million people worldwide with diabetes and that in 2030 there will be 439 million adults with diabetes.[1] Different molecular targets have been proposed in order to modulate important aspects of DM pathogenesis. Specially, the activation of peroxisome proliferator-activated receptors (PPARs) offers a promising strategy for the treatment of DM and metabolic diseases.[2]

PPARs are nuclear lipid-activated transcription factors that have been identified as major regulators in glucose and lipid metabolism; thereby they contribute significantly to some disorders such as diabetes, obesity and cardiovascular complications.[3] The expression of three different PPAR subtypes, PPARα, PPARδ and PPARγ has been described.[3,4] Each subtype shows a different tissue distribution pattern, and pharmacological response.[3] For example, PPARα is amply expressed in the liver, kidney, heart, and skeletal muscle. Experimental data indicates that this subtype has a critical role in the uptake and oxidation of fatty acids and also in lipoprotein metabolism. Activation of PPARα decreases serum triglycerides and increases serum high-density lipoprotein cholesterol, which together restores lipid and glucose homeostasis. PPARγ is predominantly expressed in adipose tissue and is a key modulator of adipocyte differentiation. Activation of PPARγ improves glycemic control by increasing insulin sensitivity. PPARδ also seems to play an important role in the regulation of lipid metabolism and cholesterol efflux.[5–7] Selective, dual and pan PPAR agonists have been studied as potential therapy against diabetes and the metabolic syndrome.[3,5,6] Progress in the development of some dual PPARα/γ agonists has been hampered because of safety issues during phase II and phase III clinical trials.[8] Therefore, the systematic study of the structural requirements to determine selective compounds toward one PPAR subtype or for the design of new dual PPAR agonists is of great significance.[5–7] Computational studies, such as pharmacophore modeling[9,10] have been reported directed towards the rationalization of the SAR of PPAR ligands.

Abbreviations: DAD, dual activity difference; DM, diabetes mellitus; EStateIndices, electrotopological state indices; PPAR, peroxisome-proliferator-activated receptor; ROCS, Rapid Overlay of Chemical Structures; SAR, structure–activity relationships; TAD, triple activity difference.

* Corresponding authors. Tel.: +1 772 345 4685; fax: +1 772 345 3649 (J.L.M.-F.).
E-mail address: jmedina@tpims.org (J.L. Medina-Franco).

0968-0896/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.bmc.2012.04.005

Systematic description of the structure–activity relationships (SAR) of data sets can be conveniently achieved using the principles and methods of activity landscape modeling.[11–13] The activity landscape has been defined as 'any representation that integrates the analysis of the structural similarity and potency differences between compounds sharing the same biological activity'.[12] The activity landscape has been compared to rolling hills or continuous SAR where small changes in molecular structure are associated with small changes in activity.[14] A discontinuous SAR or rugged activity landscape, however, is populated with molecules with small changes in structure but large changes in activity, that is, 'activity cliffs' (chemical compounds with highly similar structures but significantly different biological activities).[14] Understanding the activity landscape and early detection of activity cliffs can be crucial to the success of predictive computational models such as QSAR.[15–17] However, understanding the activity landscape of data sets tested with one or multiple biological endpoints is not straightforward because the landscape can be highly complex with a combination of smooth and rugged regions.[18] Another major challenge is the dependence of chemical space on molecular representation.[19,20]

Herein, we report a systematic characterization of the SAR of 168 compounds screened against the three PPAR subtypes. The analysis of the SAR was based on pairwise comparisons of the activity similarity and structure similarity. The activity landscape across multiple PPAR subtypes, that is, 'multi-target activity landscape modeling',[21,22] was characterized using the dual and triple activity difference (DAD/TAD) maps we recently developed.[23,24] As pointed out by Peltason and co-workers, multi-target activity landscape modeling is attractive to characterize the SAR of compound data sets associated with selectivity[22] or promiscuity, for example, 'scaffold-based promiscuity' of the targets. Single-, dual-, and triple-receptor subtype activity cliffs as well as scaffold hops (chemical compounds with highly dissimilar structures but significantly similar biological activities)[25] were identified.

2. Methods

2.1. Data set

For this study, a set of 168 unique PPAR agonists was obtained from the Binding Database,[26,27] which is a public repository of chemical compounds annotated with biological activity well suited to perform SAR studies.[28] All the 168 compounds have IC_{50} values reported against the three receptor subtypes. The SMILES representation of the structures and the pIC_{50} ($-\log IC_{50}$) values are listed in the Supplementary data (Table S1). For PPARα, the pIC_{50} values range from 4.36 to 8.32, for PPARδ from 3.40 to 8.76 and for PPARγ from 4.39 to 8.30. It is noteworthy that the activity range for the three subtypes is very similar, thus facilitating the cross-comparisons of the activity landscapes.[29]

2.2. Structure similarity

Eleven different 2D fingerprint representations were calculated for each molecule using the open-source software MayaChemTools.[30] This software is a set of Perl scripts, modules and classes for the generation of fingerprints corresponding to atom neighborhoods,[31] atom types, electrotopological state indices (EStateIndices),[32] extended connectivity,[33] MACCS (322 bits),[34] path length, topological atom pairs (TopAtomPairs),[35] topological atom torsions (TopAtomTorsions),[36] topological atom triplets (TopAtomTriplets), topological pharmacophore atom pairs (TopPh4Pairs),[37] and topological pharmacophore atom triplets (TopPh4Triplets).[38] All of them are widely used and validated molecular descriptors.

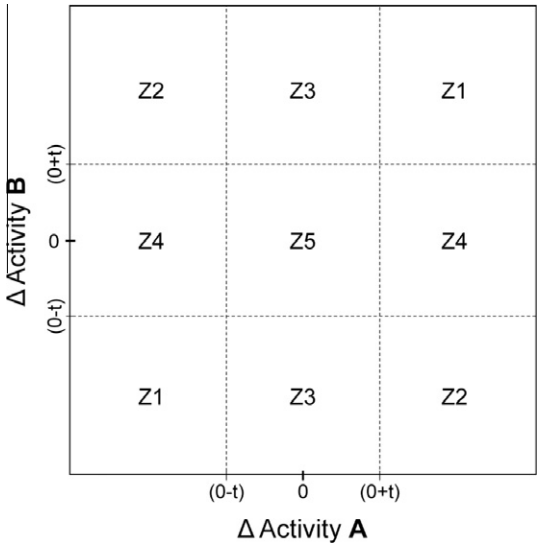


Figure 1. Prototype dual activity-difference (DAD) map divided in five major regions labeled as Z1–Z5. Region Z1 contains pairs in which structural changes affect in similar proportion the activity against both receptors. Pairs in Z2 are those in which the structural change increases the activity for one target, but decreases the activity toward the other (inverse SAR). Pairs in which a structural difference does not affect the activity against one target, but it does against the other, are located in region Z3 or Z4. Region Z5 is associated with pairs with similar activities against both targets.

