Exploring the Binding Site Structure of the PPARγ Ligand-Binding Domain by Computational Solvent Mapping[†]

Shu-Hsien Sheu,[‡] Taner Kaya,[§] David J. Waxman,[||] and Sandor Vajda*,[‡]

Departments of Biomedical Engineering, Chemistry, and Biology, Boston University, 44 Cummington Street, Boston, Massachusetts 02215

Received September 12, 2004; Revised Manuscript Received October 30, 2004

ABSTRACT: Solvent mapping moves molecular probes, small organic molecules containing various functional groups, around the protein surface, finds favorable positions, clusters the conformations, and ranks the clusters based on the average free energy. Using at least six different solvents as probes, the probes cluster in major pockets of the functional site, providing detailed and reliable information on the amino acid residues that are important for ligand binding. Solvent mapping was applied to 12 structures of the peroxisome proliferator activated receptor γ (PPARγ) ligand-binding domain (LBD), including 2 structures without a ligand, 2 structures with a partial agonist, and 8 structures with a PPAR agonist bound. The analysis revealed 10 binding "hot spots", 4 in the ligand-binding pocket, 2 in the coactivator-binding region, 1 in the dimerization domain, 2 around the ligand entrance site, and 1 minor site without a known function. Mapping is a major source of information on the role and cooperativity of these sites. It shows that large portions of the ligand-binding site are already formed in the PPARγ apostructure, but an important pocket near the AF-2 transactivation domain becomes accessible only in structures that are cocrystallized with strong agonists. Conformational changes were seen in several other sites, including one involved in the stabilization of the LBD and two others at the region of the coactivator binding. The number of probe clusters retained by these sites depends on the properties of the bound agonist, providing information on the origin of correlations between ligand and coactivator binding.

The peroxisome proliferator activated receptor γ (PPARγ)[1] is a ligand-activated transcription factor and member of the nuclear receptor superfamily that plays an important role in adipogenesis and glucose homeostasis. PPARγ and the closely related receptors PPARα and PPARδ bind a variety of fatty acids and their metabolites (*1*−*3*). Synthetic PPARγ agonists, including thiazolidinediones (TZDs), have been shown to be effective as insulin-sensitizing agents, reducing insulin resistance and lowering plasma glucose levels in patients with type-2 diabetes (*4*). The effects of ligands on PPARγ are mediated through the ligand-binding domain (LBD), a region of 270 amino acid residues in the C-terminal half of the receptor (*2*). In addition to its role in ligand binding, the LBD also contains dimerization and transactivation regions, including the transcriptional activation function 2 (AF-2) associated with helix 12 (H12). Structural (*5*−*12*) and biochemical (*13*, *14*) studies have helped to elucidate the mechanism of ligand-induced transcriptional activation by PPARγ. Upon binding of an agonist, the PPARγ LBD undergoes conformational changes, most notably in the AF-2 region. These changes result in the displacement of corepressor proteins that inhibit transcription and the recruitment of coactivator proteins that are required for transcriptional activation.

Although the basic principles that govern the activation of PPARγ are well-established (*2*, *3*), several important questions remain. First, the AF-2 region of PPARγ can be in an active-like state without a bound ligand, and because the structural differences between such an "active" apo AF-2 domain and an AF-2 domain with a bound ligand and coactivator are quite small, the origin of the activation "switch" is not obvious. Biochemical studies suggest that ligand binding stabilizes H12 in the active conformation, both by direct contacts between the ligand and residues in helix H12 and by globally stabilizing the lower half of the LBD, which itself will favor H12 stably adopting an active conformation (*2*, *13*, *14*). However, it is not clear how ligand binding stabilizes distal parts of the LBD, and vice versa, how distal secondary structural interactions contribute to the positioning of the ligand and transcriptional output as reported in the literature (*15*). Second, it is unclear why structurally similar PPARγ ligands are sometimes characterized by significantly different pharmacological profiles (*16*). For example, the TZDs rosiglitazone and pioglitazone are PPARγ agonists that are currently prescribed in clinical

[†] This investigation was supported by Grant P42 ES07381 from the National Institute of Environmental Health Sciences, Grant GM61867 from the National Institute of Health, and Grants DBI-9904834 and MRI DBI-0116574 from the National Science Foundation.

* To whom correspondence should be addressed. Telephone: 617-353-4757. Fax: 617-353-6766. E-mail: vajda@bu.edu.

[‡] Department of Biomedical Engineering.

[§] Department of Chemistry.

[||] Department of Biology.

