Journal of Computer-Aided Molecular Design **17:** 785–796, 2003.

Peroxisome Proliferator-Activated Receptors target family landscape: A chemometrical approach to ligand selectivity based on protein binding site analysis

Bernard Pirard*
Aventis Pharma Deutschland GmbH, DI&A Chemistry, Computational Chemistry, D-65926 Frankfurt am Main, Germany
(Fax: ++49-69-331399, E-mail: bernard.pirard@aventis.com)

Received 20 May 2003; accepted in revised form 17 November 2003

Summary

The Peroxisome Proliferator-Activated Receptors (PPARs) are nuclear receptors which over the last couple of years have been the focus of considerable research efforts aiming to identify compounds with well-defined selectivity profiles for the treatment of metabolic diseases. The ligand binding domains (LBD) of the three known PPAR subtypes exhibit between 60 and 70% sequence identity. To gain insight into the structural determinants of selectivity for the PPAR subtypes, a set of 13 crystal structures of PPAR LBD were classified, using the GRID/CPCA approach. As a result, nearly all of the crystal structures of each different PPAR subtype were found clustered in different regions of the CPCA score plots, and hydrophobic as well as steric interactions were identified as the major determinants of PPAR subtypes selectivity. Furthermore, interpretation of the GRID/CPCA model in structural terms led to the identification of LBD regions which could be targeted to improve the selectivity for a given PPAR subtype. Our findings are consistent with published structure–activity relationships for PPAR ligands as well as with site-directed mutagenesis results.

Introduction

During the last decade, large protein families containing many pharmaceutically relevant targets such as the kinases, the proteases, the G protein coupled receptors, the ion channels, and the nuclear receptors have been identified. The nuclear receptors (NRs) act as ligand-activated transcription factors that regulate the expression of genes controlling vital processes like cell growth, differentiation and metabolism [1]. With few exceptions, the NRs, estimated to number 48 in the human genome, consist of six functional and structural domains (A–F). Two of them exhibit the highest degree of sequence conservation, namely region C, involved in binding to specific DNA sequences, and region E, which harbors functional components such as the ligand binding domain (LBD). Based on sequence alignments of the conserved domains, the NR family has been subdivided into seven subfamilies, such as the thyroid hormone-like and the estrogen-like receptors [2].

The Peroxisome Proliferator-Activated Receptors (PPARs), which were cloned a decade ago as orphan receptors, belong to the thyroid hormone-like subfamily of NRs [3, 4]. So far, three PPAR subtypes, produced by distinct genes and designated as PPARα, PPARγ and PPARδ, have been identified. The three PPAR subtypes exhibit the same domain organization as the other NR family members, with region C showing over 80% sequence identity and region E displaying between 60 and 70% sequence identity [3, 4]. Agonist binding to the PPAR LBD triggers a cascade of events including heterodimerization with the retinoid X receptor (RXR) and co-activator recruitment, which events lead to the activation of DNA transcription [3, 5]. The availability of potent and selective synthetic ligands has helped to uncover the specific function of each PPAR subtype in the lipid and lipoprotein homeostasis. The first subtype, PPARα,

promotes lipid catabolism in the liver and in skeletal muscles while PPARγ regulates fatty acids storage in adipocytes [1, 3, 5]. PPARα is the molecular target of the fibrate class of lipid lowering agents, and PPARγ is the receptor for the thiazolidinedione type of anti-diabetic drugs [1, 3, 6, 7]. The third isoform, PPARδ, which is not known as the target of any class of available drugs, has been recently identified as a regulator of inverse cholesterol transport and high-density lipoprotein metabolism [8].

Both selective and dual PPAR ligands are currently being explored in different pharmaceutical companies as potential therapies for several metabolic diseases such as diabetes and dyslipidemia [3, 9–14]. PPAR ligands may also be useful for other therapeutic areas including oncology and inflammation [3]. The availability of several crystal structures of ligands bound to the PPAR's LBD has provided valuable insights into the molecular mechanism of receptor activation as well as a basis for the rational design of novel ligands with altered receptor pharmacological profile. Specifically, the PPAR LBD exhibits a very similar fold to the other NRs structures comprising 12 α-helices arranged in three layers to form an anti-parallel 'helical sandwich' [15–17]. Within the central core of the PPAR's LBD is a T-shaped binding cavity flanked by helices 3, 5, 7, 11, and 12 with a volume of about 1300 Å^3. This cavity is two to three times larger than those observed in other NRs [5, 18–21]. Most of the known PPAR agonists which occupy only a fraction of the T-shaped binding cavity, share a hydrophilic head group involved in key hydrogen bonds (H bonds) with several protein side chains, a central hydrophobic part and a flexible linker to the tail (Scheme 1) [3, 18–21]. One of these key H bonds involves a tyrosine of the C terminal helix, helix 12, which upon agonist binding, adopts a conformation, favoring co-activator recruitment [21]. In crystal structures of agonist–PPAR complexes, helix 12 folds back over the ligand binding pocket, whereas helix 12 of PPARγ-*apo* structures can either extend away from the ligand binding pocket or fold back over this pocket. Hence, agonist binding locks helix 12 in an active conformation [21]. On the other hand, recent crystallographic data indicate that antagonist binding to PPARα prevents helix 12 from adopting the active position [22]. Despite the therapeutic potential of ligands with different PPAR subtype selectivity profiles, few studies have dealt with the structural determinants of receptor subtype selectivity. Recently, comparison of the LBD crystal structures of the three human PPAR subtypes together with site-directed mutagenesis has revealed that mutation of a single amino acid was sufficient to affect the subtype selectivity of several chemical classes of ligands [18, 23].

Herein, we describe the application of the GRID/CPCA [24] strategy to uncover selective interactions within the LBD of the three PPAR subtypes. GRID/CPCA enables one to characterize the common binding site of a protein target family in three-dimensional (3D) structural terms and to highlight regions of this binding site, which are amenable to selective interactions with a given type of ligand functional group. This approach was recently referred to as the 'target family landscape' approach [25].

