# Prediction of the PPARα agonism of fibrates by combined MM–docking approaches

Fabio Lannutti *, Alessandro Marrone, Nazzareno Re

*Dipartimento di Scienze del Farmaco, Università "G d'Annunzio" di Chieti-Pescara, Via dei Vestini 31, I-66100 Chieti, Italy*

## ARTICLE INFO

## ABSTRACT

Fibrates are peroxisome proliferator-activated alpha receptor (PPARα) activators derived from fibric acid and are the most clinically used therapeutics in the treatment of hypertriglyceridemia. Long standing studies on these drugs have accumulated a large body of experimental data about their biological activity and, more recently, on the molecular mechanism mediating their PPARα agonism. An immense interest for the discovery of new fibrates with improved potency and PPARα selectivity has stimulated many investigations toward a deeper understanding of structure–activity relationships controlling their activity. The present study aimed at investigating the binding properties of a set of 23 fibrates, characterized by similar carboxylic heads but differing in the size and orientation of the hydrophobic portion, using computational approaches. We combined standard docking and molecular mechanics approaches to better describe the adaptation of the protein target to the bound ligand. The agonist potencies were then regressed against the calculated binding energies to elaborate predictive model equations. The obtained models were characterized by good performances realizing a fair trade-off between accuracy and computational costs. The best model was obtained with a regression procedure allowing automatic generation of a training subset from the whole set of trials and filtering out outliers, thus highlighting the importance of regression strategies.

© 2011 Elsevier Inc. All rights reserved.

## 1. Introduction

Peroxisome proliferator-activated receptors (PPARs) belong to the superfamily of nuclear receptors [1]. PPARs form a heterodimer with the retinoid X receptor (RXR) upon binding of activating ligands and transactivate genes containing a peroxisome proliferator response element (PPRE) [2]. Three subtypes of PPAR receptors, PPARα, PPARγ, and PPARδ, have been recognized to date; their endogenous ligands include fatty acids and eicosanoids [3]. PPARα is predominantly expressed in the liver, where it plays a critical role in lipid metabolism and in the regulation of transport and catabolism of fatty acids [4]. In addition to the liver, PPARα is also expressed, albeit at lesser extent, in a variety of cell types, including smooth muscle cells, endothelial cells, and macrophages, playing a pivotal role in processes such as atherosclerosis and inflammation. Several studies recognizing the reduced PPARα expression and activation as a possible explanation for cardiovascular diseases have demonstrated the use of synthetic activators as a valuable therapeutic approach, resulting in atheroprotective effects in humans [5].

Among the most effective molecules acting as PPARα activators, fibric acid derivatives are clinically used to treat hypertriglyceridemia [6a–c]. Fibrates form a class of well-known PPAR agonists currently used in therapy [6d–f]. Long standing studies on these drugs have contributed to the accumulation of a large body of experimental data about their biological activity, stimulating an increased interest in the design of increasingly more potent and selective PPAR-activating drugs for the treatment of dyslipidemia [7–12].

In particular, there is an immense interest in the design of new fibrates with improved potency and selectivity toward the activation of PPARα.

The PPARα agonism of fibrates is due to their ability to induce an activated conformation of the receptor by binding to the so-called ligand binding domain (LBD). The activated PPAR can in turn recruit a co-activator peptide and initiate the upstream biochemical cascade, leading to gene transcription. The current model for ligand-dependent activation of PPARs [1–3] proposes that agonist binding induces a specific conformation of PPARα in which the transcriptional activating function 2 (AF2) helix, together with helices 3 and 4, provide for a suitable interface for the binding of the PPARα co-activator.

Agonist binding promotes and stabilizes the conformation and orientation of the AF2 domain and helices 3 and 4, inducing and sustaining the activation state of PPARα. The investigation of the

binding properties of known PPARα modulators could thus be essential to assess the molecular basis of fibrate agonism and to support the design of new agonists.

In addition to biological and SAR results, several 3D-structures of co-crystallized complexes of human PPARα with agonists are currently available from the PDB archive, and this accessible information is very helpful in drawing the molecular basis of the PPARα agonists' mechanism of action [13–16]. In particular, the co-crystallized structure (PDB code 2P54) of the complex of PPARα with a new potent fibrate, GW590735, and the SAR investigation of a series of its analogues have been recently reported [13]. The receptor molecule found in this complex was presumably in the activated conformation because the co-activator peptide was co-crystallized together with the bound PPARα and interacted with helices 3 and 4 and with the AF2 helix, i.e., in agreement with the current knowledge of the molecular mechanism of activation of this nuclear receptor.

A detailed knowledge of how a potent fibrate such as GW590735 interacts with the PPARα activated conformation can then be used to model the binding properties of other fibrate analogues by means of target-directed rational approaches.

Among the available methods to focus on this issue [17], molecular docking is particularly appealing because of the relatively short computational time required and because it is widely accepted to calculate binding geometries and to discriminate active from non-active compounds [18]. The scoring functions implemented in most docking procedures describe the ligand–target interactions through empirical formulae that allow easy processing and ranking of molecular structure databases. However, the scoring functions implemented in the majority of the currently used docking programs, such as Glide [19a–b], Gold [19c], and AutoDock [19d], are rarely employed in lead optimization due to weak predictive power as far as congeneric ligands are concerned. On the other hand, the development of computational protocols for the target structure-based prediction of the activity of congeneric ligands would be of great interest for drug designers. How and why does the replacement of certain functional groups affect the lead activities? How could the chemical structure be changed to optimize or maintain the lead biological activity?

Several alternative computational approaches have been proposed to improve the accuracy in the investigation of the binding properties of potentially active ligands. Linear interaction energy (LIE) [20] and molecular mechanics Poisson–Boltzman Surface Area (MM-PBSA) or molecular mechanics generalized Born Surface Area (MM-GBSA) [21,22] have been recently reported as valuable solutions to this issue, balancing both the accuracy and practicability of computational efforts of binding free energy estimations. The calculation of target–ligand geometries of bound complexes can be effectively dealt with by interfacing docking results, which is able to generate conceivable first sketches of ligand binding poses with molecular dynamics (MD) simulations, leading to more accurate estimates of target–ligand binding energy.

More recently, to reduce the computationally demanding requirements and to apply these approaches to high throughput virtual screening, several workflows have been derived that replace MD sampling with simple energy minimizations, such as modified MM-PBSA and MM-GBSA versions [23a–c] and linear interaction energy with continuum electrostatics (LIECE) [23d].

Docking approaches have recently been used to investigate the binding properties of either phthalate monoesters or newly designed fibrates as potential ligands of PPARγ or PPARα, respectively [24a–b].

In the present investigation, we used an LIECE approach through the combination of docking and MM calculations to estimate the binding geometry and the molecular interactions of a series of 23 fibrates with PPARα. The description of agonist induced-fit effects responsible for PPARα activation was simplified here by the assumption that the backbone conformation of the target induced by any of the considered fibrates can be, to a certain extent, represented by that co-crystallized with GW590735 (ligand **1**). This assumption relies on the fact that all the considered agonists, characterized by similar carboxylate heads, are expected to induce a similar backbone adaptation in response to agonist binding.

The ligand binding poses were preliminarily calculated in the target conformation co-crystallized with GW590735 using a Glide docking standard approach, and the resulting best XP docking solutions were then considered for further force field relaxation of the target–ligand complexes. The local minimization of the docking complexes allows modeling of the specific adaptation of the target to each bound agonist, mainly through the optimization of the side chain conformation of the LBD residues; an improvement in the prediction of both binding geometries and binding energies is expected at this stage.

Several models for the prediction of the PPARα-fibrate binding energy were then developed by means of both monovariate and multivariate regression approaches expressing the experimental agonist efficacies ($EC_{50}$) as a function of the calculated PPARα–fibrate interaction energies.

Although it encodes the biological response to PPARα activation and is not directly connected to the PPARα–fibrate affinities, $EC_{50}$ can be correlated to the binding energies in certain conditions (see below).

Two alternative regression approaches were tested and compared for the generation of training subsets from the 23 considered fibrates, namely the user-defined training subset (UDTS) and the automatically generated training subset (AGTS) procedures. Both procedures provide model equations in which the single contributions to the target–ligand binding energy are correctly weighted for their attractive/repulsive character. The best predictions were obtained with AGTS, which was more effective in filtering out outliers and led to more specialized models.

## 2. Computational methods

### 2.1. Docking studies

The crystal structure of a complex between the PPARα binding subunit and GW590735 (**1**) was taken from the PDB archives (entry code 2P54) and employed as the target in docking calculations.

The selected PDB entry of the co-crystallized complex was graphically inspected and analyzed to find possible structural mismatches using the molecular modeling software Maestro [25] and the structural analysis tool Pymol [26].