[1] Abbreviations: PPAR, peroxisome proliferator activated receptor; LBD, ligand-binding domain; AF-2, transcriptional activation function 2; SRC-1, steroid receptor coactivator 1; TZD, thiazolidinedione; H12, helix 12; PDB, Protein Data Bank; RXR, retinoid X receptor; ACE, analytical continuum electrostatics; FDPB, finite difference Poisson−Boltzmann; ACP, atomic contact potential.

10.1021/bi048032c CCC: $30.25
Published on Web 12/30/2004

practice as insulin-sensitizing agents. Both TZDs have similar effects on normalization of glycemic levels but exert disparate actions on high-density lipoprotein (HDL) metabolism. Whereas rosiglitazone decreases HDL and increases the level of low-density lipoprotein (LDL), pioglitazone significantly raises HDL and lowers LDL (*16*). These differences may result from distinct modes of interaction between PPARγ and the coactivator protein PGC-1α induced by rosiglitazone compared to pioglitazone (*16*); however, the molecular basis for any such differences are unclear. Third, PPARγ binds to DNA as a heterodimer with the retinoid X receptor (RXR), and there is some evidence that allosteric communication occurs between the receptor partners, affecting ligand binding (*15*−*18*). However, the mechanism of interaction between PPARγ's dimerization and ligand-binding site is not clear. These observations suggest that the PPARγ LBD is a complex system in which the identity of a particular ligand is transmitted to transcriptional coregulators by ligand-induced conformational changes, both directly in the co-regulator-binding region and indirectly in distal parts of the LBD.

In the present study, we use computational solvent mapping (*19*−*21*) to identify 10 important binding sites as "hot spots" on the PPARγ LBD. We then investigate the roles, relative importance, and cooperativity of these sites in ligand and coactivator binding and try to determine the specific LBD amino acid residues that are primarily responsible for the changes in these properties. Computational solvent mapping is based on the original experimental studies of Ringe and co-workers (*22*−*24*), who determined protein structures in aqueous solutions of various organic solvents, and in each case found only a limited number of organic molecules bound to the protein in the first water layer. The power of the method arises when at least six structures of a given protein solved in different solvents are superimposed, because different solvent molecules tend to cluster in the active site, forming "consensus" sites that delineate the important subsites of the binding pocket (*22*−*26*).

We have shown that solvent mapping of proteins can be carried out computationally rather than experimentally (*19*−*21*), and we employ this computational approach here. The method moves small organic molecules containing various functional groups ("molecular probes") around the protein surface, finds favorable positions using empirical free-energy functions, clusters the conformations, and ranks the clusters based on average free energy (*21*). We employ at least six different probes, retain the five lowest free-energy clusters for each probe, and define the positions at which several clusters overlap as the consensus sites. Mapping a number of enzymes has shown that the largest consensus sites (i.e., consensus sites with the highest number of overlapping clusters) are almost always localized to major subsites of the enzyme active site, and as a result, the amino acid residues that interact with the probes also bind the specific ligands of the protein (*21*). Thus, the method can provide detailed and reliable information on important amino acid residues in the binding site. It is important to note that the mapping finds the active site even when it is not the largest pocket (*21*), and in this sense, it performs better than binding site identification methods based on purely geometric criteria (*27*−*29*).

Here, we describe the application of solvent mapping to the PPARγ LBD. Our overall aims are to find the "hot spots" of the protein where most organic probe molecules cluster, to determine if these sites play any role in the binding of ligands, coactivators, or dimeric partners, and to study the effects of the conformational changes upon the binding of various ligands. These analyses are now possible because of the availability of 12 PPARγ LBD X-ray structures, including 2 without a ligand, 2 with a partial agonist, and 8 with 6 different agonists bound. In all structures, any clustering of the probes occurs at a number of distinct sites, revealing 10 binding "hot spots". A total of 4 of the 10 sites are in the ligand-binding site, 2 in the coactivator-binding region, 1 in the dimerization domain, 2 around the ligand entrance site, and 1 minor site without a known function. After the number of probes in each of the 10 sites for all of the 12 structures was determined, we were able to obtain information on ligand-induced conformational changes, the potential interactions among "hot spots", and interactions between these sites and the rest of the protein. In particular, we find that large portions of the PPARγ-binding site are already formed in the ligand-free structure, but an important pocket near the AF-2 domain is accessible only in structures that are cocrystallized with strong agonists. Mapping also shows that the binding of agonists restricts the size of a distal pocket and thus contributes to the stabilization of the LBD. Agonist binding introduces conformational changes in several other sites, including two at the site of coactivator binding. The number of probe clusters retained by these sites is shown to depend on the properties of the agonist and hence may help to understand the observed differences in coactivator binding and transactivation profiles.