Additionally, ComboScore and ShapeTanimoto, as implemented in Rapid Overlay of Chemical Structures (ROCS) (OpenEye Scientific Software) were computed as 3D fingerprints using a single low-energy conformation calculated with the MMFF94x force field. The Combo Score was scaled by dividing the 3D similarity values by two, in order to obtain a number within the same scale as the other computed similarity measures. Recent works have shown that the use of 3D fingerprints are important in characterizing activity landscapes, even if the use of a single conformer to represent 3D structures is an approximation of the real active conformation.[39–41] However, activity landscape modeling can also be carried out using multiple 3D conformers as we have previously shown.[29] Pairwise similarity was calculated using the Tanimoto coefficient[42,43] but other similarity measures can be used. For each fingerprint representation, 14,028 pairwise similarities were calculated for the 168 compounds. Similarity values calculated with EStateIndices, TopAtomPairs, TopPh4Triplets, and ComboScore were combined through mean fusion, that is, computing the mean similarity of the four selected fingerprints (see below). Other fusion rules can be applied.

2.3. Pairwise activity relationships

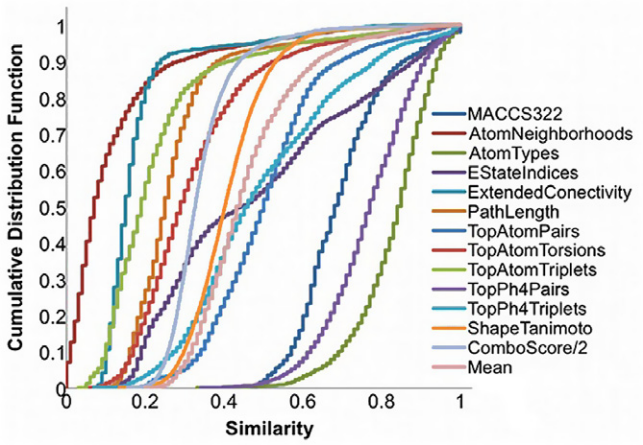
For each pair of PPAR agonists, potency differences were calculated using the expression:

$$\Delta pIC_{50}(R)_{i,j} = pIC_{50}(R)_i - pIC_{50}(R)_j$$

where $pIC_{50}(R)_i$ and $pIC_{50}(R)_j$ are the activities of the ith and jth molecules ($j > i$) against receptor subtype R. In this paper R = PPARα, δ, or γ. The positive or negative values of the expression above provided information concerning the direction of SAR.[24]

2.4. Dual- and triple-receptor activity landscapes

We recently developed DAD maps to characterize activity landscapes of data sets with biological activity for two biological endpoints.[23,44] These maps are based on pairwise activity differences



	Max	Q3	Median	Q1	Min	Mean	STD
MACCS322	1.00	0.76	0.68	0.62	0.45	0.70	0.11
Atom Neighborhoods	1.00	0.15	0.07	0.03	0.00	0.12	0.15
Atom Types	1.00	0.90	0.84	0.77	0.33	0.83	0.10
EStateIndices	1.00	0.68	0.45	0.26	0.06	0.49	0.25
ExtendedConnectivity	1.00	0.19	0.15	0.12	0.07	0.18	0.12
PathLength	1.00	0.31	0.25	0.20	0.10	0.28	0.14
TopAtomPairs	1.00	0.58	0.50	0.41	0.13	0.51	0.15
TopAtomTorsions	1.00	0.39	0.29	0.22	0.05	0.33	0.16
TopAtomTriplets	1.00	0.26	0.19	0.13	0.03	0.23	0.16
TopPh4Pairs	1.00	0.85	0.77	0.69	0.34	0.77	0.11
TopPh4Triplets	1.00	0.63	0.46	0.35	0.07	0.50	0.19
ShapeTanimoto	1.00	0.46	0.40	0.35	0.13	0.41	0.10
ComboScore/2	1.00	0.37	0.33	0.29	0.16	0.34	0.09
Mean	1.00	0.52	0.43	0.37	0.20	0.46	0.13

Figure 2. Cumulative distribution functions (CDF) of 14,028 pairwise structural similarities using different 2D and 3D structure representations. The table summarizes the information of the distributions. Q3 and Q1 indicate the third and first quartile, respectively.

against two receptors and are built plotting the activity difference for one receptor against the activity difference for the second receptor as depicted in the prototype DAD map in Figure 1. Vertical and horizontal lines at $\Delta IC_{50} \pm t$ determine boundaries for low/high potency difference for receptors **A** and **B**, respectively. Herein, we set $t = 1$ (1 log unit) so that data points were considered with low potency difference if $-1 \leqslant \Delta pIC_{50} \leqslant 1$ for each receptor. The boundaries give rise to five general zones, Z1–Z5.[44] If the structural changes between two compounds in the pair affect in similar proportion the biological activity against both **A** and **B** receptors (either an increase or decrease), this pair will be located in Z1. Pairs of compounds located in Z2 are those in which the structural change increases the activity toward one receptor, but decreases the activity toward the other; therefore Z2 is associated with an 'inverse SAR' and is the most important zone for selectivity studies. Pairs of compounds located in Z3 or Z4 denote compounds in which structural differences do not affect the activity against one receptor, but do affect the activity against the other. Data points in Z5 are associated with pairs of compounds for which structural changes have little or no impact on the activity against the two receptors. In general, the most informative data points associated with receptor selectivity are those located at the outer regions of zones Z2, Z3, and Z4. The classification of data points in Figure 1 is independent of the structure similarity. However, since DAD maps are based on pairwise comparisons, it is straightforward to incorporate structure similarity information. Here we represent structure similarity using a continuous color scale from green (low similarity) to red (high similarity). Mapping structure similarity information into the DAD maps easily reveals the presence of single-, dual-, or triple-receptor activity cliffs as well as regions with continuous SAR for one or several receptors.[44] DAD maps can be extended to triple activity-difference (TAD) maps by adding a third dimension that represents the activity difference for a third receptor. TAD maps can be interpreted as the grouping of three DAD maps for each combination of two receptors.[44]

3. Results and discussion

3.1. Distribution of similarity measures

Figure 2 summarizes the 14,028 pairwise similarities calculated with 13 2D and 3D structure representations as cumulative distribution functions (CDF). The table at the bottom of the figure summarizes the statistics of each curve indicating the maximum, third and first quartile, median, mean, and standard deviation. 2D and 3D representations showed a wide variation of distributions. Similar results have been obtained for other data sets.[29,39,40] Atom Types, TopPh4Pairs and MACCS (322 bits) showed the highest similarity values with median values of 0.84, 0.77, and 0.68, respectively. Fingerprints with the lowest similarity values were atom neighborhoods, extended connectivity, and topological atom