Solvent molecules in the crystal structure were removed, and the protein was prepared for the docking studies using the protein preparation module implemented in Maestro. This procedure adds polar hydrogen atoms and minimizes the protein within a gradient threshold of 0.30 kJ mol$^{-1}$ Å$^{-1}$ using the OPLS-AA force field [30].

The 3D structures of all the 23 ligands were built in Maestro and subsequently minimized using the MacroModel program [27]. The local minimization was performed with the all-atoms OPLS-AA force field [30] through the truncated-Newton conjugate gradient (TNCG) algorithm by applying the gradient threshold criteria at 0.05 kJ mol$^{-1}$ Å$^{-1}$, a maximum of 500 iterations, and simulating the water environment with the GB/SA polarizable continuum method implemented in the program.

All docking calculations were carried out with Glide [28,29]. This program performs a hierarchical search of ligand conformations undergoing a filtering procedure and finally minimizes in the field of the receptor using the OPLS-AA force fields [30] in conjunction with a distance-dependent dielectric model. The lowest-energy

poses obtained in this fashion were subjected to a Monte Carlo procedure to obtain the final set of docking solutions. Glide uses two concentric boxes to generate the potential grids and define the binding site. The grids are computed within the space defined by the "outer box", which encompasses all the ligand atoms. The "inner box" is defined as containing all acceptable positions for the ligand center upon docking. Both the outer and inner boxes were centered on the mass center of the crystallographic ligand with edge lengths of 28 Å and 10 Å, respectively.

Default input parameters were used in all computations (no scaling factor for the van der Waals radii of non-polar protein atoms and a scaling factor of 0.8 for non-polar ligand atoms). All compounds were docked and scored using the Glide standard-precision (SP) mode [29a]. Upon completion of each docking calculation, 30 poses per ligand were saved. The best-docked structures were ranked using a model energy score (Emodel) derived from a combination of GlideScore (Gscore, a modified and extended version of the empirically based ChemScore function), Coulombic, and van der Waals energies, and the strain energy of the ligands [29a]. The top-ranked compounds obtained in this way were docked and scored again with the Glide extra-precision (XP) mode [29b], and the best of 10 XP-docked structures was finally selected as final docking solution.

The selection of docking solutions was performed first by considering the poses best resembling the experimental binding geometry of **1**. Among them, the pose with the highest score was then chosen as the final docking solution.

The comparison between experimental and calculated binding poses was carried out through the superposition module of Maestro based on the calculation of the atomic root mean square deviations (RMSDs).

For **1**, the RMSD was calculated in-place and after pose superposition. For all other ligands, the in-place RMSDs of the common carboxylic head with respect to that of the experimental poses of **1** were also calculated.

### 2.2. Molecular mechanics

The agonist poses resulting from the docking calculations were considered for further molecular modeling of the receptor–ligand complex structure, thus taking into account part of the ligand induced fit effects of the receptor structure, which is known to be important for this system.

Indeed, the side chain conformational adaptation was partially modeled for each agonist by relaxation of the protein–ligand complex obtained from docking.

The local minimization of each receptor–agonist complex obtained after analyses of docking results was carried out in gas phase using the MacroModel program. The AMBER force field [31] was employed using the default MacroModel parameterization and including the explicit hydrogen bond term. The non-bonded contributions were calculated by applying the 7.0, 12.0, and 4.0 Å criteria for the van der Waals, electrostatic, and hydrogen bond interactions cut-off, respectively.

All optimizations were performed with the first-derivates LBFCG algorithm and setting the gradient-based convergence threshold to $0.05 \, \mathrm{kJ \, mol^{-1} \, Å^{-1}}$, with a maximum number of 8000 iterations, warranting the full convergence for each run.

The minimized bound complexes were then inspected to analyze the effect of MM relaxation on both the bound agonist pose and the protein conformation. To this aim, the ligand binding poses after minimization were superimposed in Maestro workspace with the corresponding docking poses, and both the in-place and minimum RMSD values were calculated; in the case of ligand **1**, such a comparison was also extended to the corresponding experimental pose. We also compared the X-ray and MM relaxed structures

of the protein target through superimposition and minimization of the backbone α carbons RMSD. In particular, the superimposed structures of the protein target were used to investigate the effects of the relaxation on the AF2 helix conformation and on the binding site residues, which were expected to be more deeply involved in the ligand induced fit process, and the RMSD values for AF2 helix alpha carbons and the non-hydrogen atoms of the binding site were calculated (Supporting Information).

Finally, the protein–ligand binding energies were estimated as the energy difference between the optimized complexes and the separated receptor and ligand structures calculated in the presence of a water environment within the same Amber force field employed in the complex optimizations. The aqueous solvent was simulated through the GB/SA polarizable continuum method implemented in MacroModel code by imposing the default set-up for water simulation [32].

### 2.3. Target–ligand binding energies

We used both the GlideScore values of the best XP poses ($G_{\mathrm{score}}$) and the binding energies from the MM calculations ($\Delta E_{\mathrm{ff}}$) as theoretical estimates of the affinity of each agonist to PPARα. The first two models were obtained by monovariate regression analyses of GlideScore (Model 1) and MM binding energies, $\Delta E_{\mathrm{ff}}$ (Model 2):

$$\Delta G_{\mathrm{int}} = b_0 + b_{\mathrm{score}} G_{\mathrm{score}}$$

$$\Delta G_{\mathrm{int}} = b_0 + b_{\mathrm{ff}} \Delta E_{\mathrm{ff}}.$$

$\Delta E_{\mathrm{ff}}$ breaks down into its single energy contributions, $\Delta E_{\mathrm{els}}$, $\Delta E_{\mathrm{VdW}}$, and $\Delta E_{\mathrm{HB}}$. The inclusion of the solvation free energy through the GB/SA approach and a further $\Delta E_{\mathrm{ster}}$ correction for the energy penalty required for the ligand to pass from the lowest conformation in solution to that assumed when docked to the receptor allowed the development of an LIECE-like model (Model 3) through a five parameter multivariate regression:

$$\Delta G_{\mathrm{int}} = b_0 + b_{\mathrm{els}} \Delta E_{\mathrm{els}} + b_{\mathrm{VdW}} \Delta E_{\mathrm{VdW}} + b_{\mathrm{HB}} \Delta E_{\mathrm{HB}}$$
$$+ b_{\mathrm{solv}} \Delta G_{\mathrm{solv}} + b_{\mathrm{ster}} \Delta E_{\mathrm{ster}}.$$

Further multivariate analyses were performed by reducing the number of regressors by backward variable selection (BVS) according to statistical significance ($F$), correlation ($R^2$), and positive signs for all the regressed coefficients; the latter feature was taken into account to keep the "attractive" or "repulsive" meaning of each energy contribution. The BVS analysis allowed the generation of two further models by first removing $\Delta E_{\mathrm{els}}$ and $\Delta E_{\mathrm{HB}}$ (Model 4) and then removing $\Delta E_{\mathrm{ster}}$ (Model 5).

$$\Delta G_{\mathrm{int}} = b_0 + b_{\mathrm{VdW}} \Delta E_{\mathrm{VdW}} + b_{\mathrm{solv}} \Delta G_{\mathrm{solv}} + b_{\mathrm{ster}} \Delta E_{\mathrm{ster}}$$

$$\Delta G_{\mathrm{int}} = b_0 + b_{\mathrm{VdW}} \Delta E_{\mathrm{VdW}} + b_{\mathrm{solv}} \Delta G_{\mathrm{solv}}.$$

Each model equation was validated by an external validation approach in which the whole ligand set was divided in two subsets, each containing 60–70% extended fibrates and 40–30% smaller fibrates. In particular, ligands **1**–**10** were used as a user-defined training set (UDTS), whereas ligands **11**–**23** were used as an external validation set. UDTS included six extended fibrates [13], three classical fibrates (fenofibrate, clofibrate, and ciprofibrate) and one fibrate-like agonist (gemfibrozyl) in which both structural and potency variability were representative of the whole ligand set. The UDTS model effectiveness was evaluated in terms of both predictive squared correlation coefficient ($Q^2$) and prediction error (*PE*), calculated as the standard error in the regression between the experimental values and those predicted by the model of the whole set (training plus validation).

An alternative and less biased validation approach was also considered. This was based on an iterative leave-one-out test performed with an in-home R script. Such a test searches for effective training subsets starting from the whole data matrix. The data entry (corresponding to a ligand) with the highest impact on $PE$ was iteratively removed while $R^2$, $F$, and $PE$ were monitored. Removing one point per iteration, a validation set was assigned such that the number $k$ of leave-one-out iterations corresponds to the number of trials in the final validation set, whereas the training set was formed by the remaining $N$-$k$ trails, where $N$ is the total number of trials. This method led to automatically generated training subsets (AGTS) and validation subsets to be used to calculate the corresponding $Q^2$ and $PE$ values.