MATERIALS AND METHODS

Uniform Solvent Mapping. The goal of this method is to obtain information on the regions of a protein that are most likely to bind organic solvents used as molecular probes. No a priori assumptions on the location or properties of the binding site are made. The algorithm consists of five computational steps as follows (*21*).

Step 1. Rigid-Body Search. A multistart simplex method is used to move the probes around the protein, starting from evenly distributed points over the protein surface, generating over 6000 docked conformations for each probe. The scoring function in the search is given by $\Delta G_s = \Delta E_{elec} + \Delta G_{des} + V_{exc}$, where ΔE_{elec} denotes the direct (Coulombic) part of the electrostatic energy, ΔG_{des} is the desolvation free energy, and V_{exc} is an excluded volume penalty term such that $V_{exc} = 0$ if the ligand does not overlap with the protein. The electrostatic energy is determined by the expression $\Delta E_{elec} = \sum_i \Phi_i q_i$, where q_i is the charge of the ith probe atom and Φ_i is the electrostatic field of the solvated protein at that point (*30*−*32*). The electric field Φ is calculated by a finite difference Poisson−Boltzmann (FDPB) method (*30*) using the CONGEN program (*32*). Dielectric constants $\epsilon = 4$ and 78 are used for the protein and solvent, respectively. We use the template partial charges provided by the Quanta program (*33*) for the probe molecules. The desolvation term, ΔG_{des}, is obtained by the atomic contact potential (ACP) model, (*34*) an atomic-level extension of the Miyazawa−Jernigan potential (*35*).

Table 1: PPARγ Structures in the PDB, Studied by Computational Solvent Mapping

chain[a]	X-ray structure	H12[b]	ligand	binding affinity	ligand efficacy (μM)	ref
1prg (a)	homodimer	on	none			*5, 6*
1prg (b)	homodimer	off	none			
4prg (a)	homodimer	on	GW0072 (partial agonist)	$IC_{50} = 110$ nM in 100 nM rosiglitazone, $K_i = 70$ nM	15−20% of rosiglitazone	*7*
4prg (b)	homodimer	off	GW0072			
2prg (a)	homodimer and SRC-1 peptide	on	rosiglitazone (TZD)	$IC_{50} = 0.44 \pm 0.04$ μM	$EC_{50} = 0.043$	*5*
1fm6 (d)	PPARγ/RXRα heterodimer with SRC-1 peptide	on	rosiglitazone (TZD)	$IC_{50} = 0.44 \pm 0.04$ μM	$EC_{50} = 0.043$	*8*
1fm9 (d)	PPARγ/RXRα heterodimer with SRC-1 peptide	on	farglitazar (GI262570) (L-Tyr derivative)	$K_i = 1.0$ nM	$EC_{50} = 0.000\,20$	*8*
1k74 (d)	PPARγ/RXRα heterodimer with SRC-1 peptide	on	GW409544 (L-Tyr derivative)		$EC_{50} = 0.000\,28$	*9*
1i7i (a)	homodimer	on	tesaglitazar (AZ242)	$IC_{50} = 0.2$ μM in 250 nM rosiglitazone	$EC_{50} = 1.3$	*10*
1i7i (b)	homodimer	off	tesaglitazar (AZ242)			
1nyx (a)	homodimer	on	ragaglitazar (DRF2725)	$IC_{50} = 0.092 \pm 0.003$ μM	$EC_{50} = 0.6$	*11*
1knu (a)	homodimer	on	3q (carbazole analogue of ragaglitazar)	similar to ragaglitazar	$EC_{50} = 0.17$	*12*

[a] PDB code. The number in parentheses specifies the chain studied. [b] Position of the H12 helix. "On" and "off" denote active (or active-like) and inactive positions of the helix, respectively.

Step 2. Minimization and Rescoring. Step 1 produces a large number of protein−ligand complexes at various local minima of ΔG_s. The free energy of each complex is minimized using the more accurate free-energy potential $\Delta G = \Delta E_{elec} + \Delta E_{vdw} + \Delta G^*_{des}$, where ΔE_{vdw} denotes the receptor−ligand van der Waals energy, and the superscript in ΔG^*_{des} emphasizes that the desolvation term includes the change in the solute−solvent van der Waals interaction energy. The sum $\Delta E_{elec} + \Delta G^*_{des}$ is obtained by the analytic continuum electrostatics (ACE) model (*33*), as implemented in version 27 of Charmm (*37*, *38*) using the parameter set from version 19 of the program. The minimization is performed using an adopted basis Newton−Raphson method (*37*). During the minimization, the protein atoms are held fixed, while the atoms of the probe molecules are free to move.