Table 1
Correlation matrix between molecular representations

	AtomNeighborhoods	AtomTypes	EStateIndices	ExtendedConectivity	MACCS322	PathLength	TopAtomPairs	TopAtomTorsions	TopAtomTriplets	TopPh4Pairs	TopPh4Triplets	ShapeTanimoto	ComboScore/2
AtomNeighborhoods	1.000												
AtomTypes	0.491	1.000											
EStateIndices	0.556	0.447	1.000										
ExtendedConectivity	0.896	0.413	0.472	1.000									
MACCS322	0.623	0.365	0.479	0.642	1.000								
PathLength	0.800	0.350	0.456	0.901	0.704	1.000							
TopAtomPairs	0.751	0.804	0.503	0.709	0.612	0.659	1.000						
TopAtomTorsions	0.859	0.601	0.523	0.791	0.583	0.726	0.824	1.000					
TopAtomTriplets	0.865	0.587	0.497	0.845	0.639	0.791	0.920	0.865	1.000				
TopPh4Pairs	0.511	0.415	0.347	0.471	0.527	0.442	0.611	0.474	0.572	1.000			
TopPh4Triplets	0.527	0.383	0.363	0.591	0.435	0.617	0.545	0.496	0.553	0.579	1.000		
ShapeTanimoto	0.310	0.126	0.202	0.334	0.218	0.304	0.281	0.313	0.369	0.159	0.142	1.000	
ComboScore/2	0.480	0.252	0.290	0.497	0.359	0.456	0.455	0.466	0.548	0.285	0.270	0.820	1.000

triplets with median values of 0.07, 0.15, and 0.19, respectively. Several fingerprints had a normal distribution as revealed by the sigmoidal shape of the corresponding curves and the similar median and mean values.

Table 1 shows the correlation matrix of the Pearson's correlation coefficient between the 14,028 pairwise similarities for each pair of the 13 2D and 3D representations. High correlations between representations occur for topological atom pairs and topological atom triplets; atom neighborhoods and extended connectivity; path length and extended connectivity; topological atom triplets and topological atom torsions; topological atom triplets and atom neighborhoods; topological atom torsions and atom neighborhoods (correlation >0.85). EStateIndices showed low correlation (<0.55) with all other representations. This result is explained because EStateIndices is the only fingerprint used in this study that encodes the intrinsic electronic state of the atom as perturbed by the electronic influence of all other atoms in the molecule.[32] In contrast, other fingerprints with similar design showed high correlations, for example, topological atom pairs and topological atom triplets (correlation of 0.92).

3.2. Selection of structure representations for consensus models

In order to reduce the dependence of structure representation with chemical space, we combined similarity measures obtained by different methods using the principles of data fusion. In this work we computed the mean similarity, that is, mean fusion, of four selected representations. Our group has applied this approach to derive consensus models of the activity landscape of other data sets.[29,40,44] We computed the mean similarity of EStateIndices, topological atom pairs, topological pharmacophore triplets, and Combo Score (scaled). Following a similar approach recently reported,[29,45] from the initial pool of 13 2D and 3D similarities (cf. Section 2) we selected fingerprints that capture different aspects of the chemical structures[46] and showed relatively low linear correlations for the 14,028 pairwise structure similarities. The maximum correlation between any of the four selected representations was <0.55 (Table 2).

3.3. Consensus models of activity landscapes

Figure 3 shows DAD and TAD maps with the 14,028 pairwise potency differences of the 168 compounds of the data set. The distribution of the data points in the plots is independent of the structure similarity (see above). Data points are colored by mean structure similarity using a continuous scale from less similar (green) to more similar (red) structures. Mapping similarity values into the DAD maps clearly revealed activity differences associated with large and small differences in chemical structures making straightforward the identification of single-, dual-, and triple-receptor activity cliffs and scaffold hops (see below).

Figure 4 shows DAD and TAD maps displaying only 534 molecule pairs with the highest mean structure similarity (>0.75). A total of 393 pairs had potency difference lower than 1 log unit for all three receptors subtypes and are associated with a continuous SAR; 78 pairs were identified as single-receptor activity cliffs, 48 pairs as dual-receptor activity cliffs with 'direct SAR', 4 pairs as dual-receptor activity cliffs with 'inverse SAR', and 11 pairs as triple-receptor activity cliffs. Table 2 summarizes the number of single- and dual-receptor activity cliffs for each subtype identified in the DAD and TAD maps. For the set of 168 compounds studied in this work, the larger number of activity cliffs for PPARγ as compared to the number of cliffs for PPARα and δ suggests that the SAR for PPARγ is more discontinuous than the SAR of the other two subtypes. This is also an indicative that it would be more challenging to develop quantitative models such

Table 2
Counts of activity cliffs and scaffold hops

Type	PPAR subtype	Scaffold Hops	Activity cliffs[a]		
		Total	Total	Direct	Inverse
Single-target	α	14	18 (3)		
	δ	20	19 (0)		
	γ	30	41 (15)[b]		
Dual-target	α–δ	4	8 (2)	6 (2)	2 (0)
	α–γ	23	31 (16)	30 (15)	1 (1)
	δ–γ	7	13 (5)	12 (5)	1 (0)
Triple-target (pan cliff)	α–δ–γ	11	11 (7)	10 (7)	1 (0)

[a] Number in parenthesis indicate activity cliffs with potency difference of more than 1.5 log units.
[b] Six activity cliffs have potency difference of more than 2 log units.