AGTS has some advantages with respect to UDTS: (i) the training set selection is automatic and unbiased; (ii) the number of iterations can be optimized until improved models are gained; and (iii) the model search is oriented to minimal prediction errors.

The final validation was performed after five iterations, leading to 17 and 6 data entries in the training and validation subsets, respectively, such that approximately the 75% of data were used as training, the remaining 25% were used as validating. Both mono-variate (GlideScore and $\Delta E_{ff}$) and multivariate (contributions from $\Delta E_{ff}$ breakdown) regression analyses passed through the AGTS validation procedure; the BVS approach was again used in the multivariate analyses for the variables selection.

A total of 10 models were generated and compared: models 1–5 refer to the UDTS procedure, and models 6–10 refer to the AGTS procedure, reported in Table 2a and b, respectively.

## 3. Results and discussion

### 3.1. Binding geometries

Fibrates interact with PPARα receptors through binding sites located in the central core of the LBD, a large T-shaped cavity flanked by helices 3, 5, 7, 11, and 12 [1]. All the observed structures for the complexes of these ligands with PPARα show the same binding mode: (i) the carboxylic head interacts mainly with AF2; (ii) the aryl group directly bound to the head does not form appreciable hydrophobic interactions but does behave as a spacer; and (iii) the terminal tail extends toward the lower or upper distal cavity and forms the most important and strongest hydrophobic interactions. In particular, the hydrogen bond between the carboxylic head and Tyr464 in the AF2 helix is conserved in all agonist complexes and has been proposed to directly stabilize the AF2 helix in a conformation allowing co-activator recruitment. The carboxylic head is also hydrogen bonded to Ser280, Tyr314, and His440, belonging to helices 3, 5, and 11, respectively, which participate in the construction of the co-activator binding site. Indeed, in all known nuclear receptor crystal structures, the most notable structural differences between apo- and agonist-bound LBDs are observed in the C-terminal helix AF2.

Sierra et al. [13] have investigated a series of fibrates structurally correlated to **1** and inferred the structure–activity relationships affecting this class of PPAR agonists.

Due to their extended poly-aryl group attached to their fibrate moieties, the members of this series are expected to interact with the same PPARα binding site residues of other classical fibrates and to also be involved in further interactions that could be of interest in the design of new fibrate candidates.

In this study, a combination of docking and molecular mechanics approaches was addressed to investigate the binding properties of a set of 23 fibrates toward the PPARα receptor (see Table 1).

The whole set of activators was obtained by taking 12 candidates (**1–6** and **12–17**) from Ref. [13], 4 commonly marketed fibrate
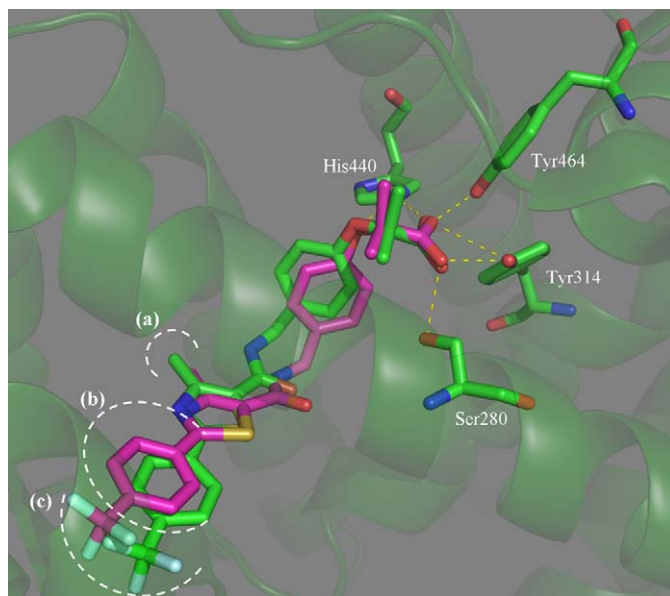


**Fig. 1.** Calculated binding poses of the 2P54 agonist (green carbons) superimposed with the corresponding experimental geometry (magenta carbons). The composition of hydrophobic pockets (a), (b) and (c) is given in the text.

agonists (**7**, **8**, **10**, **11**) [33–35] and other entries represented by gemfibrozyl [34], **9**, and six chiral analogues of classical fibrates [35], **18–23**, obtained by removing either the pro-S or pro-R α-methyl group from the typical α,α-dimethyl structure of fibrates. The latter molecules were included to extend the structural variety inside the ligands set in which all members, sharing the same carboxylic head and lacking any other ionisable functional group in their structure, are expected to be monoanions in physiological conditions.

Docking simulations were performed for all these ligands as discussed in the computational methods to provide for a first sketch of the fibrate–PPAR binding geometries.

The reliability of our docking approach in the prediction of the binding geometries of the considered fibrates to the PPARα receptor was evaluated by the comparison of the binding geometry calculated for **1** with that from the corresponding crystal structure (see Table S1). The Glide-predicted geometry is shown in Fig. 1 superimposed with the X-ray structure of the **1**-PPARα complex and indicates a good agreement, as determined by the small RMSD values of 0.55 Å and 0.36 Å (Table S1) calculated in-place and after pose superposition, respectively. All the hydrogen bond interactions of the carboxylate head, involving residues Tyr464, Ser280, His440 and responsible for the activation of agonist signaling, were correctly predicted.

The position of the heterocyclic biaryl tail of **1** was also correctly reproduced in the calculated pose. This bulky moiety occupies an extended hydrophobic pocket showing close contacts with three groups of protein residues: (a) Leu344 ⇔ CH$_3$ substituent of the thiazole ring; (b) Cys275, Ile272, Val332 ⇔ the distal aromatic ring; and (c) Val255, Leu247, and Ile241 ⇔ p-CF$_3$ substituent of the distal aromatic ring.

All the remaining ligands (**2–23**) maintain the same hydrogen bond interactions of the carboxylate head observed in ligand **1**; the extended fibrates structurally analogous to **1** also show an orientation of the heterocyclic biaryl tail similar to most of the remaining ligands. On the other hand, the biphenyl tail of ligands **22** and **23** could not adapt to the extended hydrophobic pocket like the tail of the other fibrates and thus was hosted in a different hydrophobic pocket formed by Phe318, Ile317, Val324, and Met330 residues (see Supporting Information).

Because all the considered fibrates share the same carboxylate head of **1**, which is one of the most potent agonists, we considered the docking pose of the latter as a reference structure for all the other agonists. We therefore calculated the RMSDs of the carboxylate heads for the docking poses of the other 22 fibrates (Supporting Information). The position and orientation of the carboxylate heads of the majority of ligands resembled that of the reference ligand **1**, whereas the carboxylate heads of ligand **22** and **23** were, as expected, the most displaced (higher RMSD) because of the different binding fashion induced by the biphenyl moiety.

Our calculations showed that all the considered fibrates bind to PPARα receptors through the same structural pattern observed in **1**. The carboxylate head interactions with the AF2 domain are the most conserved, in agreement with the role played by these interactions in the mechanism of PPARα agonism. The contribution of these interactions to the PPARα-agonist binding energy is thus either approximately constant or very similar in all the considered fibrates. On the other hand, the distal aryl moieties can make contacts with several groups of binding site residues, mainly in the large distal hydrophobic pocket but also, as with **22** and **23**, with an alternative hydrophobic site closer to the AF2 domain, with the orientation toward the two hydrophobic sites being controlled by the structure of the central aryl spacer.

Our results suggest that the extent and composition of the bridging moiety, which determines the reciprocal adaptation of the carboxylate head and the hydrophobic tail poses, may be critical for the development of fibrates with improved binding properties. The design of new fibrates by the modulation of head–tail bridging is currently ongoing in our laboratories.

The 23 docking complexes were subsequently optimized in a force field to include a description of induced-fit effects in the PPAR–fibrate binding model. The effect of MM relaxation on the fibrates' binding poses was analyzed by the superposition of the calculated bound complexes on the experimental 2P54 structure (see Section 2). As reported in Table S1, the MM relaxation gave rise to a slight increase in the RMSD of the carboxylate head poses with respect to that experimentally observed for ligand **1**. Indeed, small changes in the reciprocal positions of protein domains may occur during the MM optimization because the input complexes are still biased by both the binding of ligand **1** and the crystal packing. These changes may affect the shape and the volume of the LBD, thus provoking the translation of ligand pose with respect to that experimentally observed; such an effect is mainly responsible for the increase in the RMSD of carboxylate head poses upon MM relaxation of docking complexes. Indeed, an appreciable adaptation of the binding site volume to the size of the bound agonist was detected (Table S1).