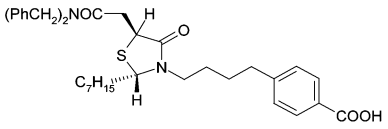
Step 3: Clustering and Ranking. The minimized probe conformations from Step 2 are grouped into clusters based on Cartesian coordinate information. The method creates a number of clusters such that the maximum distance between a cluster's hub and any of its members (the cluster radius) is smaller than half of the average distance between all of the existing hubs. We have slightly modified this algorithm by introducing an explicit upper bound of 4 Å on the cluster radius. Only clusters with more than 8 entries are retained (*19*−*21*). For each retained cluster, we calculate the partition function $Q_i = \sum_j \exp(-\Delta G_j/RT)$, obtained by summing the Boltzmann factors over the conformations in the *i*th cluster only. The clusters are ranked on the basis of their average free energies $\langle \Delta G \rangle_i = \sum_j p_{ij} \Delta G_j$, where $p_{ij} = \exp(-\Delta G_j/RT)/Q_i$ and the sum is taken over the members of the *i*th cluster.

Step 4. Determination of Consensus Sites. For each solvent, we select the minimum free-energy conformation in each of the five lowest average-free-energy clusters. The structures are superimposed, and the positions at which several such probes overlap are defined as consensus sites. The consensus sites are ranked according to the number of structures that they contain, each representing a low-energy probe cluster.

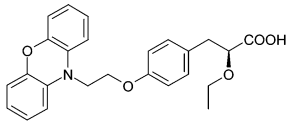
Step 5. Subcluster Analysis. For each ligand, the clusters at the consensus sites are further divided into subclusters based on probe orientations and free energies. The latter are included, because similar conformations with very different free energies usually have different mechanisms of binding (e.g., different hydrogen-bonding interactions), and hence it is preferable to group them into different subclusters (*19*). The subclusters of the *i*th cluster are ranked on the basis of the probabilities $p_{ij} = Q_{ij}/Q_i$, where Q_i is the sum of the Boltzmann factors over all conformations of the *i*th cluster and Q_{ij} is obtained by summing the Boltzmann factors over the conformations in the *j*th subcluster only. Each subcluster with $p_{ij} > 0.05$ was represented by a single conformation. The HBPLUS program (*39*, *40*) of Thornton and co-workers was used to find the nonbonded interactions and hydrogen bonds formed between each probe conformation and the protein. After counting all interactions, we have determined their distribution among the residues of the protein.

Targeted Solvent Mapping. Uniform mapping is adequate for finding relatively shallow binding regions on approximately spherical proteins but is far from ideal for PPARγ, which has a very deep binding site. To improve sampling, even in the uniform mapping, we restricted considerations to a box around the lower half of the LBD. The consensus sites from the uniform mapping identify the locations that bind clusters of organic solvents. The goal of targeted mapping is to determine, with higher accuracy, the number of clusters that stay in each such site after the probes have been moved around by the search. For each probe, we start 30 simplex minimization runs from the center of each site, thereby generating 30 docked conformations. Thus, considering seven different organic molecules as probes, 210 probes are distributed in each site. We have tested whether this number is sufficiently high by repeating some of the mapping calculations by increasing the number of starting points to seven and found no significant changes in the mapping results.

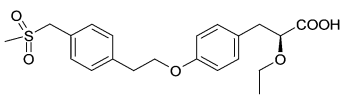
PPARγ Structures Studied. The structures mapped (Table 1) are the ligand-free PPARγ LBD [PDB code 1prg, (*5*, *6*)], LBD cocrystallized with the partial agonist GW0072 [PDB code 4prg (*7*)], and LBD structures with six different agonists, including roziglitazone [PDB codes 2prg and 1fm6,