Methods

Protein structures

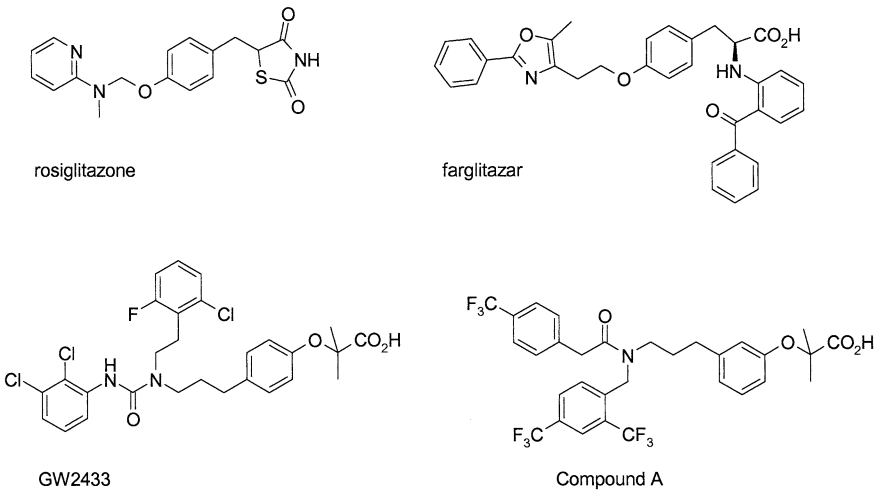
The X-ray crystal structure 3PRG of the human PPARγ LBD [26] was used as a query to search the protein-ligand database Relibase+ [27, 28] for crystal structures with resolution $\leq$3 Å and sharing at least a 40% sequence identity with the query. The result was 14 hits: three PPARα, eight PPARγ, and three PPARδ X-ray structures, as summarized in Table 1.

Alignment of 3D structures

Since the application of GRID/CPCA requires protein 3D structures to be aligned in a common reference frame, a global unbiased alignment based on secondary structures and structurally conserved regions (SCRs) was carried out for the 14 PPAR LBD structures. The coordinates for both RXRα and the accessory proteins were deleted. When a structure was found to contain more than one PPAR LBD in its asymmetric unit, only one of them was considered for the alignment. Next, the SYBYL [29] implementation of Composer [30] was run to align the PPAR LBD sequences to 1FM9 (PPARγ). The identified seed residues served as the basis for a 3D structural alignment, leading to the identification of SCRs among the entire set of proteins. In two iterations, residues with a deviation greater than 2 standard deviations (SD) were rejected.

Molecular interaction fields

After removing the ligands and the structural water molecules, the aligned PPAR LBD structures were



Scheme 1. Chemical structures of selected PPAR agonists.

checked and corrected for missing parts, without being neutralized. Although the crystal waters could play an important role in selective interactions, we have not found any water molecule which is conserved in all the superimposed binding sites. Detailed output from this analysis is supplied as supporting information. It should be mentioned that water molecules were also removed from the binding sites of other proteins previously analyzed by GRID/CPCA [24, 25]. Hydrogens were added to the structures using the program GRIN. All GRID [31, 32] calculations were carried out with a 1 Å grid size in a box enclosing all relevant parts of the PPAR LBD. During the GRID calculations, all the amino acids were kept rigid (GRID directive Move = 0). Protein flexibility was taken into account by considering several distinct structures for a given PPAR subtype. The main advantage of this approach is that conformational changes induced by a bound ligand are accounted for. The following probes were found to be the most informative [25] in identifying binding site regions which confer selectivity: the hydrophobic (DRY), the methyl (C3, steric interactions), the amide nitrogen (N1, H bond donor), the carbonyl oxygen (O, H bond acceptor), and the phenolic oxygen (OH, H bond acceptor and donor) probes. The data from GRID molecular interaction fields were organized as described previously [24, 25] in a vector containing x GRID probes $\times$ n GRID interaction points for each PPAR LBD three-dimensional structure as a row.

Consensus Principal Component Analysis (CPCA)

GRID produces too many data to enable visual inspection of common interaction patterns. Therefore, chemometrical tools such as principal component analysis (PCA) were found to be useful for extracting relevant information from molecular interaction field (MIF) maps. Principal component analysis on multivariate GRID descriptors was previously applied to uncover differences between two binding sites with respect to their GRID probe interaction patterns for targets like DNA [33], dihydrofolate reductase [34], matrix metalloproteinases [35, 36], and cyclooxygenase [37]. However, the original formulation of the GRID/PCA approach suffers from several drawbacks, namely the difficulty of determining the relative importance of the probes for selectivity as well as its limitation to pairs of targets. To overcome these limitations, an improved formulation of GRID/PCA, called GRID/Consensus PCA (CPCA) [24, 38], has been developed and applied to compare the binding sites of members of several target families including the serine proteases [39], the cytochrome P450 2C [40], the kinases [25], and the matrix metalloproteinases [41]. Briefly, CPCA is a PCA variation to analyze a data set that is intrinsically organized in blocks. In this case, each block contains the MIF values between a GRID probe and each of the protein binding sites. The aim of CPCA is to capture both the global structure of the MIF matrix and the structure of each individual block. In other words, CPCA provides global scores and loadings representing the 'consensus' of all the probes, as well as block scores and loadings express-

Table 1. Overview of PPAR LBD crystal structures retrieved from Relibase.

Subtype	PDB code	Ligand	Heterodimer	Accessory protein
PPARα	1I7G	AZ242 (agonist)	no	no
	1K7L	GW409544 (agonist)	no	SRC-1 (coactivator)
	1KKQ	GW6471 (antagonist)	no	SMRT (corepressor)
PPARγ	1PRG	*apo* form	no	no
	2PRG	rosiglitazone (agonist)	no	SRC-1 (coactivator)
	3PRG	*apo* form	no	no
	4PRG	GW0072 (partial agonist)	no	no
	1FM6	rosiglitazone (agonist)	RXRα	SRC-1 (coactivator)
	1FM9	farglitazar (agonist)	RXRα	SRC-1 (coactivator)
	1I7I	AZ242 (agonist)	no	no
	1K74	GW409544 (agonist)	RXRα	SRC-1 (coactivator)
PPARδ	1GWX	GW2433 (agonist)	no	no
	2GWX	*apo* form	no	no
	3GWX	EPA (fatty acid)	no	no

ing the 'point of view' of each probe. A weight matrix, which expresses the contribution of each probe to the global model, is also produced. Prior to the CPCA run, the matrix containing the MIF was submitted to a pretreatment step implemented in GOLPE [42, 43]. As a result, only favorable interaction energies from GRID (negative values) were kept in order to focus the analysis on attractive protein ligand interactions. Furthermore, 2- and 3-level variables were discarded, and columns with SD < 0.01 kcal mol^{-1} were rejected. Block unscaled weights (BUW) were also applied to normalize the importance of all the interactions in the final CPCA model. In addition, the cut-out tool available in GOLPE was used to focus the CPCA on the region defined by the amino acids found within 4.0 Å of farglitazar or GW2433 (Table 1). The MIF passing this pretreatment step served as input for the CPCA module of GOLPE. The resulting CPCA models were analyzed using score plots, which unveiled clustering of the similar PPAR LBD which favorably interacted with particular GRID probes. It should be noted that any grouping of targets is based on a similar protein-ligand 3D interaction pattern and not on 1D sequence similarity. Furthermore, the GOLPE active plots enable one to focus on relative differences between different target score plots, and hence on interactions separating different targets.