Changes induced on the target structure were also analyzed by the superposition of MM relaxed complexes through the calculations of RMSD on specific frames of protein structure. The average RMSD on the backbone atoms of whole PPARα (Table S1) was estimated to be $1.13 \pm 0.06$ Å; the small value of standard deviation suggests that backbone adaptation has occurred in an almost similar extent for all the bound complexes. On the other hand, the average RMSD calculated on the backbone atoms of AF2 only was $1.90 \pm 0.20$ Å, indicating that, as expected, this protein frame was more responsive to the agonist binding.

The RMSD analysis was also performed on both whole protein and LBD side chain atoms. The side chain responsiveness was quite homogeneous, as indicated by the small difference between the average RMSD values calculated on whole protein and LBD. It is worth noting that, despite their higher flexibility, the average RMSDs on side chain atoms were smaller than those calculated for the AF2 backbone atoms in both cases (Table S1).

Thus, our combined docking–MM approach was able to account for the adaptation of PPARα to fibrate binding and to clearly iden-

tify the AF2 backbone as the most responsive protein domain to agonist binding, which is in agreement with the currently accepted mechanism of PPARα activation.

### 3.2. Binding energies

The experimental activities of the considered fibrates [13,35,36], expressed in terms of their $EC_{50}$ values through $-\log_{10} EC_{50}$ ($pEC_{50}$), were used to develop predictive model equations by means of linear regression approaches. As mentioned above, $EC_{50}$ encodes for the biological response to PPARα activation but is not directly connected to PPARα–fibrate binding strength, which is usually represented by the binding affinity constant $K_A$. However, $EC_{50}$ and $K_A$ may often be reciprocally connected because until bound to an agonist, the target structure will be responsive to an activation signal and *vice versa*. According to the operational model of pharmacological agonism [37], $EC_{50}$ (or $[A_{50}]$ in Ref. [37]) and $K_A$ are connected through the following equation:

$$EC_{50} = \frac{K_A}{((2 + \tau^n)^{1/n} - 1)}. \tag{1}$$

The $\tau$ term is called the transducer ratio and expresses the operational efficacy for an agonist, whereas $n$ is connected to the shape of the dose–response curve. Eq. (1) can be rewritten in logarithmic form to make explicit the dependence of $EC_{50}$ on target–agonist binding, hence:

$$-\log_{10} EC_{50} = -\log_{10} K_A + \log_{10}((2 + \tau^n)^{1/n} - 1) \tag{2}$$

$$pEC_{50} = pK_A + \log((2 + \tau^n)^{1/n} - 1). \tag{3}$$

Considering $pK_A = \Delta G_A / 2.3RT$, where $\Delta G_A$ is the free energy for the target–agonist association, relation (4) can be written at constant temperature:

$$pEC_{50} \propto \Delta G_A + \log((2 + \tau^n)^{1/n} - 1). \tag{4}$$

$pEC_{50}$ values and target–agonist binding free energies can be considered in an approximately linear relationship when $\tau$ and $n$ values are similar for all ligands and, thus, the logarithm term variability in relation (4) is low.

We thus hypothesized a linear dependence between $pEC_{50}$ values and the calculated PPARα–fibrate binding energies by assuming that, as a consequence of the high structural similarity affecting the examined agonists, the logarithm term in relation (4) is approximately constant.

The regression equations were initially tested through a validation scheme that relies on to the grouping of ligand set in two subsets that share almost the same chemical information and thus describe a similar variance spread (see Section 2). The first subset, including ligands **1**–**10**, was used as a user-defined training subset (UDTS) undergoing the linear regression analysis; the remaining 13 ligands were used as a validation subset to feedback the model equations developed on the training set and to test the corresponding prediction accuracy.

The experimental data underwent simple linear regression analyses against the target–ligand interaction energies calculated by both docking ($G_{score}$), Model 1, and MM ($\Delta E_{ff}$), Model 2. Model 1 shows a relatively poor correlation with $R^2 = 0.38$ and $Q^2 = 0.40$ and a high $PE$ (see Table 2a). This was not surprising in view of the well-known poor performances of scoring functions in the evaluation of biological activity. Indeed, these functions are generally effective in selecting the correct poses or discriminating active from non-active compounds but are less reliable in ranking different activities in structurally and biologically similar ligands.

Model 2 shows an improved, but still unsatisfactory, correlation with $R^2 = 0.56$ and $Q^2 = 0.55$ and slightly more reliable predictions, with a $PE$ of about one.
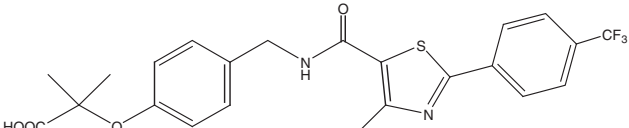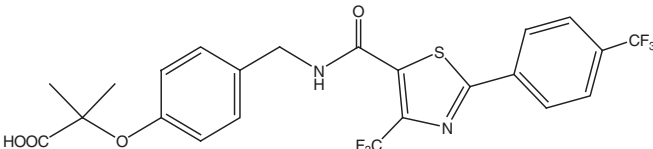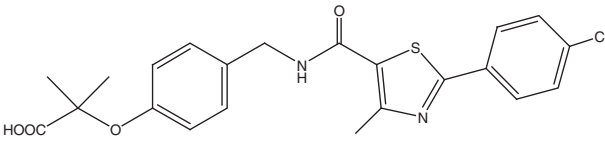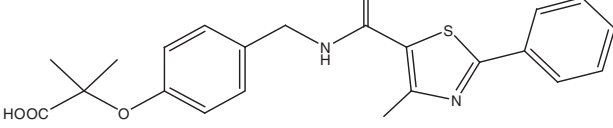
**Table 1**
Ligand set.

| Formula | Ligand | $EC_{50}$ ($\mu$M) | p$EC_{50}$ | Reference |
|---|---|---|---|---|
|  | **1** | 0.004 | 8.41 | [13] |
|  | **2** | 0.004 | 8.41 | [13] |
|  | **3** | 0.01 | 8.01 | [13] |
|  | **4** | 0.12 | 6.93 | [13] |
|  | **5** | 0.21 | 6.69 | [13] |
|  | **6** | 9.5 | 5.03 | [13] |
|  | **7** | 30 | 4.53 | [13] |
|  | **8** | 55 | 4.26 | [13] |
|  | **9** | 76 | 4.12 | [36] |
|  | **10** | 9 | 5.05 | [36] |
|  | **11** | 50 | 4.31 | [13] |
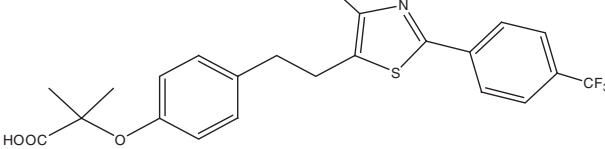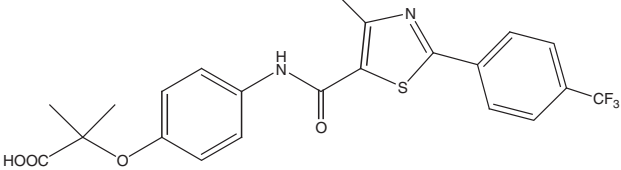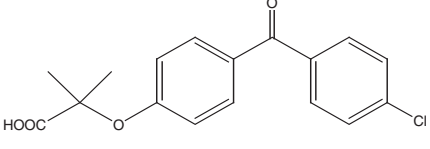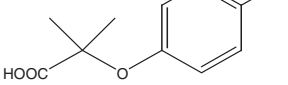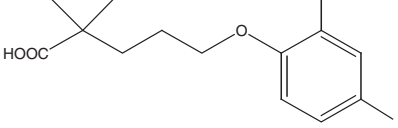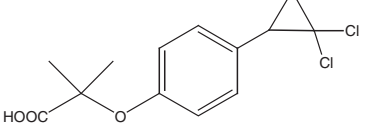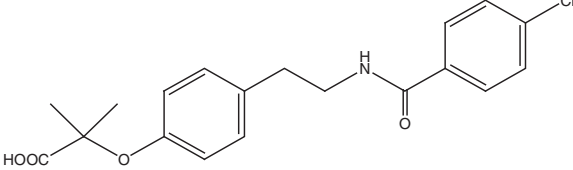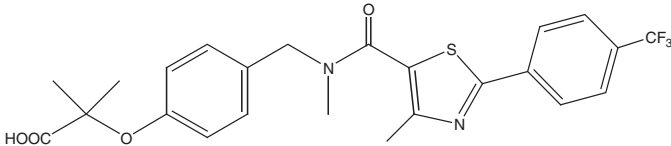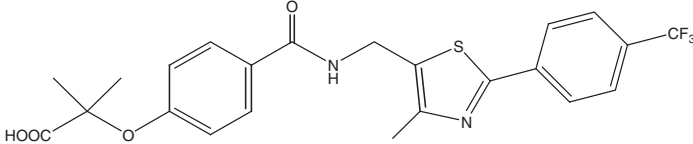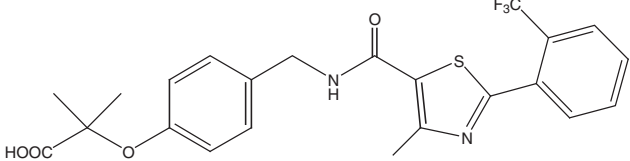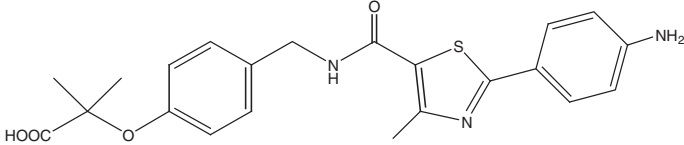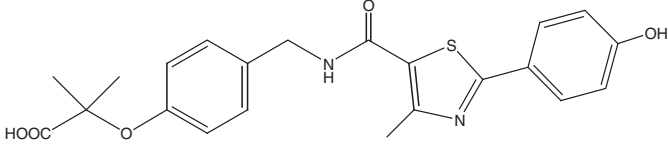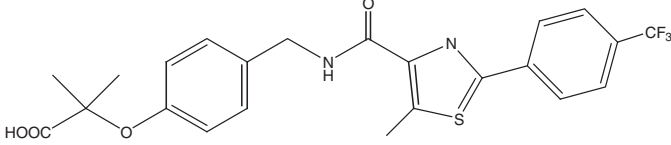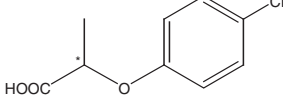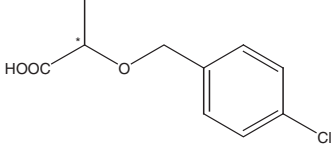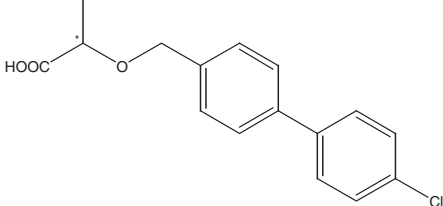
Table 1 (*Continued*)