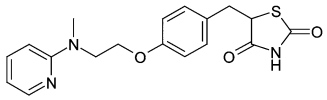
GW0072 (Partial agonist), PDB code 4prg



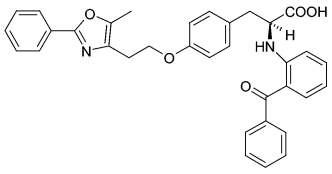
Ragaglitazar, PDB code 1nyx



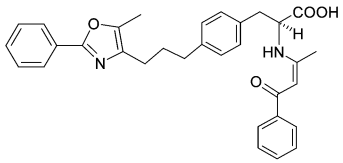
AZ242 (Tesaglitazar), PDB code 1i71



Rosiglitazone, PDB code 2prg and 1fm6



GI262570 (Farglitazar), PDB code 1fm9



GW409544, PDB code 1k74

FIGURE 1: PPARγ agonists included in this work, including the code of the corresponding structure in the PDB (*38*).

see refs *5* and *8*, respectively], two L-tyrosine derivatives GI252670, also called farglitazar [PDB code 1fm9 (*8*, *9*)] and GW409544 [PDB code 1k74 (*9*)], the dihydrocinnamate derivative tesaglitazar or AZ242 [PDB code 1i7i (*10*)], ragaglitazar [PDB code 1nyx (*11*)], and 3q, an agonist similar to ragaglitazar [PDB code 1knu (*12*)]. The structures of these agonists are shown in Figure 1, with the exception of 3q, which is very similar to ragaglitazar. PPARγ binds to DNA as a heterodimer with the RXR. Three of the X-ray structures (1fm6, 1fm9, and 1k74) are heterodimers and also include short fragments of the coactivator protein SRC-1. In the remaining six X-ray structures, the PPARγ LBD is a homodimer. In 1prg, 4prg, and 1i71, H12 is in the active conformation in chain a and in the inactive conformation in chain b. Therefore, we have mapped both chains in these three structures, resulting in a total of 12 mapping calculations.

RESULTS

Structure of the PPARγ LBD. To provide context for our results, we first summarize the structural properties of the PPARγ LBD (*5*−*12*). The domain is comprised of a three-layer antiparallel α-helical sandwich of 13 helices and a small four-stranded β sheet (Figure 2A). This architecture is very similar to that of other nuclear receptors, with the exception of an extra helix, designated H2′, between the first β strand and H3. The three long helices (H3, H7, and H10/H11) form the two outer layers of the sandwich. The middle layer of helices (H4, H5, H8, and H9) occupies the top half of the domain and is absent from the bottom half, thereby creating a very large cavity (∼1400 Å^3) for ligand binding. This large ligand-binding cavity has a distinct, three-arm Y shape, allowing PPARγ to bind ligands with multiple branches or singly branched ligands in multiple conformations. On its lower half, the right-hand side of the LBD is sealed by a two-stranded β sheet and, on the left-hand side, by the short C-terminal α helix (H12) of the receptor, which constitutes the receptor's ligand-dependent AF-2 region. A critical step during the activation process involves ligand-induced alteration of the conformation of H12 to an active or "on" position, shown in Figure 2A, which acts as a molecular switch and creates a binding cleft on the receptor for the coactivator (*2*, *3*). A cleft for the corepressor is formed in the same surface region when H12 is in the inactive or "off" position. In the unliganded LBD, the flexible pocket is believed to be in an equilibrium of conformational states and can adopt the active state even in the absence of an agonist (*2*, *13*, *14*).

Uniform Mapping. Seven solvent molecules were used as probes: acetone, acetonitrile, urea, methanol, 2-propanol, *tert*-butyl alcohol, and phenol. The initial probe positions were distributed in a box covering the entire lower half of the LBD such that the box included any point within 8 Å of the ligand, with no further a priori assumptions on the location of the binding site. We performed two sets of uniform mapping, first with boxes defined for each structure separately and then with a generic box applicable to all structures. Because the original algorithm for selecting initial probe positions assumes a globular protein shape with relatively small pockets, it was clear that there would be a limited number of points placed in the interior of the large and deep binding site. Although the results of the two sets of mapping calculations showed some dependence on the size of the box, both methods identified the same 10 locations in the lower half of the PPARγ LBD as consensus sites (Table 2 and Figure 2B). While no consensus site was found beyond these 10 positions, none of the 12 structures has all of the 10 sites. For reference, Figure 2B also shows the bound agonist farglitazar (GI262570) from the structure 1fm9 (see Table 1).