Docking

Compound A (Scheme 1), a subnanomolar PPARγ/δ agonist, for which no crystal structure is available, was docked in the PPARδ crystal structure 1GWX (Table 1), using the 1.2 release of GOLD [44]. We chose 1GWX as target for docking of compound A because this protein structure is derived from a complex with GW2433, a ligand similar to compound A (Scheme 1). This choice is likely to increase the chances of success of docking in a rigid protein [45].

Briefly, GOLD uses a genetic algorithm to search the configuration space and allows for full ligand and for limited protein side chain flexibility [46]. Because of the random nature of the genetic algorithm, 25 docking solutions were generated, considering the algorithm default parameters. After visual inspection of the docking solutions, we selected the solution with the same hydrogen bond pattern around the ligand carboxylic acid as was observed in the crystal structures of other PPAR agonists and which showed the minimum number of steric clashes with the LBD. The selected docking solution was subsequently submitted to 100 steps of steepest descent energy minimization followed by 2000 steps of Powell minimization, to relieve steric clashes with the protein. At each stage, the energy was evaluated using the MMF94S force field [47]. The energy minimizations were carried out using SYBYL 6.8 [29].

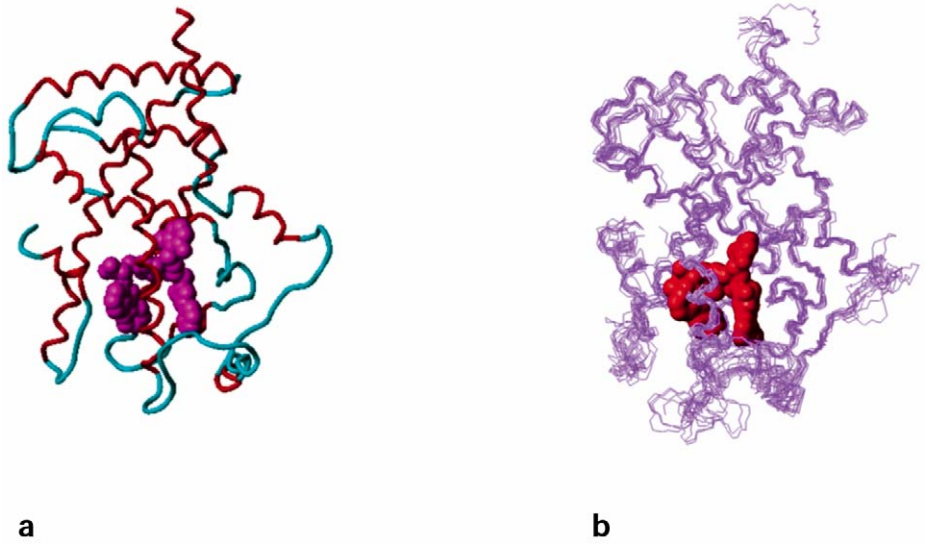


Figure 1. (a) Typical PPAR LBD fold illustrated using the PPARγ crystal structure 1FM9. (b) Structure alignment of 14 PPAR LBD crystal structures, using the PPARγ structure 1FM9 as reference. Farglitazar and GW2433 as observed in the crystal structures 1FM9 and 1GWX are displayed as space filling models. Only GRID-derived probe interaction energies found within 4.0 Å of these ligands are used for the chemometrical analysis.

Results and discussion

Structural alignment and GRID/CPCA model

The alignment strategy produced a reasonable superposition of the LBD for the 14 structures. Characteristic features of the PPAR LBD fold are illustrated in Figure 1 (left). The alignment of the 14 protein structures is shown in Figure 1 (right), where farglitazar and GW2433, selected to define the consensus binding region are displayed as space filling. At this point, we should mention that the *apo*-PPARγ structure, 3PRG, where two key amino acids, His323 and His449, show poor geometries, was discarded.

Only 4062 GRID points found within 4.0 Å of farglitazar or GW2433 were kept for analysis using PCA and CPCA. Using five GRID probes (DRY, C3, N1, O, OH), a significant CPCA model was obtained. Globally, PC1 and PC2 explain 25.6% of the variance. Within each block, this percentage varies between 21 and 29%. Detailed output from GOLPE is supplied as supporting information. Applications of the GRID/CPCA methodology to other proteins within our group or within other laboratories have produced similar values for the percentage of the variance explained by PC1 and PC2. This results from the noise produced by points located outside the ligand's binding cavity. Using the cut-out tool implemented in GOLPE leads to the elimination of some (but not all) of the 'noisy' points. It should also be mentioned that the percentage of the variance explained by PC1 and PC2 is not an indicator of the quality of a model. This percentage helps to decide how many PCs have to be selected for the interpretation step. In our opinion, the separation of the objects and the interpretability of a model are more important criteria to assess the quality of a model.

The combined super-scores from CPCA (related to a regular PCA score plot) are plotted in Figure 2 with the first principal component (PC1) on the *x*-axis and the second one (PC2) on the *y*-axis. The first PC separates all the PPARγ but 1PRG and 4PRG from the PPARδ LBDs, while the second discriminates between PPARα and PPARγ (except 1PRG and 4PRG), δ LBD. Interestingly, the PPARγ LBDs crystallized with an agonist occupy the same region of the PC1-PC2 super-scores plot, while the *apo* form (1PRG) and the one with the partial agonist GW0072 (4PRG) map very close to the PPARα LBDs. The LBD structures of a given PPAR subtype are spread, which reflects conformation or orientation variations of protein side chains and backbone regions. The backbone of the loop between helices 3 and 4 as well as the backbone of helix 12 exhibit the largest conformational variations.

Individual score plots per GRID probe from CPCA only show minor differences in the orientation of

