

CALCULATION OF THE BINDING AFFINITY OF HUMAN PPAR γ ACTIVATORS USING THE LINEAR INTERACTION APPROXIMATION

Elena Fioravanzo[§], Andrea Ciacci[†], Natalina Dell'Uomo[†], Emanuela Tassoni[†], Grazia Gallo[†], Fabio Giannessi[†],

Massimo Mabilia[§], Marco Daniele Parenti[§] and Paolo Carminati[†]

[†] Direzione Ricerca e Sviluppo Sigma-Tau S.p.A., via Pontina km 30,400, I-00040 Pomezia (RM), Italy

[§] S.I.N - Soluzioni Informatiche S.a.s., via Salvemini 9, I-36100 Vicenza, Italy



INTRODUCTION

Peroxisome proliferator-activated receptors (PPARs) belong to the nuclear receptor superfamily of ligand-activated transcription factors. There are three known subtypes, namely, PPAR α , PPAR γ , and PPAR δ distributed among the different tissues.

The recent discovery that PPAR α and PPAR γ are the primary target for thiazolidinediones (TZDs) and fibrates respectively, and the rapid progress in functional analysis of PPARs, have established that they play a central role in regulating carbohydrates and lipid metabolism in animals and in humans.

Compounds can be identified and studied, that have both PPAR α and PPAR γ agonist properties (PPAR α/γ dual agonists), which combine the benefits of a TZD plus a fibrate in a single molecule. A number of PPAR α/γ dual agonists have appeared in the literature, such as KRP-297, tesaglitazar, ragaglitazar, LY-510929, GW-409544 and muraglitazar. Nevertheless, up to-day an optimal ratio of PPAR α and PPAR γ agonist activity has not been identified because of the species differences in lipid and lipoprotein metabolism, in PPARs distribution, in PPAR-responsive genes regulation and in drug pharmacokinetic properties and metabolism.

We identified and tested for their potential anticatabolic/hypolipemic effect, two different classes of synthetic PPAR γ ligands: substituted 2-phenylthio-2-(methyl)-propanoic acid (thio- isobutyrate) and 2-ethoxy-3-phenyl-propanoic acid (a-ethoxy-propionate) derivatives [1].

Accurate ranking of binding affinities is crucial in the lead optimization phase of pharmaceutical research in order to develop potent, effective drug candidates. The ability to predict ligand-protein binding affinities has obvious utility in drug discovery. Thus, several methods have been developed ranging from rapid QSAR-based scoring functions to computationally intensive free energy perturbation (FEP) calculations for addressing this difficult problem [2]. Up to now, any of these approaches hasn't fully met the needs of researcher and developers. QSAR-type approaches, though rapid, involve many approximations and produce large errors in binding energy predictions. FEP approaches are more accurate, but cannot be used when ligand structures vary significantly. They also incur substantial CPU costs. Linear interaction approximation (LIA) [3] is a way of combining molecular mechanics calculations with experimental data to build a model scoring function for the evaluation of ligand-protein binding free energies and represents a good compromise between accuracy and computational speed. In this study, docking and LIA calculations were applied to build a predictive model scoring function using a set of human PPAR γ activators.

COMPUTATIONAL METHODS

Data Sets

Table 1: Ligand Structures co-crystallized with PPAR γ				
Structure	Name	pK _i	Structure	Name
	Ragaglitazar			GI-262570
	AZ-242			

Table 2 (A): Thio-isobutyrate - para substitution				
Structure	Name	pK _i	Structure	Name
	ST2740	5.0		ST2733
	ST2652	4.31		ST2652

Table 2 (B): Thio-isobutyrate - meta substitution				
Structure	Name	pK _i	Structure	Name
	ST2753	4.65		ST2653
	ST2618	5.89		ST2609
	ST2518	4.55		ST2651

Table 2 (C): a-ethoxy-propionates				
Structure	Name	pK _i	Structure	Name
	ST2672 (B)	4.56		ST2609 (B)
	Tesaglitazar (S) ST2751	4.80		

As computed binding affinities depend mainly on 3D structures, three different X-ray crystal structures of PPAR γ were considered, 1N1X, 1FM9 and 1I71, respectively co-crystallized with ragaglitazar [4], GI-262570 [5] and AZ-242 (tesaglitazar)[6] (Table 1 and Figure 1)

We considered 13 compounds, members of two different structural class (Table 2), whose in vitro agonist activity (pK_i) on mouse PPAR γ was determined.[1] 10 compounds are para or meta substituted thio-isobutyrate (Table 2 A and B) and 3 compounds are meta substituted a-ethoxy-propionates (Table 2 C). To better understand the structure-activity relationship and to enhance selectivity vs PPAR α isoform, these molecules were docked on the three X-ray crystal structures of PPAR γ using the software GLIDE (grid-based ligand docking with energetics) [7]. Glide has been designed to perform as close to an exhaustive search of the positional, orientational, and conformational space available to the ligand as is feasible while retaining sufficient computational speed. Glide uses a series of hierarchical filters to search for possible locations of the ligand in the active-site region of the receptor. The shape and properties of the receptor are represented on a grid by different sets of fields that provide progressively more accurate scoring of the ligand pose.