| Formula | Ligand | $EC_{50}$ (μM) | $pEC_{50}$ | Reference |
|---------|--------|----------------|-----------|-----------|
| | **12** | 0.24 | 6.63 | [13] |
| | **13** | 1.32 | 5.89 | [13] |
| | **14** | 0.465 | 6.34 | [13] |
| | **15** | 5.64 | 5.25 | [13] |
| | **16** | 10 | 4.31 | [13] |
| | **17** | 1.89 | 6.63 | [13] |
| | **18** (*R*) | 520 | 3.29 | [35] |
| | **19** (*S*) | 950 | 3.03 | |
| | **20** (*R*) | 15 | 4.83 | [35] |
| | **21** (*S*) | 36 | 4.45 | |
| | **22** (*R*) | 54 | 4.27 | [36] |
| | **23** (*S*) | 420 | 4.38 | |

The $\Delta E_{ff}$ breakdown into the single non-bonded energy contributions was further considered:

$$\Delta E_{ff} = \Delta E_{els} + \Delta E_{VdW} + \Delta E_{HB} + \Delta G_{solv}. \quad (5)$$

The energy terms on the right in Eq. (5) were retrieved by recalculating the Amber force field energy of the relaxed bound complexes and the energy of the separated target and ligand in the presence of a polarizable continuum simulating the water environment (see Section 2). The binding process scheme assumed to derive Eq. (5) can be outlined through the Born–Haber cycle reported in Fig. 2.

$\Delta E_{ff}$ is an estimation of the receptor–ligand interaction in gas phase ($\Delta E_{int}^{gas}$) and the solvation free energy variation ($\Delta\Delta G_{solv}$ is reported as $\Delta G_{solv}$ in Eq. (5) for simplicity).

Therefore, a further energy contribution, $\Delta E_{ster}$, corresponding to the energy required by the ligand and receptor to reach their

**Table 2**
Model equations: (a) regressed by the UDTS procedure; (b) regressed by the AGTS procedure (see text).

| Model | Model coefficients | | | | | | | | Training set | | Whole set | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $b_0$ | $b_{score}$ | $b_{ff}$ | $b_{els}$ | $b_{VdW}$ | $b_{HB}$ | $b_{solv}$ | $b_{ster}$ | $R^2$ | $F$ | $Q^2$ | $PE$ |
| *(a)* | | | | | | | | | | | | |
| 1 | 3.40 | −0.665 | – | – | – | – | – | – | 0.38 | 6 | 0.40 | 1.16 |
| 2 | 2.28 | – | −0.0274 | – | – | – | – | – | 0.56 | 12 | 0.55 | 1.03 |
| 3 | 19.0 | – | – | 0.00776 | −0.0174 | −0.0958 | −0.0523 | −0.00741 | 0.94 | 27 | 0.64 | 0.709 |
| 4 | 17.9 | – | – | – | −0.0195 | – | −0.0416 | −0.0167 | 0.95 | 64 | 0.66 | 0.697 |
| 5 | 17.3 | – | – | – | −0.0183 | – | −0.0400 | – | 0.95 | 86 | 0.64 | 0.719 |
| *(b)* | | | | | | | | | | | | |
| 6 | 3.23 | −0.470 | – | – | – | – | – | – | 0.38 | 11 | 0.44 | 1.33 |
| 7 | 2.04 | – | −0.0249 | – | – | – | – | – | 0.45 | 14 | 0.56 | 1.03 |
| 8 | 11.7 | – | – | 0.00381 | −0.0337 | −0.656 | −0.0333 | −0.0288 | 0.86 | 20 | 0.77 | 0.491 |
| 9 | 17.5 | – | – | – | −0.0300 | – | −0.0451 | −0.0637 | 0.85 | 32 | 0.71 | 0.512 |
| 10 | 12.0 | – | – | – | −0.0281 | – | −0.0297 | – | 0.93 | 100 | 0.74 | 0.581 |
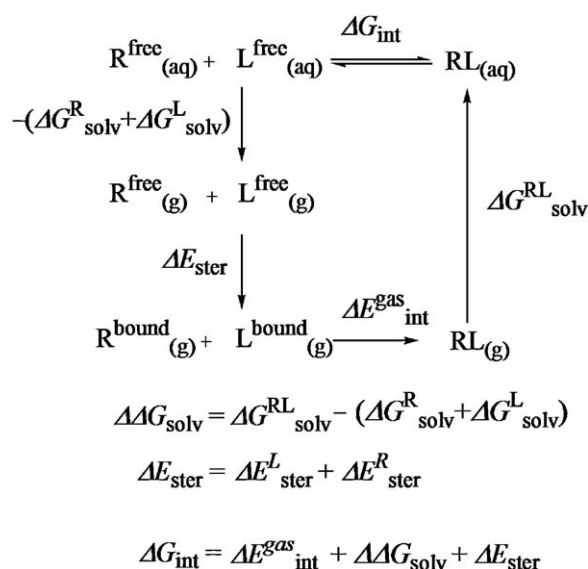


**Fig. 2.** Receptor–ligand binding process and contributions to binding energy.

binding conformation from the lowest conformation in solution, was also included in Model 3. In this case, we assumed that $\Delta E_{ster}^{R}$ is approximately constant because all the considered fibrates are expected to induce almost similar PPAR conformations; almost all variability of $\Delta E_{ster}$ comes from the ligand term $\Delta E_{ster}^{L}$, calculated as the difference between the force field energy of the binding and the global minimum conformation in solution. For this reason, the development of Model 3 (and Model 8, see below) was performed by the implicit assumption $\Delta E_{ster} \approx \Delta E_{ster}^{R}$, which corresponds to including $\Delta E_{ster}^{R}$ in the intercept constant $b_0$.

The force field contributions $\Delta E_{els}$, $\Delta E_{VdW}$, and $\Delta E_{HB}$, whose sum ($\Delta E_{ff}$) is an estimate of the receptor–ligand interaction energy, and the $\Delta G_{solv}$ and $\Delta E_{ster}$ contributions were then employed to predict PPARα agonist activity through a multivariate regression analysis. Model 3 was developed through the regression of the experimental activities against all five contributions to binding energy. The regression analysis leads to a reparameterization of the calculated binding energy in which the regression coefficients ($b$ in Table 2) represent the weights of each energy contribution.