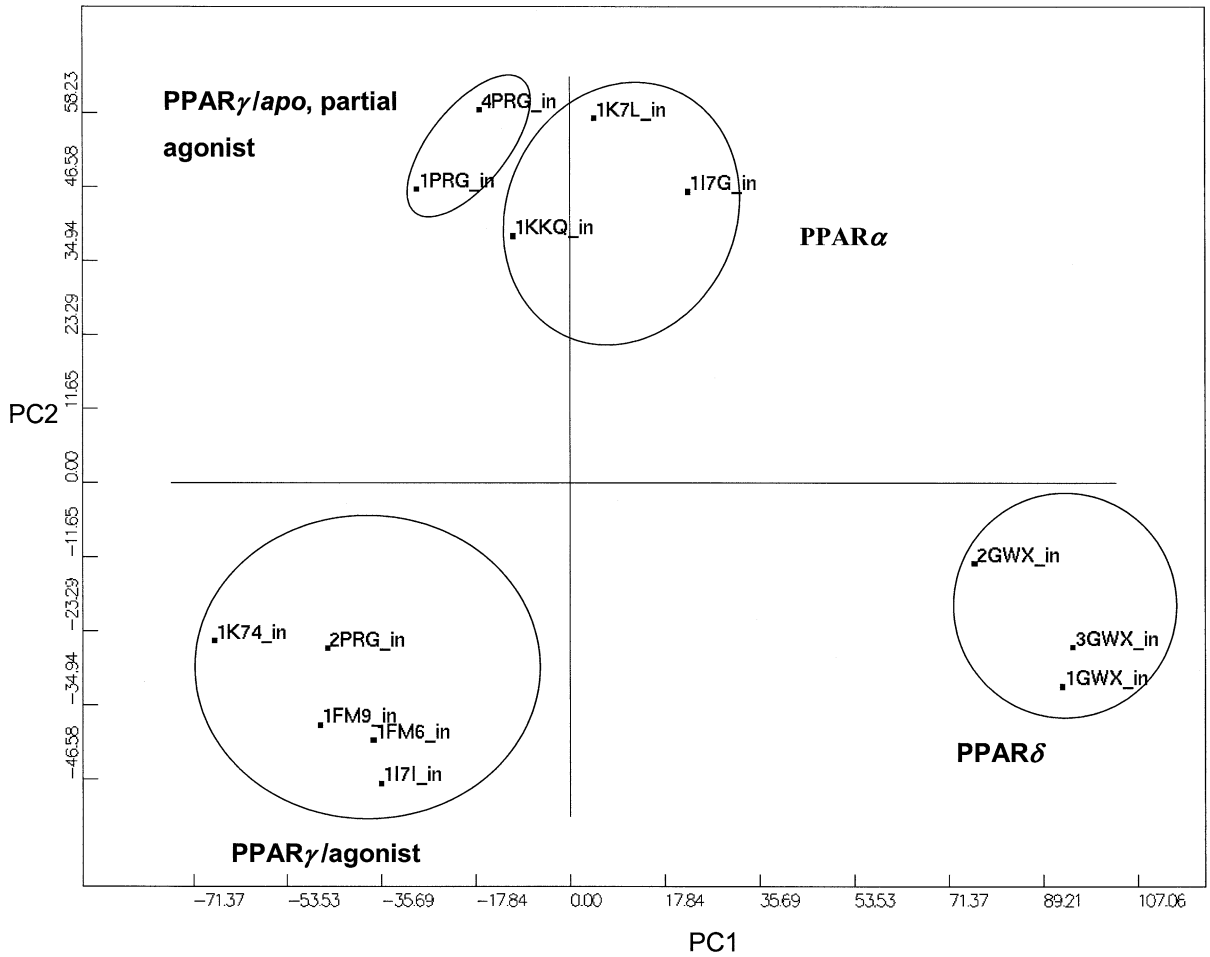


Figure 2. PCA score plot for analysis of all five different GRID probes (DRY, C3, N1, O, OH). The plot illustrates the difference between the different PPAR LBD subtypes in the chemometrical space.

the PPAR subtypes, and hence do not fundamentally change the interpretation. The CPCA superweight plot highlights the relative importance of each block (each individual GRID probe, in this case) in the overall model, for each model dimension. As illustrated in Figure 3, the probes are grouped in three clusters, including the DRY probe, the steric C3 and the three polar probes (N1, O and OH), which show weaker contributions. Thus, hydrophobic and steric interactions contribute mostly to the discrimination between the three PPAR subtypes. The DRY, C3 and OH probes were subsequently selected for further analysis because they represent each of the three clusters found in the CPCA superweight plot.

Interpretation of the GRID/CPCA model in structural terms

The PPAR subtype selectivity differences are plotted in Figure 4 for PC1 and PC2 for the DRY, C3 and OH probes derived from the PCA model. The cyan contours correspond to negative regions, and the yellow contours indicate positive regions. The crystallographically determined binding orientation of the PPARγ and δ agonists farglitazar and GW2433 from the Protein Data Bank files 1FM9 [46] and 1GWX [20] are shown for comparison. In PC1, negative contours (cyan) characterize regions with preferences for PPARγ for a given GRID probe, while positive contours (yellow) highlight favorable and selective interactions with PPARδ. In PC2, negative contours (cyan) indicate parts of the LBD with preferences for PPARγ and δ, whereas positive contours (yellow) correspond

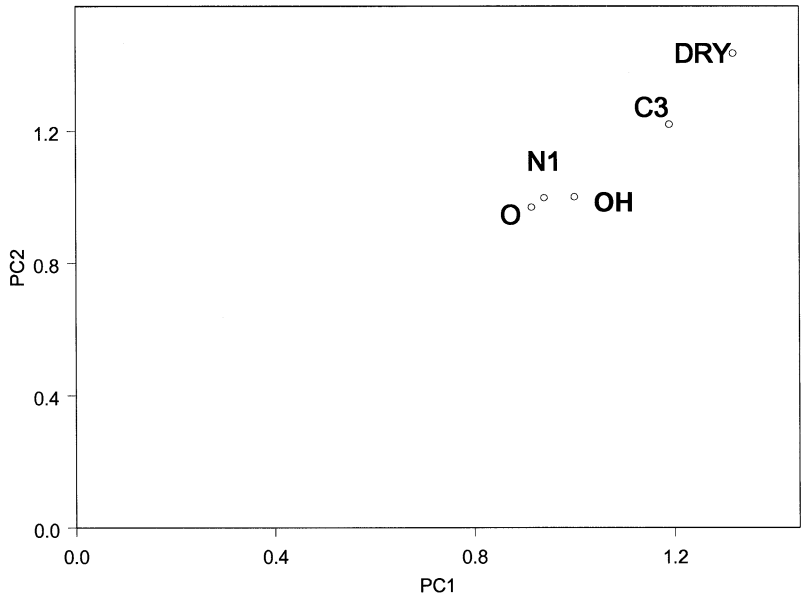


Figure 3. CPCA superweights obtained for the analysis with the five GRID probes. This plot highlights the influence of the different blocks (GRID probes) into the principal components.

to subsites where selectivity for PPARα can be gained. In the following discussion, we define four main regions: the head region binding the carboxylic acid, the left and right distal cavities interacting with both of the GW2433 urea substituents, and the linker between the head and the distal cavities. Several PPARα and γ ligands shown in Scheme 1 exhibit in α of their acidic function a substituent that binds to a side pocket which is referred to as the left proximal cavity. We should also mention that the amino acid positions correspond to those in the Protein Data Bank file 1FM9. A list of the amino acids contributing to the separation between the three PPAR subtypes is also supplied as supplementary information.