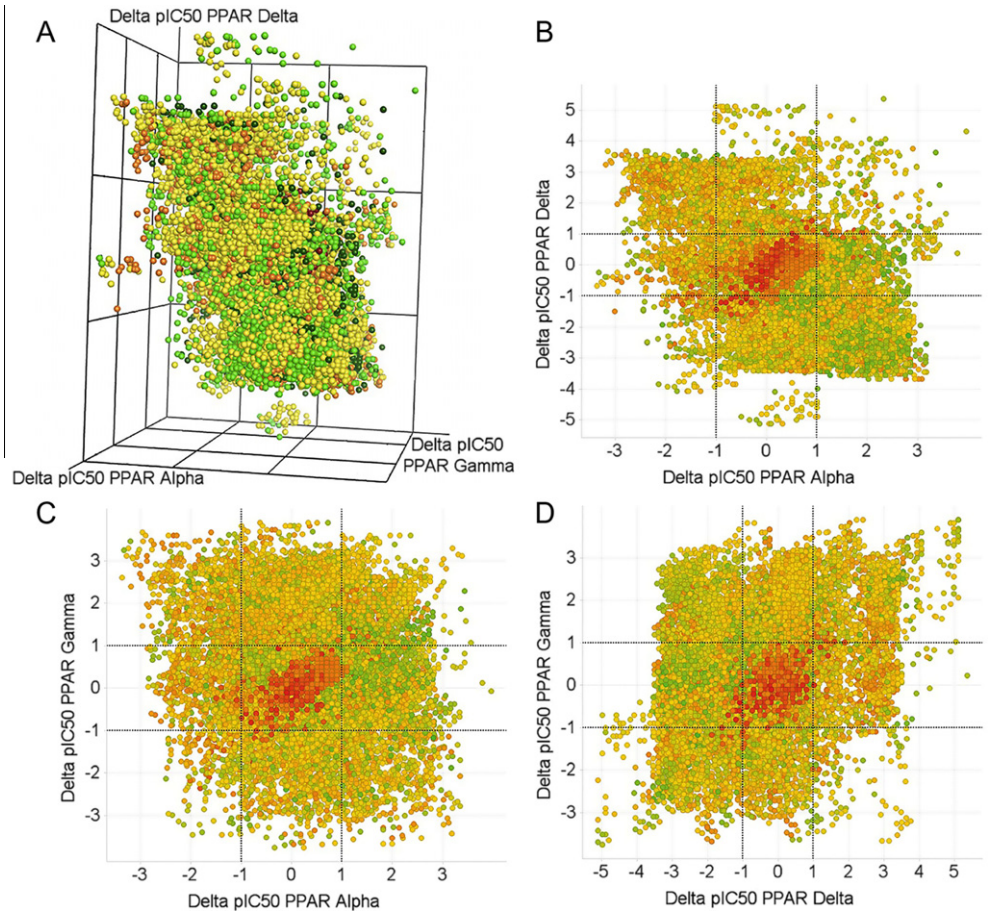


Figure 3. (A) Triple activity-difference maps for the three PPAR subtypes. Dual activity-difference maps for PPAR: (B) α–δ; (C) α–γ; and (D) δ–γ. Each map contains 14,028 data points resulting from the pairwise comparisons of 168 compounds. Data points are color-coded by the mean structure similarity using a continuous scale from more similar (red) to less similar (green).

as QSAR for PPARγ.[16] Of note, 15 out of the 41 activity cliffs for PPARγ have a potency difference of more than 1.5 log units and six cliffs of more than 2 log units (Table 2). Activity cliffs with a very large potency difference can be further identified as 'deep activity cliffs'.[40] The number of deep cliffs for PPARα and PPARδ is rather low (Table 2).

The α–δ DAD maps had a lower number of dual-receptor activity cliffs than the other two DAD maps (Table 2). This result suggests more of an agreement between the SAR of the 168 compounds tested against PPARα and δ. In contrast, the α–γ DAD map showed a larger number of dual-receptor cliffs indicating more disagreement between the SAR of the compounds tested against these two subtypes.

Selected pairs of compounds which are examples of single-, dual- and triple-receptor activity cliffs are labeled in Figure 4. Figures 5 and 6 show the chemical structures of the selected pairs. The corresponding values of potency difference for the three subtypes, and the structure similarity from different representations are summarized in Table 3.

3.3.1. Single-receptor activity cliffs

Figure 5 shows examples of single-target activity cliffs for the three PPAR subtypes. Pairs **32_36** and **90_93** (Fig. 5A) are representative examples of single-target activity cliffs for PPARα. In both cases compounds in the pair have a high structural similarity (mean similarity >0.75). The only structural difference is the

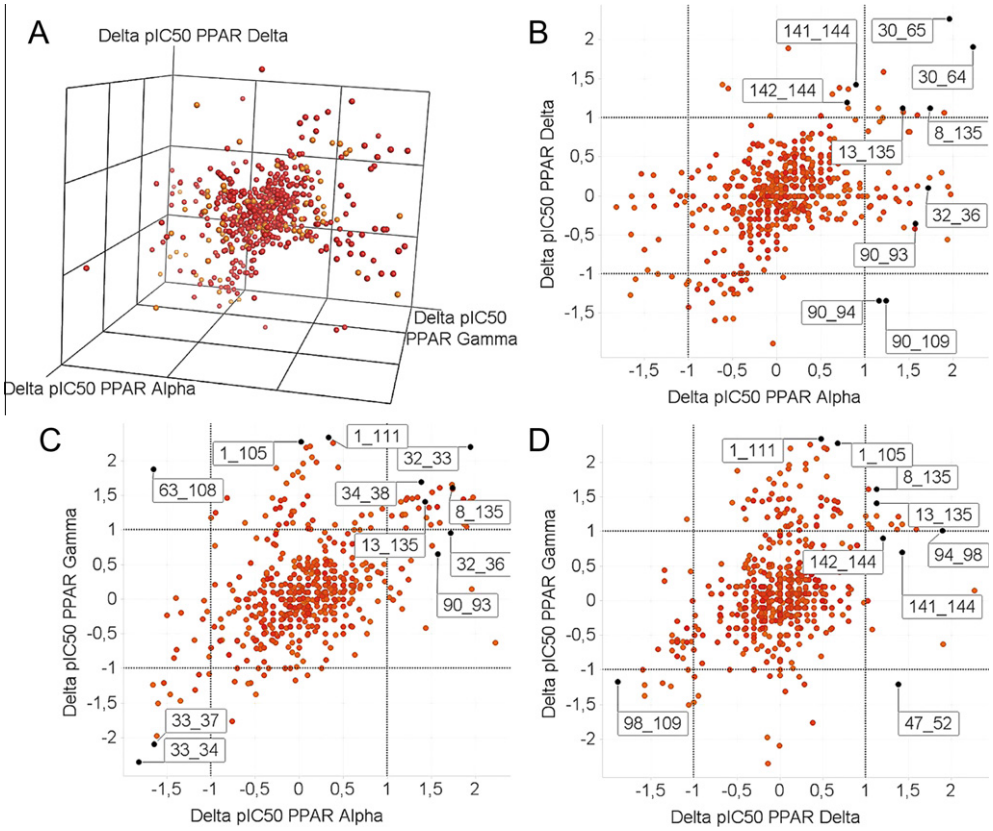


Figure 4. (A) Triple activity-difference map for the three PPAR subtypes. Dual activity-difference maps for PPAR: (B) α–δ; (C) α–γ; and (D) δ–γ. Each graph shows 534 data points, all with a mean structure similarity >0.75. Selected points are examples of activity cliffs.

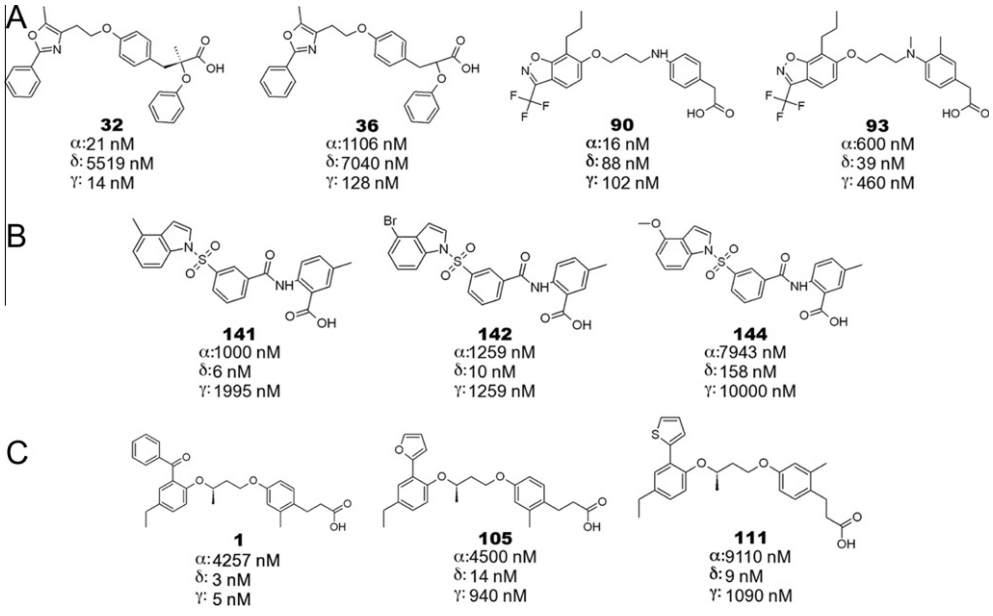


Figure 5. Chemical structures and IC_{50} values for single-target activity cliffs of PPAR (A) α (compound pairs **32_36** and **90_93**); (B) δ (pairs **141_144** and **142_144**); and (C) γ (pairs **1_105** and **1_111**). The position of each pair in the DAD maps is depicted in Figure 4.