Four of the sites identified (P1−P4) are within the ligand-binding site (Figure 2B). Site P1 is on the left arm of the Y-shaped cavity, adjacent to the carboxyl group of the PPARγ ligand GI262570, which makes hydrogen bonds with

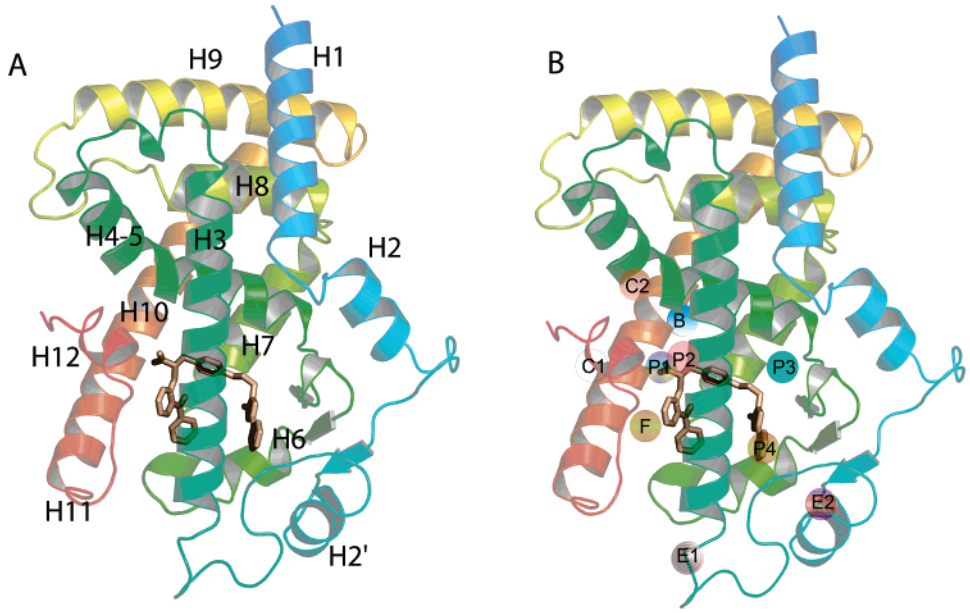


FIGURE 2: Structure of the PPARγ LBD. (A) Polypeptide backbone is shown as a cartoon, indicating the 12 α helices that comprise the domain. The agonist farglitazar (GI262570, from PDB structure 1fm9) is shown as a stick representation. (B) Approximate location of the "hot spots" or sites identified by the uniform mapping. Sites P1−P4 are subsites of the ligand-binding site of PPARγ, with P1 located at the TZD or carboxyl headgroup of the bound agonist, P2 slightly to the right, and P3 and P4 at the upper and lower distal ends, respectively, of the ligand-binding site. Sites F and B are, respectively, in the front and back of the LBD, with the latter in the dimerization domain. Sites C1 and C2 are located in the region of coactivator binding, with C2 overlapping with the SRC-1 peptide. Finally, sites E1 and E2 are in channels leading to the binding site, with E2 overlapping the putative ligand entrance. Figures were prepared using PyMol (*43*).

Table 2: "Hot Spots" of PPARγ Identified by Uniform Mapping

site[a]	description	surrounding residues[b]
P1	binding the TZD and carboxyl groups of agonists and interacting with helix H12	F282, C285, Q286, S289, H323, Y327, F363, H449, L469, Y473
P2	buried between H3 and H5 and overlapping the middle of the agonists	C285, R288, S289, I326, Y327, L330, F363, M364
P3	upper distal end of the binding site between H2 and H3 and reached only by the partial agonist GW0072	R288, E291, A292, E295, M329, E343
P4	hydrophobic pocket close to the entrance and close to the distal-end group of agonists	S255, E259, F264, H266, V277, A278, R280, I281, G284, F287, I341, S342, M348
B	surface pocket in the back between H7 and H10/H11 and overlapping the dimerization region	P366, E369, F370, K373, D441, Q444, I445, E448
F	surface pocket between H3 and H12	Q283, Q286, F287, S464, H466, L469
C1	surface pocket between H10 and H12 and possibly contributing to cofactor binding	V450, L453, Q454, K457, M463, L465, Q470, Y473, K474, D475
C2	between H5 and H12 and overlapping with the binding site of the coactivator peptide SRC-1	V293, T297, V315, L318, K319, T447, L468, E471, I472
E1	pocket defined by the lower ends of helices H3, H7, and H10	E276, I279, R357, F360, M364, E460, D462
E2	putative ligand entrance between H2′ and the β sheet	K244, F247, G258, E259, K261, I262, F264, K265, I341, S342, Q345, G346

[a] "Site" and "hot spot" will be used interchangeable for the 10 positions shown in Figure 2B. Sites P1−P4 are in the ligand-binding pocket; B and F, in the back and front of the LBD, respectively; C1 and C2, in the coactivator-binding region; and E1 and E2, in the ligand-entrance region. [b] One letter amino acid codes.