As can be seen from the CPCA differential contour plots (Figure 4), PC1 and PC2 are dominated by selectivity subsites mainly located in the distal cavities. Within the left distal cavity, the different distribution of hydrophobic and bulky side chains can be exploited to modulate the selectivity for any of the three PPAR subtypes. In particular, two bulky side chains at positions 262 and 284, which are located at the left end of the left distal pocket, limit its size in PPARδ, while in PPARα and γ, the size of this pocket is restricted by mutations at positions 281 and 339. Mutations at positions 348 and 364, as well as variations in side chain conformations, also affect the shape of the left distal pocket, near its junction with the linker region. In addition, hydrogen bonding within the left distal cavity can contribute to increase the selectivity for PPARδ. The size of the PPARγ right distal cavity is reduced by the mutation at position 288, which also affects its hydrogen bonding properties. Hydrogen bonding with Thr292 can also help to increase the selectivity for PPARα and δ. This latter amino acid as well as those in positions 228 and 333 can also be involved in selective interactions with the DRY and C3 probes. Most of the differences observed in the CPCA differential contour plots around the linker region arise from changes in protein side chain conformations. However, hydrogen bonding with Tyr327 is favorable for the selectivity for PPARγ. Within the head region, which binds the acidic function of most PPAR ligands, only mutation of residue 323 in PPARα results in an increased steric bulk as well as in a change in hydrogen bonding properties. The neighboring left proximal pocket is accessible for ligand binding in both PPARα and γ, whereas in PPARδ, the side chain of residue 453 causes steric hindrance at the entrance of this pocket. Deeper in the left proximal pocket, hydrophobic interactions with amino acid 363 are favorable for increasing the selectivity for PPARγ.

Analysis of the CPCA differential plots is also useful to localize the structural differences that are responsible for the separation between the PPARγ LBDs with a bound agonist and the outliers, 1PRG and 4PRG (Figure 2, Table 1). This separation results from conformational changes in the five regions of the LBD

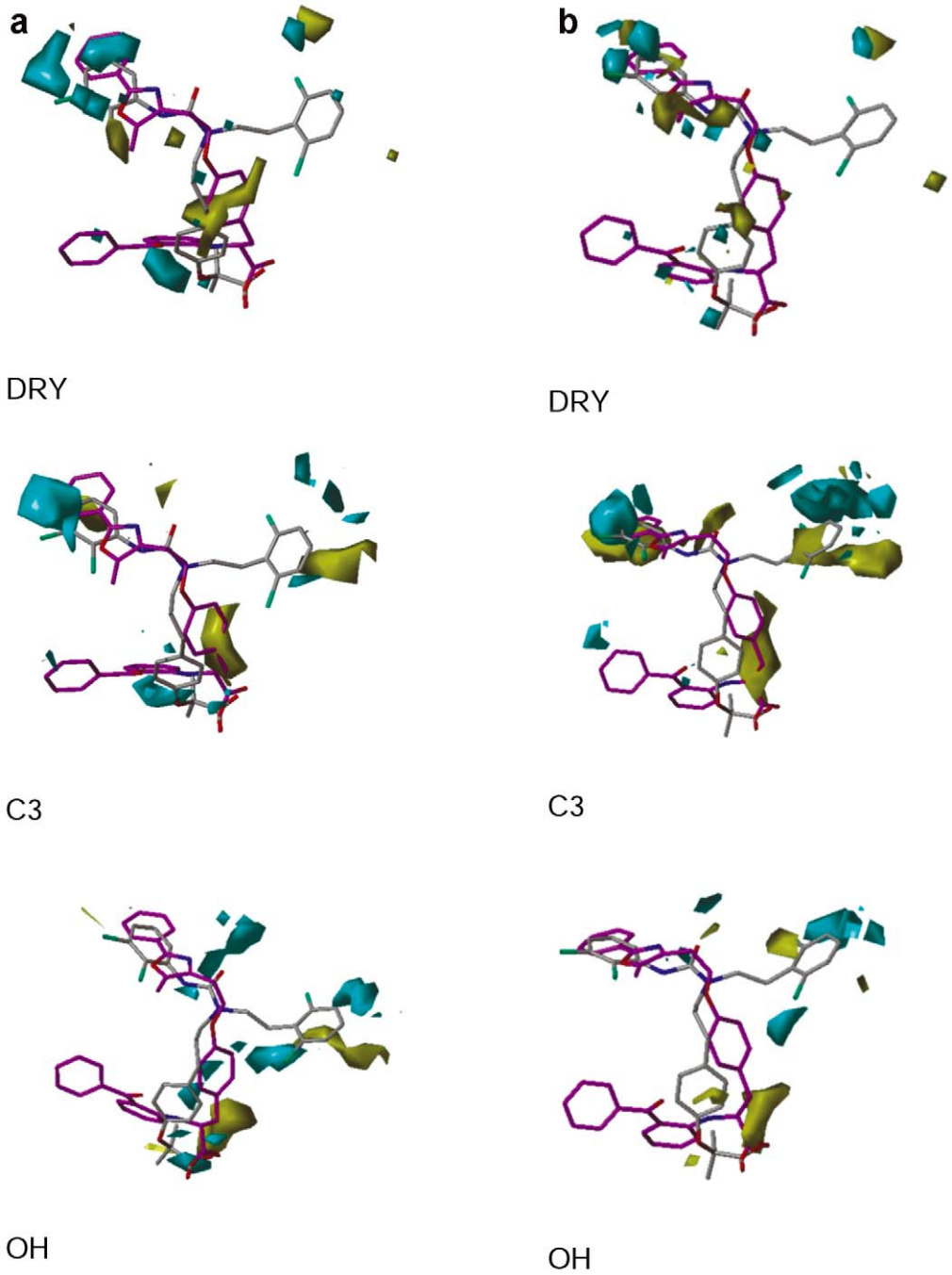


Figure 4. CPCA differential plots for the first (a) and second component (b) obtained for three representative probes (DRY, C3 and OH). The plots obtained for PC1 highlight differences between PPARγ and PPARδ, while the PC2 plots indicate differences between PPARγ/δ and PPARα. Cyan and yellow contours correspond to negative and positive regions, respectively. The X-ray structures of farglitazar and GW2433 taken from their complex with PPARγ and PPARδ are shown for comparison. Hydrogens and protein atoms are omitted for clarity.