addition of a methyl in the first pair and the change of a secondary amine for a tertiary amine in the second one. These small structural changes are associated with a potency difference of more than 1.5 log units against PPARα. Similarly, pairs **141_144** and **142_144** (Fig. 5B) are examples of single-target activity cliffs for PPARδ. In these cases, the substitution of the methoxy by a methyl or a bromide is enough to produce a significant change in the activity against PPARδ. Pairs of compounds **1_105** and **1_111** (Fig. 5C), in which the change of the phenylcarbonyl in **1** by a furan in **105** or a thiophene in **111** produces a decrease of more than 2 log units in the activity, exemplify single-target activity cliffs for PPARγ. It is important to mention that in all these examples the structural

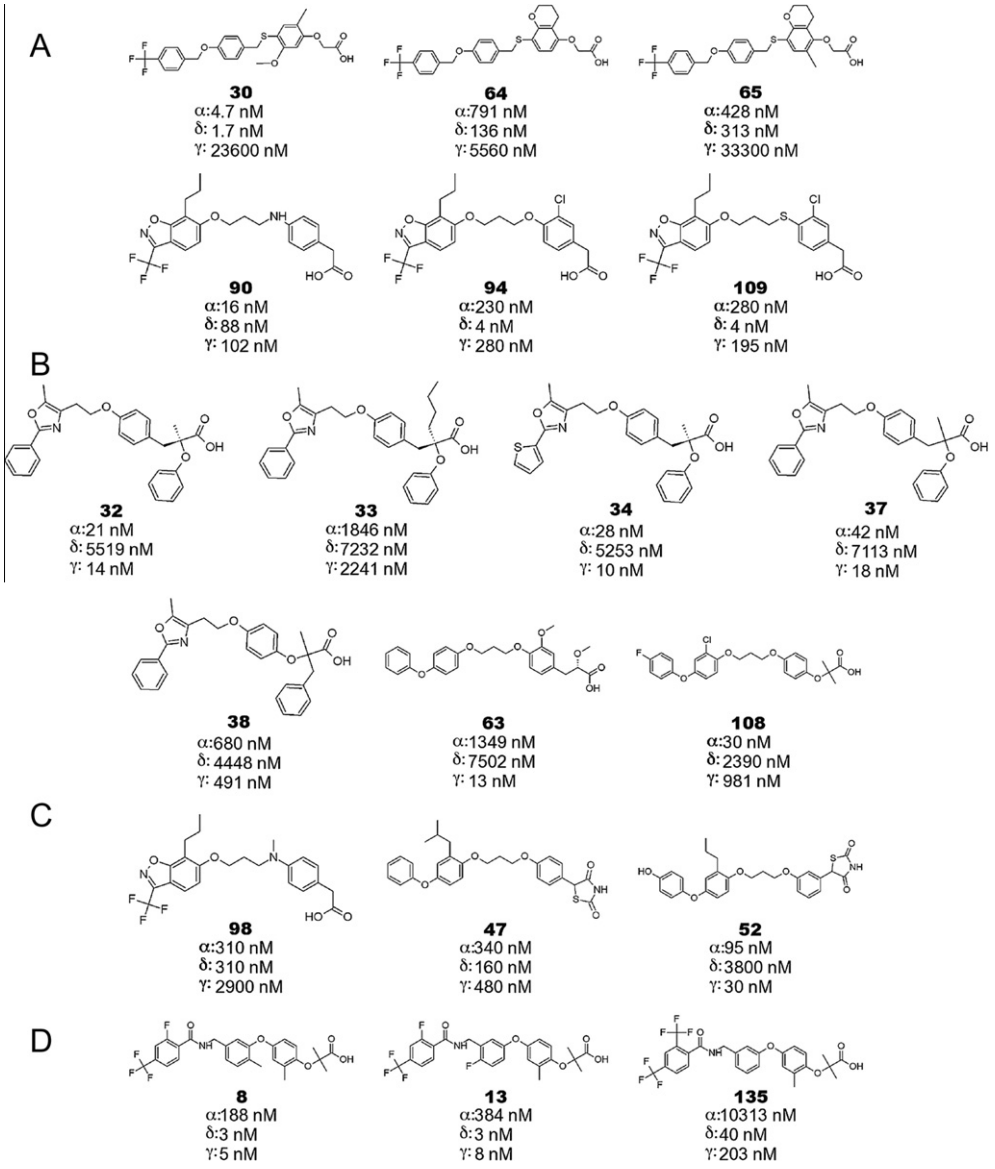


Figure 6. Chemical structures and IC_{50} values for compound pairs that form dual-target activity cliffs of PPAR (A) α–δ; (B) α–γ; and (C) δ–γ. Panel D represent the chemical structures and IC_{50} for compound pairs that form triple-target activity cliffs. The position of each pair in the DAD maps is depicted in Figure 4. See text for details.

differences only affect the activity against one PPAR subtype with a small impact in the other subtypes.

3.3.2. Dual- and triple-receptor activity cliffs

The pairs of compounds **30_64**, **30_65**, **90_94** and **90_109** (Fig. 6A) are examples of dual-receptor activity cliffs for PPARα and δ. Pairs **30_64** and **30_65** exhibit a continuous SAR as they are located in region Z1 of the α–δ DAD map. In other words, structural changes in these compounds affect in the same way the activity against the two subtypes. In these examples, the ring opening of the benzopirane in **64** and **65** into a methoxybenzene, as in **30**, results in an increase of the activity against both subtypes. On the other hand, structural changes between pairs **90_94** and **90_109**, located near region Z2, represent an inverse effect in the activity; the addition of a chloride and the change of the secondary amine in **90** by an oxygen or sulfur atom in **94** and **109**, respectively, decrease the biological activity towards PPARα whereas the same changes increase the activity toward PPARδ. Based on this observation we can suggest that the activity for PPARα depends on the heteroatom in the following order: NHR_2 > O > S.