We generated 10 poses for each ligand docked on each receptor structure. By means of a visual inspection the best pose for each ligand in each receptor structure was selected. Resulting complex structures were used to build a binding affinity model employing five different descriptors implemented in Ligand & Structure-Based Descriptors (LSBD) module [8]. LSBD generates descriptors with the following programs: (1) MacroModel - Ligand binding affinities are rank-ordered or predicted with the aid of mBrAcE calculations; (2) Prime - MM-GBSA calculations; (3) Liaison - Linear interaction (LIA) descriptors and LiaisonScore empirical scoring function are calculated; (4) LigParse groups and atoms count (5) QikProp - over 35 pharmaceutically relevant ADME properties are calculated.

Figure 1: X-Ray structures of PPAR γ receptor, in complex with three different agonist



Liaison applies linear interaction approximation (LIA) to accurately compute binding affinities for series of ligands with similar binding modes. At the heart of this method is the assumption that free energy of binding can be derived from considering only the two end points of the thermodynamic cycle of ligand binding. Unlike FEP-based packages, Liaison can be applied to ligands with large variations in structures, allowing for a thorough investigation of modified leads. LIA only requires simulations of the ligand's bound and free states. The binding event is viewed as a replacement of the ligand's aqueous environment with a mixed aqueous/protein environment. Only interactions between the ligand and the protein or between the ligand and the aqueous environment enter into the quantities accumulated during the simulation. The protein-protein and protein-water interactions are considered part of the reference Hamiltonian

The Prime MM-GBSA approach is used to predict the free energy of binding for a receptor and a set of ligands. MM-GBSA is an acronym for a method that combines OPLS molecular mechanics energies, an SGB solvation model for polar solvation, and a nonpolar solvation term composed of the nonpolar solvent accessible surface area and van der Waals interactions. The total free energy of binding is then expressed as: $\Delta G_{bind} = G_{complex} - (G_{protein} + G_{ligand})$. The ligand in the unbound state is minimized in SGB solvent but is not otherwise sampled. In the calculation of the complex, the ligand is minimized in the context of the receptor. The protein is currently held fixed in all calculations.

mBrAcE calculates ligand-receptor binding energies by molecular mechanics energy minimization of the complex and the separated receptor and ligand, with or without continuum solvation. The calculation is performed first on the receptor, then on the ligand, and finally on the complex. The energy difference is then calculated using the equation: $\Delta E = E_{complex} - E_{ligand} - E_{protein}$. The full effects of relaxation and solvation are included in calculation.

QikProp produces a list of 44 descriptors related to absorption, distribution, metabolism and excretion. These descriptors include properties like skin permeability and octanol/water partition coefficients, and counts of important functional groups.

Ligparse generate counts of various substructures, and include counts of a wide range of functional groups (defined by single SMARTS pattern), counts of composite groups (defined by multiple SMARTS patterns), and some other counts. Over 100 functional groups are identified by ligparse.

Both PLS [9] and regression analysis [10] were used to investigate likely correlations between descriptors and experimental pK_i values. The optimal number of components in each PLS model was determined by SIMCA-P+ default cross-validation procedure. Variables selection was carried out on the basis of VIP parameter and coefficient values.

RESULTS

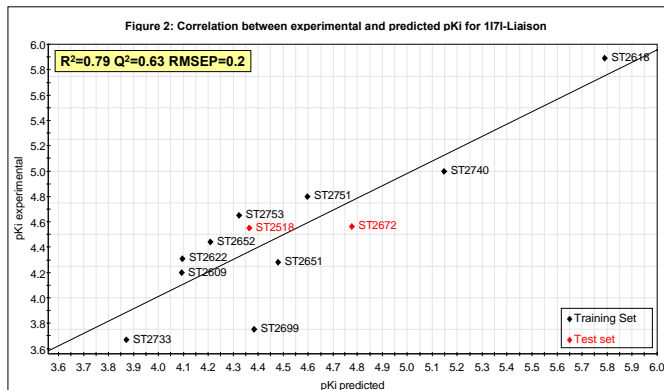


Table 3: 1I71-Liaison model datasets

Training set		
Name	pK _i exp	pK _i pred
ST2753	4.65	4.32
ST2751	4.8	4.6
ST2740	5	5.15
ST2733	3.67	3.87
ST2699	3.75	4.38
ST2652	4.44	4.21
ST2651	4.28	4.48
ST2622	4.31	4.1
ST2618	5.89	5.79
ST2609	4.2	4.09

Test set		
Name	pK _i exp	pK _i pred
ST2672	4.56	4.78
ST2518	4.55	4.36

For each receptor structure a single pose for each ligand was selected, so three different sets of docking poses were available. Three different LSBD descriptors matrices were calculated and PLS analysis was carried out on each matrix in order to generate QSAR models.