residues S289, H323, H449, and Y473. This largely hydrophilic pocket is important for the binding of all strong agonists and accommodates the polar TZD or carboxyl headgroup that interacts with H12. In the structure with GI262570, the cavity extends downward to accommodate the benzophenone tail of the agonist. The P2 site is buried between helices H3 and H5, slightly to the right of P1. Although sites P1 and P2 share a number of amino acids, some important residues, primarily Q286, H323, and Y473, are part of P1 but not of P2 (Table 2). Site P3 is at the end of the right arm of the Y-shaped cavity, defined by helix H3 from the left, helix H2 from the right, the loop connecting helices H1 and H2 from the top, and helix H5 from the back. P3 is out of the reach of the agonist GI262570, toward to the top of the LBD domain. The only ligand known to reach into this site is the partial agonist GW0072, which places one of the two benzylamide groups in P3. The P3 site is a large pocket, which is at least partially open in all PPARγ structures. Site P4 surrounds the 5-methyl-2-phenyloxazole group on the distal end of GI262570. The largely hydrophobic pocket is defined by H3 from the left and by the β sheet and helix H2′ from the right. Short ligands such as rosiglitazone do not extend far enough to reach the P4 site.

The other "hot spots" found by the mapping are outside the PPARγ ligand-binding site. Site B is at the back of LBD, between helices H7 and H10 (see Figure 2B) and is part of the dimerization region, involved in forming the PPARγ−RXRα heterodimer. Site F is between helices H3 and H12

Table 3.

Ranking of Probe Clusters within Consensus Sites for Ligand-Free and Partial-Agonist-Bound Structures of PPARγ Based on the Results of Targeted Mapping

PDB ID and chain[b]	consens. site[c]	site[d]	probe[a]						
			acetonitrile	acetone	*tert*-butanol	phenol	methanol	2-propanol	urea
1prg (a)	1	P3	1 (4.73)	1 (8.87)	1 (6.44)	1 (7.72)	1 (9.61)	1 (8.48)	3 (8.71)
			5 (8.67)			2 (5.90)			
	2	P2	4 (1.05)	4 (0.93)	2 (0.88)	5 (0.46)	5 (0.60)	4 (1.22)	4 (0.50)
				5 (0.51)					5 (0.51)
	3	P4	3 (2.30)	3 (3.33)	3 (0.25)	3 (2.73)	3 (4.04)	3 (1.82)	2 (3.83)
					4 (0.61)				
	4	C1	2 (7.77)	2 (8.14)	5 (6.76)	4 (8.35)	2 (8.11)	2 (8.23)	1 (7.84)
	5	C2					4 (11.99)	5 (11.08)	
1prg (b)	1	P2	3 (1.13)	3 (0.71)	3 (0.61)	3 (0.41)	3 (2.09)	3 (1.12)	4 (0.92)
					4 (0.47)	4 (0.43)	5 (0.38)	4 (0.66)	
	2	P3	1 (5.27)	2 (9.00)	1 (6.42)	2 (8.15)	1 (9.94)	2 (8.91)	2 (9.20)
						5 (6.36)			3 (8.392)
	3	P4	2 (7.03)	1 (4.01)	2 (0.42)	1 (3.15)	4 (4.41)	1 (1.91)	1 (3.20)
									5 (4.39)
	4	C2		4 (8.12)				5 (9.08)	
	5	B	4 (11.72)						
			5 (14.28)						
4prg (a)	1	P3	5 (2.84)	3 (0.61)	1 (1.72)	1 (0.64)	1 (1.26)	2 (0.52)	5 (1.08)
						2 (1.56)			
	2	F	2 (10.8)	2 (11.39)	3 (9.73)	4 (10.96)	4 (11.23)	4 (10.82)	4 (10.72)
	3	P4		1 (0.29)	2 (0.17)	5 (0.37)		1 (0.16)	1 (0.17)
	4	C1	4 (12.35)	4 (12.10)				3 (12.49)	2 (11.08)
	5	P2	1 (1.68)				3 (4.00)		3 (2.81)
	6	B			4 (11.08)	3 (11.86)	5 (11.88)		
	7	C2		5 (14.73)	5 (13.91)			5 (15.49)	
4prg (b)	1	P4	1 (0.61)	2 (0.70)	1 (0.56)	1 (0.37)	1 (2.06)	1 (0.28)	1 (0.54)
	2	P3	2 (4.26)		2 (0.74)	4 (0.53)	3 (3.03)	2 (2.28)	4 (1.62)
	3	F	3 (8.91)	1 (9.53)		3 (6.77)		3 (8.79)	
				5 (6.64)					
	4	P2		3 (0.43)			4 (1.17)	4 (0.71)	2 (2.61)
									3 (1.01)
	5	B	4 (12.48)			2 (13.33)	5 (13.90)		5 (11.51)