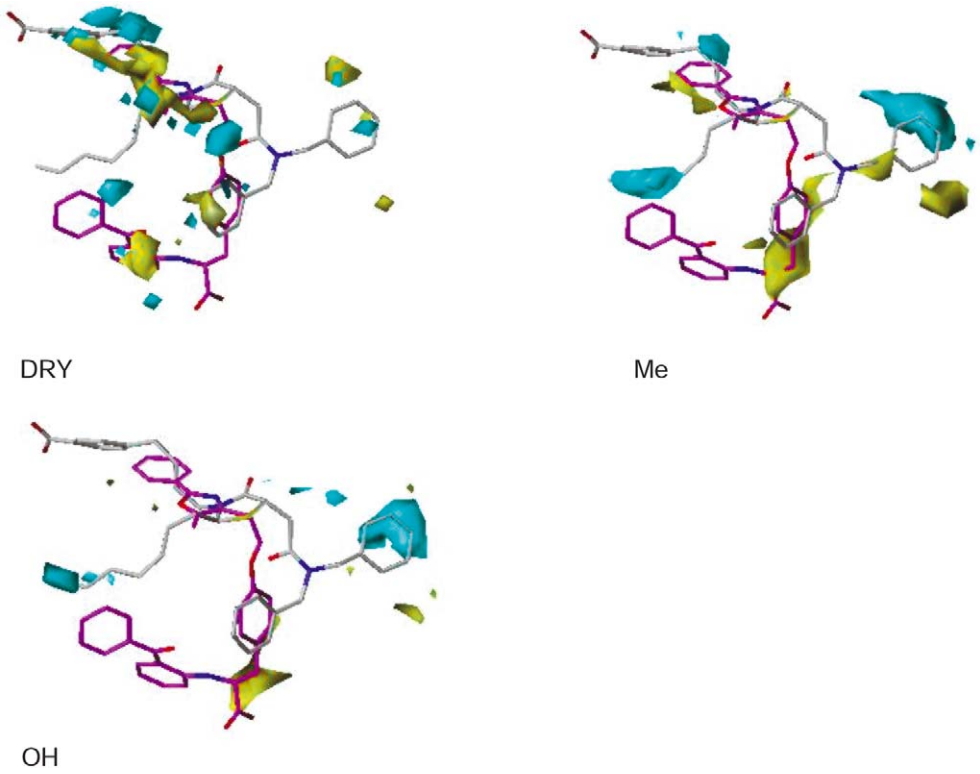


Figure 5. CPCA differential plots illustrating the differences between 1FM9 and 4PRG, selected as representatives of the two clusters of PPARγ structures. Cyan and yellow contours correspond to negative and positive regions, respectively. The X-ray structures of farglitazar and GW0072 taken from their complex with PPARγ are shown for comparison. Hydrogens and protein atoms are omitted for clarity.

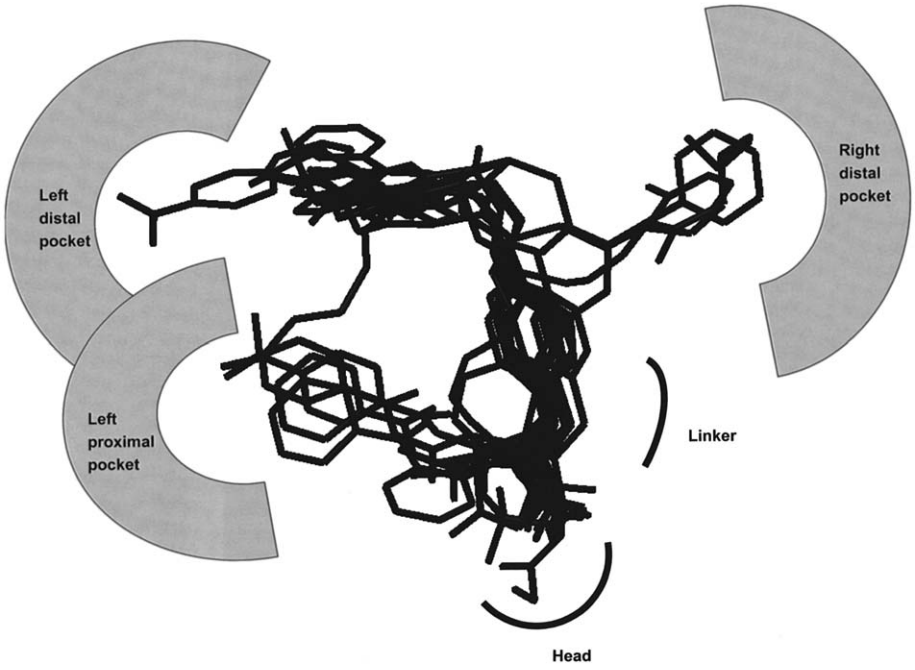


Figure 6. Superposition of 11 PPAR ligands obtained after the structure alignment of the PPAR LBD crystal structures.

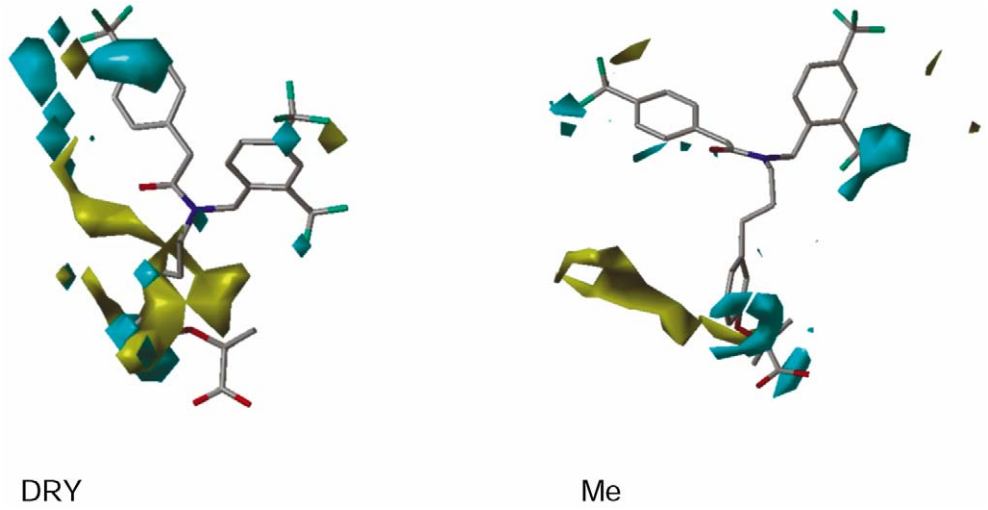


Figure 7. Docking of compound A into PPARδ LBD. The PC2 differential contours obtained for the DRY and C3 probes are displayed to highlight regions of the binding site where selectivity for PPARγ/δ can be gained (cyan). Yellow contours correspond to regions where selectivity for PPARα is increased.