Compared to Model 1 and 2, Model 3 shows an improved correlation and statistical significance, as indicated by the higher values of both $R^2 = 0.94$ and $F = 27$ for the training set and $Q^2 = 0.64$ and $PE = 0.7$ for the whole set (see Table 2a). However, further analyses showed that both $\Delta E_{els}$ and $\Delta E_{HB}$ regression coefficients were affected by low statistical significance such that by application of

the BVS method, these regressors were subsequently dropped off in the next models. The low significance of both the $\Delta E_{els}$ and $\Delta E_{HB}$ coefficients is somewhat surprising and worthy of further analyses because of the recognized importance of these interactions in the PPARα complexes of these ligands. In particular, the cross-correlation between the single energy contributions and their corresponding degree of variability was analyzed. The analysis of the cross-correlation between the single energy contributions to binding energies (See Table S2) indicated that $\Delta E_{els}$ and $\Delta G_{solv}$ are not linearly independent, showing a high correlation between these two contributions. In these circumstances, the variance encoded by one term can in fact be described by the other, i.e., the two dependent variables can be resolved to a single variable dependence. On the other hand, $\Delta E_{HB}$ was the only analyzed variable affected by a low degree of variability, in agreement with the hypothesis that agonist signaling is due to a well-conserved hydrogen-bond pattern between the carboxylate head of fibrate agonists and the residues previously mentioned in the PPARα binding site. Implicitly, the negligible correlation of the H-bond term indicates that, inside the analyzed set of fibrates, the modulation of agonist potency is mainly controlled by steric/hydrophobic (VdW term) and electrostatic (solvation term) effects.

The removal of $\Delta E_{els}$ and $\Delta E_{HB}$ from the regression gave rise to Model 4 and Model 5 by further BVS removal of $\Delta E_{ster}$, which also had low statistical significance. The latter model shows the best results, with a high correlation and significance in the training set ($R^2 = 0.95$ and $F = 86$) and good performances in the prediction. The predictive performances of Model 5 are excellent and comparable with those reported for different protein–ligands systems by others [21,22]. Predictive models for the estimation of binding free energies for a series of neuraminidase inhibitors were obtained by Monte Carlo simulations and through the application of the LIE method [21]; the best performing three-variable model, including $\Delta E_{els}$, $\Delta E_{vdW}$, and $\Delta SASA_{prot}$ (solvent-accessible surface area from bound to unbound protein), was characterized by a $Q^2$ of 0.62. Recently, a modified LIE method based on continuum solvation has been used to develop predictive models of the HIV-1RT-inhibitors binding free energies [22], with a cross-validation correlation coefficient of 0.66 together with an overall RMS error of 1.3 kcal mol$^{-1}$ for a series of 20 ligands. However, in both reported instances, the estimate of binding free energy contributions are calculated from either stochastic or molecular dynamic simulations that typically require a long calculation time. In contrast, our approach, ensuring almost comparable performances in a much shorter time, would more suitably support the rational drug design of new fibrates.

A second set of linear regression analyses was then performed through an alternative validation scheme, AGTS, consisting of an iterative leave-one-out procedure that filters out the trial point at each iteration, giving rise to minimal prediction error.
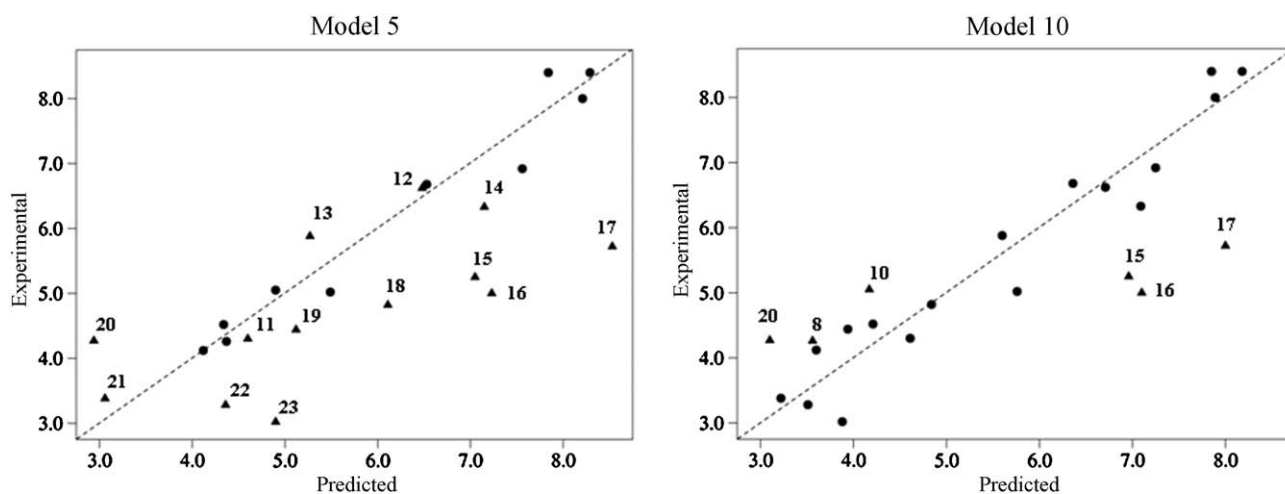
**Fig. 3.** Correlation between experimental and predicted agonist activities obtained by UDTS (Model 5) and AGTS (Model 10) validation. Full circle, training set; triangle, validation set.

Starting from the whole set of 23 trail points, the iterative leave-one-out procedure produced improved models after 5 iterations, leaving a training set with 17 trails and a validation set with 6 trials. The results of these cross-validated models (models 6–10) are reported in Table 2b.

These models share the UDTS procedure equation form and lead to similar values of the corresponding regression coefficients, indicating that both validation schemes convey similar predictive models.

In particular, the multivariate regression models 8-10 were again the best performing, with a high correlation and significance in the training set ($R^2 = 0.86$–$0.93$ and $F = 20$–$100$) and good performances in the prediction set ($Q^2 = 0.71$–$0.77$ and $PE = 0.49$–$0.58$).

The slightly better performances of the AGTS models are mainly due to a higher number of trials in the training sets, which allowed the inclusion of a larger amount of variance in the resulting models.

As shown in Fig. 3, the comparison of the best models obtained by the UDTS and AGTS procedures indicates Model 10 with only three outliers, i.e. **15**, **16**, and **17**. These ligands are structurally similar to the most active ligand **1**, but differ by either the presence of electron donor substituents, i.e., $NH_2$ and OH in **15** and **16**, replacing the electron withdrawing terminal $CF_3$ group, or the interchanged position of N and S in the thiazole ring in ligand **17**. The molecular modifications yielding these three ligands from **1** are expected to affect their agonist potencies through electronic rather than steric effects, and for the same reason, they cannot be explicitly treated by MM-based approaches.

Thus, we found that AGTS is able to filter out as outliers those fibrates whose binding is electronically rather than sterically affected; in this respect, AGTS allowed the building of a specialized model that performs better in predicting modulations of PPARα agonism of a steric nature, such as the inclusion of hindering substituents, variations in the spacer structure and increased sizes of hydrophobic moieties.

## 4. Conclusions

In the present investigation, a combination of docking and MM calculations was used to estimate the binding geometries and the target–ligand interaction energies for a series of PPARα fibrate agonists. Computational approaches based on the refinement of docking binding geometries through further MM relaxation of the target–ligand complexes provided for a fair description of the adaptation of PPARα to fibrate binding and clearly identified the AF2 backbone with the protein domain as being the most responsive to fibrate binding.

The considered fibrates showed similar structural patterns in the binding to PPARα, in which the carboxylate head interactions with the AF2 domain was the most conserved, which is in agreement with the role played by these interactions in the mechanism of PPARα agonism. The distal aryl moieties made contacts with several groups of binding site residues, mainly in the large distal hydrophobic pocket but also, in some cases, with an alternative hydrophobic site closer to the AF2 domain. The reciprocal orientation of the ligand hydrophobic tail and carboxylate head takes the major role in the modulation of PPARα–fibrate binding strength.

Model equations for the prediction of PPARα agonism were obtained by means of both monovariate and multivariate regression approaches. Among the calculated energy contributions to the PPARα–fibrate binding energy, van der Waals and solvation were found to be linearly independent and thus were only significantly correlated to $pEC_{50}$ in the best models. AGTS was the most effective regression procedure, as it was able to yield more accurate predictions of fibrate $EC_{50}$ values. Moreover, this procedure was able to automatically filter out entries characterized by a diverse (outlying) modulation between binding and agonist potency. In particular, Model 10 was more specialized in the ranking of fibrate agonists whose binding is mainly controlled by steric rather than by electronic modulation. The potential employment of Model 10 to gain a preliminary insight of the agonist potency of newly designed fibrates is currently under investigation.