Several dual-receptor activity cliffs were identified for PPARα and γ. Representative examples are the compound pairs **32_33**, **33_37**, **33_34**, **34_38** and **63_108** (Fig. 6B) of which the first four exhibit a direct SAR, whereas the last one exhibits an inverse SAR. The main structural difference in pairs **32_33**, **33_37**, located in region Z1 of the α–γ DAD map, is the length of the aliphatic chain. A longer chain would be associated with a decrease of the activity against PPARα and γ, but no effect on the activity against δ. In the same way, pair **33_34** presents two structural differences: the length of the aliphatic chain and the substitution of a phenyl by a thiophene ring. The pair **32_34** shows that replacing a phenyl by a thiophene ring does not impact activity, indicating that the length of the aliphatic chain is the main cause of the activity cliff. Similarly, in pair **34_38**, also presenting a direct SAR, the important difference between the structures of these compounds is the phenyl ether, which increases the activity toward PPARα and γ. In contrast, pair **63_108**, associated with an inverse effect in activity, is located in region Z2 of the α–γ DAD map. The presence of halogens and the lack of methoxyl groups in **108** are associated with an increase of the activity against PPARα and with a decrease against γ. On the

Table 3
Potency difference and structure similarity of selected single-, dual- and triple-receptor activity cliffs and scaffold hops

	Pair	$\Delta pIC_{50}\alpha$	$\Delta pIC_{50}\delta$	$\Delta pIC_{50}\gamma$	Mean Similarity	TopAtomPairs	TopPh4Triplets	Estate Indices	Combo Score/2
Activity cliffs									
Single α	**32_36**	1.722	0.106	0.961	0.86	0.92	0.98	1.00	0.56
Single α	**90_93**	1.574	−0.353	0.654	0.79	0.85	0.92	0.98	0.43
Single δ	**141_144**	0.900	1.421	0.700	0.87	0.95	0.91	0.99	0.65
Single δ	**142_144**	0.800	1.199	0.900	0.87	0.94	0.95	0.99	0.62
Single γ	**1_111**	0.330	0.477	2.338	0.78	0.85	0.84	0.84	0.58
Single γ	**1_105**	0.024	0.669	2.274	0.77	0.85	0.84	0.81	0.59
Dual α–δ direct	**30_64**	2.226	1.903	−0.628	0.76	0.86	0.99	0.99	0.22
Dual α–δ direct	**30_65**	1.959	2.265	0.150	0.77	0.85	0.97	0.99	0.29
Dual α–δ inverse	**90_94**	1.158	−1.342	0.439	0.87	0.89	0.93	0.96	0.70
Dual α–δ inverse	**90_109**	1.243	−1.342	0.281	0.84	0.89	0.94	0.97	0.57
Dual α–γ direct	**32_33**	1.944	0.117	2.204	0.78	0.86	0.92	1.00	0.35
Dual α–γ direct	**34_38**	1.385	−0.072	1.691	0.83	0.94	0.98	0.98	0.43
Dual α–γ direct	**33_37**	−1.643	−0.007	−2.095	0.80	0.86	0.92	1.00	0.43
Dual α–γ direct	**33_34**	−1.819	−0.139	−2.350	0.79	0.79	0.88	0.98	0.51
Dual α–γ inverse	**63_108**	−1.653	−0.497	1.878	0.78	0.80	0.93	0.82	0.57
Dual δ–γ direct	**94_98**	0.130	1.889	1.015	0.81	0.87	0.98	0.96	0.45
Dual δ–γ direct	**98_109**	−0.044	−1.889	−1.172	0.81	0.86	0.94	0.97	0.45
Dual δ–γ inverse	**47_52**	−0.554	1.376	−1.204	0.82	0.88	0.96	0.93	0.50
Triple	**8_135**	1.739	1.125	1.609	0.81	0.81	0.99	0.87	0.58
Triple	**13_135**	1.429	1.125	1.404	0.81	0.81	0.94	0.97	0.53
Scaffold hops									
Single α	**7_62**	0.082	2.835	2.232	0.25	0.33	0.17	0.17	0.35
Single α	**97_126**	0.067	−1.366	−2.632	0.25	0.35	0.24	0.19	0.24
Single δ	**95_151**	2.421	0.000	1.169	0.23	0.33	0.21	0.16	0.22
Single δ	**95_165**	2.721	0.097	1.369	0.25	0.35	0.21	0.18	0.26
Single γ	**57_168**	1.092	−2.635	0.087	0.24	0.31	0.13	0.19	0.33
Dual α–δ direct	**31_135**	0.319	−0.477	−1.832	0.25	0.38	0.29	0.07	0.24
Dual α–γ direct	**57_141**	0.092	−3.061	−0.313	0.24	0.19	0.20	0.17	0.40
Dual δ–γ direct	**33_83**	−1.698	−0.581	0.065	0.25	0.44	0.13	0.25	0.19
Triple	**92_143**	0.698	−0.455	−0.079	0.25	0.28	0.27	0.18	0.29
Triple	**110_131**	−0.067	−0.097	−0.691	0.23	0.27	0.22	0.20	0.25

other hand, the absence of halogens and the presence of methoxyl groups in **63** improve the activity towards PPARγ and decreases the activity toward PPARα. It is important to say that there is no significant difference in the activity of these pairs toward PPARδ as they are located in regions Z4 and Z3 of the α–δ and δ–γ DAD maps, respectively.

Compound pairs **94_98** and **98_109** exemplify dual-target activity cliffs for PPARδ and γ with direct SAR. These pairs present a high mean molecular similarity (0.81). As it can be seen in Figure 6C, the only difference is the substitution of the tertiary amine in **98** by an ether in **94** or thioether in **109** and the addition of the chloride. These changes confer a major activity against PPARδ and γ with a minor effect in the activity against PPARα. On the other hand, pair **47_52** represents a dual-target activity cliff for PPARδ and γ with and inverse SAR. In this example a different substitution pattern and a hydroxyl group decrease the activity against PPARδ while the activity against PPARγ is increased in more than 1 log unit.

Examples of triple-target activity cliffs are pairs **8_135** and **13_135** (Fig. 6D), which have high structure similarity and are located in region Z1 of all three DAD maps (Fig. 4). In these pairs, the substitution of the trifluoromethyl in **135** by a single fluoride, as in **8** and **13**, significantly increases the activity against the three PPAR subtypes.

3.3.3. Scaffold hops

Figure 7 shows DAD and TAD maps displaying 126 compound pairs with low mean structure similarity (<0.25). A total of 109 pairs had a potency difference lower than 1 log unit and are associated with scaffold hopping. Table 2 summarizes the number of single-, dual- and triple-receptor scaffold hops. Figure 7 shows the position of selected compound pairs in the DAD maps. The corresponding values of potency difference for the three subtypes, and the structure similarity from different representations are summaried in Table 3. Figure 7E and Figure S1 in the Supplementary data shows the chemical structures of the selected pairs.

Examples of compounds with low structure similarity but low potency difference for single receptors subtypes are pairs **7_62** and **97_126** for PPARα, pairs **95_151** and **95_165** for PPARδ, and pair **57_168** for PPARγ; all of them located in region Z3 or Z4 of the DAD maps (Fig. 7). Additionally, pairs **31_135**, **57_141** and **33_83**, located in region Z5 in only one DAD map, are examples of dual-receptor scaffold hops for α–δ, α–γ and δ–γ, respectively. Compound pairs **92_143** and **110_131**, which are located in region Z5 in all three DAD maps are examples of triple-target scaffold hops (Fig. 7E). It is worth noting that all four compounds in Figure 7E have some selectivity towards PPARδ. The scaffolds present in these four compounds are attractive scaffolds to develop selective compounds towards this isoform.