(Figure 5). In the head region, the backbone and side chains of helix 12 exhibit most of the conformational variability. Variations in side chain conformations account for differences observed in the linker region, the left proximal and right distal cavities. The differences observed in the left distal cavity result from conformational changes of the C-terminal portion of helix 3 and of the loop between helix 3 and helix 4. Interestingly, this loop adopts similar conformations in 1PRG, 4PRG and in two PPARα structures (1K7L and 1KKQ), while it is missing in the third PPARα structure (1I7G). It is also noteworthy that conformational changes of helix 12, which represent the main difference between the agonist-bound and antagonist-bound forms of PPARα (Table 1), do not cause such a large separation between PPARα structures in the score plot (Figure 2).

PPAR ligands

Because of their therapeutic potential, many pharmaceutical companies are working to discover either selective or mixed PPAR ligands. As shown in Figure 6, all the ligands that have been so far crystallized with the different PPAR LBD, occupy at least three of the five LBD subsites, namely the head, the linker and the left distal cavity. Only the left proximal pockets of PPARα and γ are filled with ligands, while so far PPARδ (1GWX) and PPARγ (4PRG) are the only family members with ligands bound to the right distal pocket. Our GRID/CPCA computations reveal that most of the selective interactions are formed with both distal pockets (Figure 4). This observation agrees with the SAR published by Liu and colleagues, who showed that the selectivity between PPARα and PPARγ could be fine-tuned by changing the flexibility or the substitution pattern of the ligands tail [11]. Another example is compound A (Scheme 1), a dual subnanomolar PPARγ/δ agonist, which exhibits more than 30-fold selectivity *versus* PPARα in both binding and functional assays [12]. When docked in PPARδ, the amide substituents of compound A that bind to the left and right distal pockets overlap negative contours of the PC2 differential CPCA plots obtained for the DRY and C3 probes (Figure 7). In PC2, these negative contours highlight regions of the LBD where selectivity for PPARγ and PPARδ can be increased. Hydrophobic and steric interactions turn out to be the major determinants of PPAR subtype selectivity. However, the GRID/CPCA results suggest that differences in hydrogen bonding properties of the distal pockets could also be exploited to alter PPAR subtype selectivity (Figure 4). Although the three remaining subsites of the PPAR LBD exhibit less variation than the distal pockets, the GRID/CPCA calculations reveal some scope for selective interactions, in particular with the OH and C3 probes within the linker and head regions. Again, our analysis agrees well with site-directed mutagenesis experiments, resulting in the identification of residue 323 (Tyr in PPARα, His in PPARγ and δ) as a key determinant of PPAR subtype selectivity for

thiazolidinedione and nonthiazolidinone ligands [18]. Furthermore, the steric hindrance generated in PPARδ by Met453 at the entrance of the left proximal pocket (Figure 4) accounts for the lack of PPARδ activity of tyrosine-based compounds like farglitazar. In addition, a Met to Val change at position 453 generates a PPARδ mutant, responsive to fibrates. Reciprocally, a Val to Met mutation at position 453 of PPARα results in a loss of response to fibrates.

Conclusions

The GRID/CPCA approach was used successfully to examine the LBD of the three PPAR subtypes which share between 60 and 70% sequence identity. As a result, regions of the LBD amenable to selective interactions with given functional groups were identified. In particular, both distal pockets exhibit considerable differences among the three PPAR subtypes and are expected to be important in the design of ligands with well-defined selectivity profile. On the other hand, differences between only two PPAR subtypes are observed within the other regions of the binding site. This target family landscape-derived classification is also in strong agreement with published structure-activity relationships for different PPAR ligands, as well as with site-directed mutagenesis results. We anticipate to provide valuable guidelines on both the subsites and the type of interactions which could be used to design PPAR ligands with defined selectivity profiles.

Acknowledgements

Thanks to Hans Matter (Aventis) and Gabriele Cruciani (University of Perugia, Italy) for many interesting discussions, and to the two anonymous reviewers for constructive comments.

Supporting Information Available

Structural analysis of the PPAR crystal data, variable explaining each component in the GRID/CPCA model and list of amino acids contributing to the separation between the three PPAR subtypes.

References

1. Chawla, A., Repa, J.J., Evans, R.M. and Mangelsdorf, D.J., Science, 294 (2001) 1866.
2. Horn, F., Vriend, G. and Cohen, F.E., Nucleic Acids Res., 29 (2001) 346.
3. Willson, T.M., Brown, P.J., Sternbach, D.D. and Henke, B.R., J. Med. Chem., 43 (2000) 527.
4. Available on the web: http://receptors.ucsf.edu/NR/.
5. Corton, C.J., Anderson, S.P. and Stauber, A., Annu. Rev. Pharmacol. Toxicol., 40 (2000) 491.
6. Otto, C., Lehrke, M. and Göke, B., Pharmacogenomics, 3 (2002) 99.
7. Perry, C.G. and Petrie, J.R., Exp. Opin. Emerging Drugs, 7 (2002) 165.
8. Oliver, W.J. Jr., Shenk, J.L., Snaith, M.R., Russell, C.S., Plunket, K.D., Bodkin, N.L., Lewis, M.C., Winegar, D.A., Sznaidman, M.L., Lambert, M.H., Xu, H.E., Sternbach, D.D., Kliewer, S.A., Hansen, B.C., and Willson, T.M. A., Proc. Natl. Acad. Sci. USA, 98 (2001) 5306.
9. Rami, H.K. and Smith, S.A., Exp. Opin. Ther. Patents, 10 (2000) 623.
10. Brown, P.J., Stuart, L.W., Hurley, K.P., Lewis, M.C., Winegar, D.A., Wilson, J.G., Wilkison, W.O., Ittoop, O.R. and Willson, T.M., Bioorg. Med. Chem. Lett., 11 (2001) 1225.
11. Liu, K.G., Smith, J.S., Ayscue, A.H., Henke, B.R., Lambert, M.H., Leesnitzer, L.M., Plunket, K.D., Willson, T.M. and Sternbach, D.D., Bioorg. Med. Chem. Lett., 11 (2001) 2385.
12. Liu, K.G., Lambert, M.H., Leesnitzer, L.M., Oliver, W. Jr., Ott, R.J., Plunket, K.D., Stuart, L.W., Brown, P.J., Willson, T.M. and Sternbach, D.D., Bioorg. Med. Chem. Lett., 11 (2001) 2959.
13. Brooks, D.A., Etgen, G.J., Rito, C.J., Shuker, A.J., Dominianni, S.J., Washawsky, A.M., Ardecky, R., Paterniti, J.R., Tyhonas, J., Karanewsky, D.S., Kauffman, R.F., Broderick, C.L, Oldham, B.A., Montrose-Rafizadeh, C., Winneroski, L.L., Faul, M.M. and McCarthy, J.R., J. Med. Chem., 44 (2001) 2061.
14. Lohray, B.B., Lohray, V.B., Bajji, A.C., Kalchar, S., Poondra, R.R., Padakanti, S., Chakrabarti, R., Vikramadithyan, R.K., Misra, P., Juluri, S., Mamidi, N. V.S.R. and Rajagopalan, R., J. Med. Chem., 44 (2001) 2675.
15. Steinmetz, A.C.U., Renaud, J.-P. and Moras, D., Annu. Rev. Biophys. Biomol. Struct., 30 (2001) 329.
16. Bourguet, W., Germain, P. and Gronemeyer, H., Trends Pharma Sci., 21 (2000) 381.
17. Weatherman, R.V., Fletterick, R.J. and Scanlan, T.S., Annu. Rev. Biochem., 68 (1999) 559.
18. Xu, H.E., Lambert, M.H., Montana, V.G., Plunket, K.D., Moore, L.B., Collins, J.L., Oplinger, J.A., Kliewer, S.A., Gampe, R.T. Jr., McKee, D.D., Moore, J.T. and Willson, T.M., Proc. Natl. Acad. Sci. USA, 98 (2001) 13919.
19. Cronet, P., Petersen, J.F.W., Folmer, R., Blomberg, N., Sjöblom, K., Karlson, U., Lindstedt, E.-L. and Bamberg, K., Structure, 9 (2001) 699.
20. Xu, H.E., Lambert, M.H., Montana, V.G., Parks, D.J., Blanchard, S.G., Brown, P.J., Sternbach, D.D., Lehmann, J.M., Wisely, G.B., Willson, T.M., Kliewer, S.A. and Milburn, M.V., Mol. Cell., 3 (1999) 397.
21. Nolte, R.T., Wisely, G.B., Westin, S., Cobb, J.E., Lambert, M.H., Kurokawa, R., Rosenfeld, M.G., Willson, T.M., Glass, C.K. and Milburn, M.V., Nature, 395 (1998) 137.
22. Xu, H.E., Stanley, T.B., Montana, V.G., Lambert, M.H., Shearer, B.G., Cobb, J.E., McKee, D.D., Galardi, C.M., Plunket, K.D., Nolte, R.D., Parks, D.J., Moore, J.T., Kliewer, S.A., Wilson, T.M. and Stimmel, J.B., Nature, 415 (2002) 813.