## Acknowledgements

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmgm.2011.02.002.

## References

[1] I. Issemann, S. Green, Activation of a member of the steroid hormone receptor superfamily by peroxisome proliferators, Nature 347 (1990) 645–650.

[2] B.P. Kota, T.H.-W. Huang, B.D. Roufogalis, An overview on biological mechanisms of PPARs, Pharmacol. Res. 51 (2005) 85–94.

[3] (a) B.M. Forman, J. Chen, R.M. Evans, Hypolipidemic drugs, polyunsaturated fatty acids, and eicosanoids are ligands for peroxisome proliferator-activated receptors α and δ, Proc. Natl. Acad. Sci. U.S.A. 94 (1997) 4312–4317;
(b) S.A. Kliewer, S.S. Sundseth, S.A. Jones, P.J. Brown, G. Bruce Wisely, C.S. Koble, P. Devchand, W. Wahli, T.M. Willson, J.M. Lenhard, J.M. Lehmann, Fatty acids and eicosanoids regulate gene expression through direct interactions with peroxisome proliferator-activated receptors α and·γ, Proc. Natl. Acad. Sci. U.S.A. 94 (1997) 4318–4323.

[4] C. Duval, M. Müller, S. Kersten, PPARα and dyslipidemia, Biochim. Biophy. Acta 1771 (2007) 961–971.

[5] (a) D.H. van Raalte, M. Li, P.H. Pritchard, K.M. Wasan, Peroxisome proliferator-activated receptor (PPAR)-alpha: a pharmacological target with a promising future, Pharmaceut. Res. 21 (2004) 1531;
(b) Z. Israelian-Konaraki, P.D. Reaven, Peroxisome proliferator-activated receptor-alpha and atherosclerosis: from basic mechanisms to clinical implications, Cardiology 103 (2005) 1–9;
(c) P. Lefebvre, G. Chinetti, J.C. Fruchart, B. Bart Staels, Sorting out the roles of PPARα in energy metabolism and vascular homeostasis, J. Clin. Invest. 116 (2006) 571–580;
(d) J.C. Fruchart, Novel Peroxisome proliferator activated receptor-α agonists, Am. J. Cardiol. 100 (2007) S41–S46;
(e) A.Y.Y. Cheng, L.A. Leiter, PPAR-alpha: therapeutic role in diabetes-related cardiovascular disease, Diabetes Obes. Metab. 9 (2008) 691–698;
(f) B. Staels, M. Maes, A. Zambon, Fibrates and future PPARα agonists in the treatment of cardiovascular disease, Nat. Clin. Pract. Cardiovasc. Med. 5 (2008) 542–553.

[6] (a) K. Schoonjans, B.M. Staels, J. Auwerx, Role of the peroxisome proliferator-activated receptor (PPAR) in mediating the effects of fibrates and fatty acids on gene expression, J. Lipid Res. 37 (1996) 907–925;
(b) B. Staels, J. Auwerx, Role of PPAR in the pharmacological regulation of lipoprotein metabolism by fibrates and thiazolidinediones, Curr. Pharm. Design 3 (1997) 1–14;
(c) J.C. Fruchart, P. Duriez, B. Staels, Molecular mechanism of action of the fibrates, J. Soc. Biol. 193 (1999) 67–75;
(d) M.J. Chapman, Fibrates in 2003: therapeutic action in atherogenic dyslipidaemia and future perspectives, Atherosclerosis 171 (2003) 1–13;
(e) G. Chinetti-Gbaguidi, J.C. Fruchart, B. Staels, Pleiotropic effects of fibrates, Curr. Atheroscler. Rep. 7 (2005) 396–401;
(f) P.J. Barter, K.-A. Rye, Is there a role for fibrates in the management of dyslipidemia in the metabolic syndrome? Arteriosclerosis, Thrombosis, and Vascular Biology 28 (2008) 39–46.

[7] (a) P.J. Brown, L.W. Stuart, K.P. Hurley, M.C. Lewis, D.A. Winegar, J.G. Wilson, W.O. Wilkison, O.R. Ittoop, T.M. Willson, Identification of a subtype selective human PPAR agonist through parallel-array synthesis, Bioorg. Med. Chem. Lett. 11 (2001) 1225–1227.

[8] (a) M. Nomura, T. Tanase, T. Ide, M. Tsunoda, M. Suzuki, H. Uchiki, K. Murakami, H. Miyachi, Design, synthesis and evaluation of substituted phenylpropanoic acid derivatives as human peroxisome proliferator activated receptor activators. Discovery of potent and human peroxisome proliferator activated receptor α subtype-selective activators, J. Med. Chem. 46 (2003) 3581–3599;
(b) H. Miyachi, H. Uchiki, Bioorg. Med. Chem. Lett. 13 (2003) 3145–3149.

[9] Y. Xu, D. Mayhugh, A. Saeed, X. Wang, R.C. Thompson, S.J. Dominianni, R.F. Kauffman, J. Singh, J.S. Bean, W.R. Bensch, R.W. Barr, J. Osborne, C. Montrose-Rafizadeh, R.W. Zink, N.P. Yumibe, N. Huang, D. Luffer-Atlas, D. Rungta, D.E. Maise, N.B. Mantlo, Design and synthesis of a potent and selective triazolone-based peroxisome proliferator-activated receptor α agonist, J. Med. Chem. 46 (2003) 5121–5124.

[10] G.Q. Shi, J.F. Dropinski, Y. Zhang, C. Santini, S.P. Sahoo, J.P. Berger, K.L. MacNaul, G. Zhou, A. Agrawal, R. Alvaro, T.-q. Cai, M. Hernandez, S.D. Wright, D.E. Moller, J.V. Heck, P.T. Meinke, Novel 2,3-dihydrobenzofuran-2-carboxylic acids: highly potent and subtype-selective PPARα agonists with potent hypolipidemic activity, J. Med. Chem. 48 (2005) 5589–5599.

[11] J.M. Matthews, X. Chen, E. Cryan, D.J. Hlasta, P.J. Rybczynski, K. Strauss, Y. Tang, J.Z. Xu, M. Yang, L. Zhou, K.T. Demarest, Design and synthesis of indane-ureido-thioisobutyric acids: a novel class of PPARα agonists, Bioorg. Med. Chem. Lett. 17 (2007) 6773–6778.

[12] (a) T. Asaki, T. Aoki, T. Hamamoto, Y. Sugiyama, S. Ohmachi, K. Kuwabara, K. Murakami, M. Todo, Structure–activity studies on 1,3-dioxane-2-carboxylic acid derivatives, a novel class of subtype-selective peroxisome proliferator-activated receptor α (PPARα) agonists, Bioorg. Med. Chem. 16 (2008) 981–994;
(b) T. Aoki, T. Asaki, T. Hamamoto, Y. Sugiyama, S. Ohmachi, K. Kuwabara, K. Murakami, M. Todo, Discovery of a novel class of 1,3-dioxane-2-carboxylic acid derivatives as subtype-selective peroxisome proliferator-activated receptor α (PPARα) agonists, Bioorg. Med. Chem. Lett. 18 (2008) 2128–2132.

[13] M.L. Sierra, V. Beneton, A.B. Boullay, T. Boyer, A. Brewster, F. Donche, M.C. Forest, M.H. Fouchet, F.J. Gellibert, D.A. Grillot, M.H. Lambert, A. Laroze, C.L. Grumelec, J.M. Linget, V.G. Montana, V.L. Nguyen, E. Nicodeme, V. Patel, A. Penfornis, O. Pineau, D. Pohin, F. Potvain, G. Paulain, C.B. Ruault, M. Saunders, J. Toum, X.E. Xu, R.X. Xu, P.M. Pianetti, J. Med. Chem. 50 (2007) 685.

[14] H.E. Xu, T.B. Stanley, V.G. Montana, M.H. Lambert, B.G. Shearer, J.E. Cobb, D.D. Mckee, C.M. Galardi, K.D. Plunket, R.T. Nolte, D.J. Parks, J.T. Moore, S.A. Kliewer, T.M. Willson, J.B. Stimmel, Structural basis for antagonist-mediated recruitment of nuclear co-repressors by PPARalpha, Nature 415 (2002) 813.

[15] P. Cronet, J.F. Petersen, R. Folmer, N. Blomberg, K. Sjoblom, U. Karlsson, E.L. Lindstedt, K. Bamberg, Structure of the PPARalpha and -gamma ligand binding domain in complex with AZ 242; ligand selectivity and agonist activation in the PPAR family, Structure 9 (2001) 699–706.