Structure ST2653 was detected as an outlier in all generated models, therefore this result depends neither on active site structure nor on nature of descriptors employed. In fact it is difficult to explain so different pK_i values, 6.85 for ST2653 and 4.55 for ST2518, on the basis of steric or electrostatic interactions in the active site if the chemical structure difference is a single chlorine atom. We are further studying ST2653 behavior looking, e.g., for a different binding mode. In the present study this molecule was excluded from the training set.

The final QSAR model was generated using a training set of 10 molecules and validated with a test set of 2 molecules (Table 3); best model was obtained employing Liaison descriptors and 1I71 receptor structure ($R^2=0.79$ $Q^2=0.63$ $RMSEP=0.2$, Figure 2).

Not surprisingly, the best results were obtained using the 3D structure of the receptor co-crystallized with tesaglitazar (1I71). In fact, this agonist is very similar to the molecules in our dataset.

CONCLUSIONS

In all generated models, GlideScore is always outperformed by Liaison descriptors in generating a model able to correlate with experimental data (data not shown). This is probably not a flaw unique to GlideScore but is more likely an indication of the poor performance that is common to most docking scoring functions. It is known that GlideScore has been primarily optimized to yield accurate binding poses, and in all cases here it gave optimal poses. This is an important task, since from our tests it is clear that a high degree of confidence is required in the binding modes being used as input for the Liaison protocol.

Although the very high level of complexity of the chosen target (e.g. wide and less-characterized active site, opening/closing of hydrophobic pocket, different 3D structure available, etc) and the lack of experimental data available, in particular for the test set, we obtained a fascinating model that highlight the capability of this method to reveal the hidden relationship between the agonist-receptor complex and the in vitro activity.

In conclusion, the protocol of using Glide for pose generation and Liaison protocol for rescoring appear promising for the application to structure-based lead optimization of chemical series for PPAR γ activators.

REFERENCE

1. Patent n°: WO 2003/059864 (alfa-ethoxy-propionates) and WO 2004/056355 (thio-isobutyrate); Dell'Uomo, N. et al. ChemMedChem 2006, 1, 49-53.
2. Lazaridis, T. Binding Affinity and Specificity from Computational Studies. Curr. Org. Chem. 2002, 6, 1319-1332.
3. Aquist, J.; Marelius, J. The Linear Interaction Energy Method for Predicting Ligand Binding Free Energies. Comb. Chem. High Throughput Screening 2001, 4, 613-626.
4. Ebdrup, S.; Petersson, L.; Rasmussen, H.B.; Deussen, H.-J.; Frost Jensen, A.; Mortensen, S.B.; Fleckner, J.; Pridal, L.; Nygaard, L.; Sauerberg, P. Synthesis and biological and structural characterization of the dual-acting peroxisome proliferator-activated receptor alpha/gamma agonist ragaglitazar. J.MED.CHEM. v46 pp.1306-1317, 2003
5. Gampe Jr., R.T.; Montana, V.G.; Lambert, M.H.; Miller, A.B.; Bledsoe, R.K.; Milburn, M.V.; Kiewer, S.A.; Willson, T.M.; Xu, H.E. Asymmetry in the PPARgamma/RXRalpha crystal structure reveals the molecular basis of heterodimerization among nuclear receptors. Mol.Cell v5 pp.545-555, 2000
6. Cronet, P.; Petersen, J.F.; Folmer, R.; Blomberg, N.; Sjoblom, K.; Karlsson, U.; Lindstedt, E.L.; Bamberg, K. Structure of the PPARalpha and -gamma ligand binding domain in complex with AZ 242; ligand selectivity and agonist activation in the PPAR family. Structure v9 pp.699-701, 2001
7. Glide, version 4.0, Schrödinger, LLC, New York, NY, 2005.
8. Schrödinger Suite 2006 Ligand and Structure-Based Descriptors protocol; Liaison version 4.0, Schrödinger, LLC, New York, NY, 2005; MacroModel version 9.1, Schrödinger, LLC, New York, NY, 2005; Prime version 1.5, Schrödinger, LLC, New York, NY, 2005; QikProp version 2.5, Schrödinger, LLC, New York, NY, 2005.
9. SIMCA-P+ Version 11.0.0.0, June 12 2005 (Umetrics AB)
10. Strike, version 1.5, Schrödinger, LLC, New York, NY, 2005.