4. Conclusions and perspectives

Several structural features associated with ligand selectivity were identified from the analysis of single- and dual-activity cliffs identified in the data set of 168 compounds with reported activity against three PPAR subtypes. For example, if we consider that most of the PPAR agonists are composed by an aromatic tail, a linker or spacer, and an acidic head,[5,6] it is possible to suggest that heteroatoms in the spacer increase the activity for PPARα in the following order: NHR_2 > O > S > NR_3, for example, compounds **90**, **94**, **109**, **98**, respectively. Interestingly, the activity towards the three PPARs is increased when NHR_2 is used instead of NR_3, this may be a consequence of the hydrogen-bond donor and acceptor behavior of the secondary amine, while the tertiary amine only behaves as a hydrogen-bond acceptor. Additionally, results suggest that a large aliphatic chain near the acidic head, for example, the butyl of compound **33** compared with compound **32** and **37**, tends to reduce the

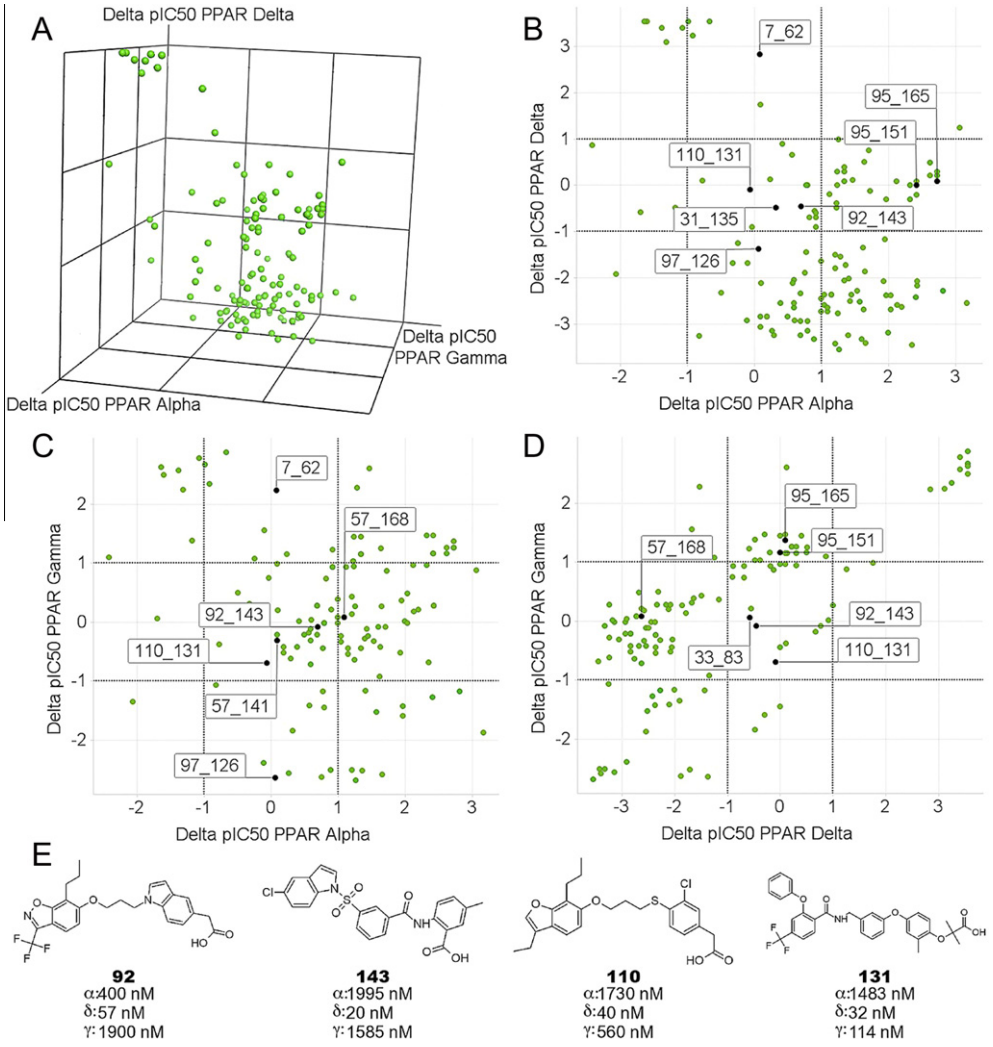


Figure 7. (A) Triple activity-difference map for the three PPAR subtypes. Dual activity-difference maps for PPAR: (B) α–δ; (C) α–γ; and (D) δ–γ. Each graph shows 126 data points, all with a mean similarity <0.25. Selected points are examples of scaffold hops. (E) Chemical structures and IC_{50} values of triple-target scaffold hops (compound pairs **92_143** and **110_131**).

biological activity toward PPARα and γ, making it an important characteristic that could be used for the design of new PPARδ selective agonists. On the other hand, an aliphatic chain near the aromatic tail seems to be important for activity and many compounds share this structural feature. Because of its importance, this aliphatic chain has been proposed before as an important pharmacophoric feature of pan PPAR agonists.[9] This work represents a first attempt to systematically analyze the SAR of PPAR ligands using the emerging approach of activity landscape modeling. Perspectives of this work include extending this approach to the SAR analysis of more compounds tested across different PPAR subtypes and the use of the models developed in this work to predict the activity landscape of other data sets.[47]

Acknowledgments

We are grateful to Jacob Waddell for proofreading the manuscript and to the anonymous reviewers for their comments. Authors thank OpenEye Scientific Software, Inc. for providing ROCS. This work was supported by CONACyT project No. 80093. O.M.-L is very grateful to CONACyT for the fellowship granted (No. 245408). J.L.M.-F thanks the State of Florida, Executive Office of the Governor's Office of Tourism, Trade, and Economic Development for funding.

Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.bmc.2012.04.005.