[16] H.E. Xu, M.H. Lambert, V.G. Montana, K.D. Plunket, L.B. Moore, J.L. Collins, J.A. Oplinger, S.A. Kliewer, R.T. Gampe Jr., D.D. McKee, J.T. Moore, T.M. Willson, Structural determinants of ligand binding selectivity between the peroxisome proliferator-activated receptors, Proc. Natl. Acad. Sci. U.S.A. 98 (2001) 13919–13924.

[17] A.R. Leach, Molecular Modelling: Principles and Applications, Addison Wesley Longman Limited, Harlow, 1996.

[18] S.F. Sousa, P.A. Fernandes, M.J. Ramos, Proteins 65 (2006) 15–26.

[19] (a) R.A. Friesner, J.L. Banks, R.B. Murphy, T.A. Halgren, J.J. Klicic, D.T. Mainz, M.P. Repasky, E.H. Knoll, M. Shelley, J.K. Perry, D.E. Shaw, P. Francis, P.S. Shenkin, Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy, J. Med. Chem. 47 (2004) 1739–1749;
(b) T.A. Halgren, R.B. Murphy, R.A. Friesner, H.S. Beard, L.L. Frye, W.T. Pollard, J.L. Banks, Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening, J. Med. Chem. 47 (2004) 1750–1759;
(c) G. Jones, P. Willett, R.C. Glen, J. Mol. Biol. 245 (1995) 43–53;
(d) M.L. Verdonk, J.C. Cole, M.J. Hartshorn, C.W. Murray, R.D. Taylor, Proteins 52 (2003) 609–623;
(e) G.M. Morris, R. Huey, W. Lindstrom, M.F. Sanner, R.K. Belew, D.S. Goodsell, A.J. Olson, J. Comput. Chem. 30 (2009) 2785–2791.

[20] (a) J. Åqvist, V.B. Luzhkov, B.O. Brandsdal, Ligand binding affinities from MD simulations, Acc. Chem. Res. 35 (2002) 358–365;
(b) M. Almlöf, J. Carlsson, J. Åqvist, Improving the accuracy of the linear interaction energy method for solvation free energies, J. Chem. Theory Comp. 3 (2007) 2162–2175;
(c) E. Stjernschantz, J. Marelius, C. Medina, M. Jacobsson, N.P.E. Vermeulen, C. Oostenbrink, Are automated molecular dynamics simulations and binding free energy calculations realistic tools in lead optimization? An evaluation of the linear interaction energy (LIE) method, J. Chem. Inf. Model. 46 (2006) 1972–1983.

[21] (a) T. Hou, S. Guo, X. Xu, Predictions of binding of a diverse set of ligands to gelatinase-A by a combination of molecular dynamics and continuum solvent models, J. Phys. Chem. B 106 (2002) 5527–5535;
(b) A. Bortolato, S. Moro, In silico binding free energy predictability by using the linear interaction energy (LIE) method: bromobenzimidazole CK2 inhibitors as a case study, J. Chem. Inf. Model. 47 (2007) 572–582;
(c) M. Naïm, S. Bhat, K.N. Rankin, S. Dennis, S.F. Chowdhury, I. Siddiqi, P. Drabik, T. Sulea, C.I. Bayly, A. Jakalian, E.O. Purisima, Solvated interaction energy (SIE) for scoring protein–ligand binding affinities. 1. Exploring the parameter space, J. Chem. Inf. Model. 47 (2007) 122–133;
(d) C.-E. Chang, M.K. Gilson, Free energy, entropy, and induced fit in host–guest recognition: calculations with the second-generation mining minima algorithm, J. Am. Chem. Soc. 126 (2004) 13156–13164.

[22] B. Kuhn, P.A. Kollman, Binding of a diverse set of ligands to avidin and streptavidin: an accurate quantitative prediction of their relative affinities by a combination of molecular mechanics and continuum solvent models, J. Med. Chem. 43 (2000) 3786–3791.

[23] (a) N. Huang, C. Kalyanaraman, K. Bernacki, M.P. Jacobson, Molecular mechanics methods for predicting protein–ligand binding, Phys. Chem. Chem. Phys. 8 (2006) 5166–5177;
(b) C.R. Guimaraes, M. Cardozo, MM.BB/SA rescoring of docking poses in structure-based lead optimization, J. Chem. Inf. Model. 48 (2008) 958–970;
(c) M.R. Lee, Y. Sun, Improving docking accuracy through molecular mechanics generalized Born optimization and scoring, J. Chem. Theory Comput. 3 (2007) 1106–1119;
(d) D. Huang, A. Caflish, Efficient evaluation of binding free energy using continuum electrostatic solvation, J. Med. Chem. 47 (2004) 5791–5797.

[24] (a) T. Kaya, S.C. Mohr, D.J. Waxman, S. Vajda, Computational screening of phthalate monoesters for binding to PPARγ, Chem. Res. Toxicol. 19 (2006) 999–1009;
(b) X. Li, H. Zou, A. Wu, Y. Ye, J. Shen, Structure-based drug design of a novel family of chalcones as PPARα agonists: virtual screening, synthesis, and biological activities in vitro, Acta Pharmacol. Sin. 28 (2007) 2040–2052.

[25] Maestro, Version 8.5, Schrodinger, LLC, New York, NY, 2008.

[26] W.L. DeLano, The PyMOL Molecular Graphics System, DeLano Scientific, Palo Alto, CA, USA, 2002, http://www.pymol.org.

[27] MacroModel, Version 9.0, Schrödinger, LLC, New York, NY, 2005.

[28] Glide, Version 3.5, Schrödinger, LLC, New York, NY, 2005.

[29] R.A. Friesner, R.B. Murphy, M.P. Repasky, L.L. Frye, J.R. Greenwood, T.A. Halgren, P.C. Sanschagrin, D.T. Mainz, Extra precision Glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes, J. Med. Chem. 49 (21) (2006) 6177–6196.

[30] (a) W.L. Jorgensen, D.S. Maxwell, J. Tirado-Rives, Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids, J. Am. Chem. Soc. 118 (1996) 11225–11236;
(b) G.A. Kaminski, R.A. Friesner, J. Tirado-Rives, W.L. Jorgensen, Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides, J. Phys. Chem. B 105 (2001) 6474–6487.

[31] (a) S.J. Weiner, P.A. Kollman, D.A. Case, U.C. Singh, C. Ghio, G. Alagona, S. Profeta, P. Weiner, A new force field for molecular mechanical simulation of nucleic acids and proteins, J. Am. Chem. Soc. 106 (1984) 765–784;

(b) S.J. Weiner, P.A. Kollman, D.T. Nguyen, D.A. Case, An all force field for simulations of proteins and nucleic acids, J. Comput. Chem. 7 (1986) 230–252.

[32] W.C. Still, A. Tempczyk, R.C. Hawley, T. Hendrickson, Semianalytical treatment of solvation for molecular mechanics and dynamics, J. Am. Chem. Soc. 112 (1990) 6127–6129.

[33] (a) P. Lefebvre, G. Chinetti, J.C. Fruchart, B. Staels, J. Clin. Invest. 116 (2006) 571–580;
(b) B. Staels, J. Dallongeville, J. Auwerx, K. Schoonjans, E. Leitersdorf, J.C. Fruchart, Mechanism of action of fibrates on lipid and lipoprotein metabolism, Circulation 98 (1998) 2088–2093.

[34] B. Staels, J.C. Fruchart, Therapeutic roles of peroxisome proliferator-activated receptor agonists, Diabetes 54 (2005) 2460–2470.

[35] A. Pinelli, C. Godio, A. Laghezza, N. Mitro, G. Fracchiolla, V. Tortorella, A. Lavecchia, E. Novellino, J.-C. Fruchart, B. Staels, M. Crestani, F. Loiodice, Synthesis, biological evaluation, and molecular modeling investigation of new chiral fibrates with PPARα and PPARγ agonist activity, J. Med. Chem. 48 (2005) 5509–5519.

[36] R. Mukherjee, S. Sun, L. Santomenna, B. Miao, H. Walton, B. Liao, K. Locke, J.-H. Zhang, S.H. Nguyen, L.T. Zhang, K. Murphy, H.O. Ross, M.X. Xia, C. Teleha, S.-Y. Chen, B. Selling, R. Wynn, T. Burn, P.R. Young, Ligand and coactivator recruitment preferences of peroxisome proliferator activated receptor α, J. Steroid Biochem. Mol. Biol. 81 (2002) 217–225.

[37] J.W. Black, P. Leff, N.P. Shankley, J. Wood, An operational model of pharmacological agonism: the effect of E/[A] curve shape on agonist dissociation constant estimation, Br. J. Pharmacol. 84 (1985) 561–571.