23. Takada, I., Yu, R.T., Xu, H.E., Lamber, M.L., Montana, V.G., Kliewer, S.A., Evans, R.M. and Umesono, K., Mol. Endocrinol., 14 (200) 733.
24. Kastenholz, M.A., Pastor, M., Cruciani, G. Haaksma, E.E.J. and Fox, T., J. Med. Chem., 43 (2000) 3033.
25. Naumann, T. and Matter, H., J. Med. Chem., 45 (2002) 2366.
26. Uppenberg, J., Svensson, C., Jaki, M., Bertlisson, G., Jendeberg, L. and Berkenstam, A., J. Biol. Chem., 273 (1998) 31108.
27. Bergner, A., Gunther, J., Hendlich, M., Klebe, G. and Verdonk, M., Biopolymers, 61 (2002) 99.
28. Relibase+ Version 1.0 available from the Cambridge Crystallographic Data Center (http://www.ccdc.cam.ac.uk/)
29. SYBYL Molecular Modeling Package, Version 6.8, Tripos, St. Louis, MO, 2001.
30. a) Blundell, T.L., Carney, D.P., Gardner, S., Hayes, F.R.F., Howlin, B., Hubbard, T.J.P., Overington, J.P., Singh, D.A., Sibanda, B.L. and Sutcliffe, M.J., Eur. J. Biochem., 172 (1988) 513. b) Sutcliffe, M.J., Haneef, I., Carney, D.P. and Blundell, T.L., Protein Eng., 1 (1987) 377. c) Sutcliffe, M.J., Hayes, F.R.F. and Blundell, T.L., Protein Eng., 1 (1987) 385.
31. a) Goodford, P.J., J. Med. Chem., 28 (1985) 849. b) Boobbyer, D.N.A., Goodford, P.J., McWhinnie, P.M. and Wade, R.C., J. Med. Chem., 32 (1989) 1083. c) Wade, R.C., Clerk, K.J. and Goodford, P.J., J. Med. Chem., 36 (1993) 140. d) Wade, R.C. and Goodford, P.J., J. Med. Chem., 36 (1993) 148.
32. GRID version 20, Molecular Discovery Ltd., Oxford, UK, 2002.
33. Cruciani, G. and Goodford, P.J., J. Mol. Graphics, 12 (1994) 116.
34. Pastor, M. and Cruciani, G., J. Med. Chem., 38 (1995) 4637.
35. Matter, H. and Schwab, W., J. Med. Chem., 42 (1999) 4506.
36. Matter, H. and Schwab, W., In Gundertofte, K. and Jørgensen, F.S. (Eds.), Molecular Modeling and Prediction of Bioactivity, Kluwer, New York, NY, 2000, pp. 123-128.
37. Filipponi, E., Ceccheti, V., Tabarrini, O. and Bonelli, D., Fravolini, J. Comput.-Aided Mol. Des., 14 (2000) 277.
38. Westerhuis, J.A., Kourti, T. and MacGregor, J.F., J. Chemom., 12 (1998) 301.
39. Matter, H., Defossa, E., Heinelt, U., Naumann, T., Schreuder, H. and Wildgoose, P., In. Höltje, H.D. and Sippl (Eds.), Rational Approaches to Drug Design, Proceedings of the 13th European Symposium on Quantitative Structure-Activity Relationships, Prous, Barcelona, 2001, pp. 177–185.
40. Ridderström, M., Zamora, I., Fjellström, O. and Anderson, T.B., J. Med. Chem., 44 (2001) 4072.
41. Terp, G.E., Cruciani, G., Christensen, I.T. and Jørgensen, F.S., J. Med. Chem., 45 (2002) 2675.
42. GOLPE 4.5., Multivariate Infometric Analysis, Srl, Perugia, Italy, 1999.
43. Baroni, M., Costantino, G., Cruciani, G., Riganelli, D., Valigi, R. and Clementi, S., Quant. Struct.-Act. Relat., 12 (1993) 9.
44. GOLD Version 1.2, available from the Cambridge Crystallographic Data Center (http://www.ccdc.cam.ac.uk/)
45. Murray, C.W., Baxter, C.A. and Frenkel, D.A., J. Comput.-Aided Mol. Des., 13 (1999) 547.
46. a) Jones, G., Willett, P. and Glen, R.C., J. Mol. Biol., 245 (1995) 43. b) Jones, G., Willett, P., Glen, R.C., Leach, A.R. and Taylor, R., J. Mol. Biol., 267 (1997) 727.
47. Halgren, T.A., J. Comput. Chem., 20 (1999) 720.