References and notes

1. Shaw, J. E.; Sicree, R. A.; Zimmet, P. Z. *Diabetes Res. Clin. Pract.* **2010**, *87*, 4.
2. Moller, D. E. *Nature* **2001**, *414*, 821.
3. Willson, T. M.; Brown, P. J.; Sternbach, D. D.; Henke, B. R. *J. Med. Chem.* **2000**, *43*, 527.
4. Arck, P.; Toth, B.; Pestka, A.; Jeschke, U. *Biol. Reprod.* **2010**, *83*, 168.
5. Zaware, P.; Shah, S. R.; Pingali, H.; Makadia, P.; Thube, B.; Pola, S.; Patel, D.; Priyadarshini, P.; Suthar, D.; Shah, M.; Jamili, J.; Sairam, K. V. V. M.; Giri, S.; Patel, L.; Patel, H.; Sudani, H.; Patel, H.; Jain, M.; Patel, P.; Bahekar, R. *Bioorg. Med. Chem. Lett.* **2011**, *21*, 628.
6. Pingali, H.; Jain, M.; Shah, S.; Patil, P.; Makadia, P.; Zaware, P.; Sairam, K. V. V. M.; Jamili, J.; Goel, A.; Patel, M.; Patel, P. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 6471.
7. Henke, B. R. *J. Med. Chem.* **2004**, *47*, 4118.
8. Rubenstrunk, A.; Hanf, R.; Hum, D. W.; Fruchart, J.-C.; Staels, B. *Biochim. Biophys. Acta, Mol. Cell Biol. Lipids* **2007**, *1771*, 1065.
9. Sundriyal, S.; Bharatam, P. V. *Eur. J. Med. Chem.* **2009**, *44*, 3488.
10. Al-Najjar, B. O.; Wahab, H. A.; Tengku Muhammad, T. S.; Shu-Chien, A. C.; Ahmad Noruddin, N. A.; Taha, M. O. *Eur. J. Med. Chem.* **2011**, *46*, 2513.
11. Bajorath, J.; Peltason, L.; Wawer, M.; Guha, R.; Lajiness, M. S.; Van Drie, J. H. *Drug Discovery Today* **2009**, *14*, 698.
12. Wassermann, A. M.; Wawer, M.; Bajorath, J. *J. Med. Chem.* **2010**, *53*, 8209.

13. Medina-Franco, J. L.; Yongye, A. B.; López-Vallejo, F. In *Statistical Modeling of Molecular Descriptors in QSAR/QSPR*; Matthias, D., Kurt, V., Danail, B., Eds.; Wiley-VCH, 2012; p 307.
14. Maggiora, G. M. *J. Chem. Inf. Model.* **2006**, *46*, 1535.
15. Guha, R.; VanDrie, J. H. *J. Chem. Inf. Model.* **2008**, *48*, 646.
16. Guha, R.; Van Drie, J. H. *J. Chem. Inf. Model.* **2008**, *48*, 1716.
17. Scior, T.; Medina-Franco, J. L.; Do, Q. T.; Martínez-Mayorga, K.; Yunes Rojas, J. A.; Bernard, P. *Curr. Med. Chem.* **2009**, *16*, 4297.
18. Peltason, L.; Bajorath, J. *Chem. Biol.* **2007**, *14*, 489.
19. Eckert, H.; Bajorath, J. *Drug Discovery Today* **2007**, *12*, 225.
20. Medina-Franco, J. L.; Martínez-Mayorga, K.; Giulianotti, M. A.; Houghten, R. A.; Pinilla, C. *Curr. Comput.-Aided Drug Des.* **2008**, *4*, 322.
21. Peltason, L.; Hu, Y.; Bajorath, J. *ChemMedChem* **2009**, *1864*, 4.
22. Wassermann, A. M.; Peltason, L.; Bajorath, J. *ChemMedChem* **2010**, *5*, 847.
23. Pérez-Villanueva, J.; Santos, R.; Hernández-Campos, A.; Giulianotti, M. A.; Castillo, R.; Medina-Franco, J. L. *Med. Chem. Comm.* **2011**, *2*, 44.
24. Medina-Franco, J. L.; Yongye, A. B.; Pérez-Villanueva, J.; Houghten, R. A.; Martínez-Mayorga, K. *J. Chem. Inf. Model.* **2011**, *51*, 2427.
25. Brown, N.; Jacoby, E. *Mini-Rev. Med. Chem.* **2006**, *6*, 1217.
26. Chen, X.; Lin, Y.; Liu, M.; Gilson, M. K. *Bioinformatics* **2002**, *18*, 130.
27. Liu, T.; Lin, Y.; Wen, X.; Jorissen, R. N.; Gilson, M. K. *Nucleic Acids Res.* **2007**, *35*, D198.
28. Scior, T.; Bernard, P.; Medina-Franco, J. L.; Maggiora, G. M. *Mini-Rev. Med. Chem.* **2007**, *7*, 851.
29. Yongye, A.; Byler, K.; Santos, R.; Martínez-Mayorga, K.; Maggiora, G. M.; Medina-Franco, J. L. *J. Chem. Inf. Model.* **2011**, *51*, 1259.
30. Sud, M. URL: www.MayaChemTools.org.
31. Filimonov, D.; Poroikov, V.; Borodina, Y.; Gloriozova, T. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 666.
32. Hall, L. H.; Kier, L. B. *J. Chem. Inf. Comput. Sci.* **1995**, *1039*, 35.
33. Rogers, D.; Hahn, M. *J. Chem. Inf. Model.* **2010**, *50*, 742.
34. Durant, J. L.; Leland, B. A.; Henry, D. R.; Nourse, J. G. *J. Chem. Inf. Comput. Sci.* **2002**, *42*, 1273.
35. Carhart, R. E.; Smith, D. H.; Venkataraghavan, R. *J. Chem. Inform. Comput. Sci.* **1985**, *25*, 64.
36. Nilakantan, R.; Bauman, N.; Dixon, J. S.; Venkataraghavan, R. *J. Chem. Inf. Comput. Sci.* **1987**, *27*, 82.
37. Renner, S.; Fechner, U.; Schneider, G. In *Pharmacophores and Pharmacophore Searches*; Wiley-VCH GmbH KGaA, 2006; p 49.
38. Bonachéra, F.; Parent, B.; Barbosa, F.; Froloff, N.; Horvath, D. *J. Chem. Inf. Model.* **2006**, *46*, 2457.
39. Medina-Franco, J. L.; Martínez-Mayorga, K.; Bender, A.; Marín, R. M.; Giulianotti, M. A.; Pinilla, C.; Houghten, R. A. *J. Chem. Inf. Model.* **2009**, *49*, 477.
40. Pérez-Villanueva, J.; Santos, R.; Hernández-Campos, A.; Giulianotti, M. A.; Castillo, R.; Medina-Franco, J. L. *Bioorg. Med. Chem.* **2010**, *18*, 7380.
41. Yongye, A. B.; Bender, A.; Martinez-Mayorga, K. *J. Comput.-Aided Mol. Des.* **2010**, *24*, 675.
42. Jaccard, P. *Bull. Soc. Vaudoise Sci. Nat.* **1901**, *37*, 241.
43. Willett, P.; Barnard, J. M.; Downs, G. M. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 983.
44. Medina-Franco, J. L.; Yongye, A. B.; Pérez-Villanueva, J.; Houghten, R. A.; Martínez-Mayorga, K. *J. Chem. Inf. Model.* **2011**, *51*, 2427.
45. Waddell, J.; Medina-Franco, J. L. *Bioorg. Med. Chem.* in press. doi:http://dx.doi.org/10.1016/j.bmc.2011.11.051.
46. Bender, A.; Jenkins, J. L.; Scheiber, J.; Sukuru, S. C. K.; Glick, M.; Davies, J. W. *J. Chem. Inf. Model.* **2009**, *49*, 108.
47. Kuwabara, N.; Oyama, T.; Tomioka, D.; Ohashi, M.; Yanagisawa, J.; Shimizu, T.; Miyachi, H. *J. Med. Chem.* **2012**, *55*, 893.