Ranking of Probe Clusters within Consensus Sites for Agonist-Bound Structures of PPARγ Based on the Results of Targeted Mapping

PDB ID and chain[b]	consens site[c]	site[d]	probe[a]						
			acetonitrile	acetone	*tert*-butanol	phenol	methanol	2-propanol	urea
2prg (a)	1	P1	2 (0.71)	3 (0.34)	1 (0.27)	1 (0.88)	4 (0.41)	1 (0.77)	2 (0.85)
	2	F		1 (6.67)	4 (7.30)	3 (6.32)	3 (6.32)	2 (6.12)	3 (6.26)
	3	C2	4 (10.22)	2 (10.71)	5 (10.04)		5 (11.65)	3 (10.83)	5 (9.52)
	4	P3	1 (7.73)			5 (6.00)	1 (8.54)	4 (7.38)	1 (7.70)
	5	P4	5 (4.20)		2 (3.86)	2 (2.57)	2 (5.16)		
						4 (3.12)			
	6	E2	3 (7.95)	4 (8.12)	3 (8.42)				4 (7.57)
	7	C1		5 (8.05)				5 (8.46)	
1fm6 (d)	1	P1	4 (0.43)	1 (0.56)	1 (0.67)	1 (0.54)	2 (0.35)	1 (0.25)	3 (0.46)
				3 (1.25)			5 (1.99)		
	2	P3	1 (8.03)	5 (7.64)	5 (6.57)	2 (7.78)	3 (9.66)	4 (7.72)	2 (7.75)
	3	C2	5 (10.35)	4 (9.46)	4 (9.58)	5 (9.62)		5 (9.97)	5 (9.36)
	4	P4	2 (5.47)	2 (5.88)	2 (3.29)	4 (0.12)		2 (5.39)	
	5	C1	3 (7.73)		3 (6.24)			3 (7.41)	
	6	E2					1 (8.80)		1 (8.03)
	7	F					4 (7.36)		4 (7.20)
1fm9 (d)	1	P4	3 (3.31)	5 (0.78)	4 (0.69)	1 (2.05)		3 (2.65)	1 (2.78)
					5 (1.13)	5 (0.61)			
	2	P1	4 (0.57)	2 (0.80)	2 (0.45)	2 (0.66)	1 (0.71)	1 (0.67)	
	3	E2	2 (6.22)	4 (5.94)			2 (7.88)	2 (6.94)	5 (6.31)
	4	C1	1 (7.67)			4 (5.94)	5 (7.95)	4 (7.78)	4 (6.20)
	5	E1	1 (7.47)		1 (7.49)		3 (8.31)		2 (7.86)
	6	P3	5 (8.19)			3 (7.38)			3 (8.01)
	7	C2		3 (9.32)	3 (9.42)			5 (10.07)	
1i7i (a)	1	P4	5 (0.53)	3 (0.81)	5 (0.54)	5 (0.37)	5 (1.16)	4 (0.90)	4 (3.94)
				5 (9.84)				5 (0.87)	5 (1.02)
	2	P1	2 (0.34)	1 (1.21)	1 (0.93)	1 (0.61)	2 (1.04)	1 (0.24)	2 (0.89)
						4 (1.10)			
	3	P3	1 (8.50)			2 (8.02)	1 (9.65)	2 (8.34)	1 (8.92)
	4	C1	3 (7.35)	2 (7.79)	3 (8.50)		3 (7.77)		3 (6.59)
	5	C2		4 (9.53)	4 (9.81)		4 (9.36)	3 (10.41)	
	6	B	4 (12.12)		2 (11.29)	3 (11.87)			

Table 3 (Continued)

Ranking of Probe Clusters within Consensus Sites for Agonist-Bound Structures of PPARγ Based on the Results of Targeted Mapping

PDB ID and chain[b]	consens site[c]	site[d]	probe[a]						
			acetonitrile	acetone	*tert*-butanol	phenol	methanol	2-propanol	urea
1i7i (b)	1	P4	2 (3.39)	1 (3.27)	1 (1.66)		1 (3.77)	1 (2.14)	2 (3.25)
			3 (0.46)				3 (0.85)		
	2	P1	4 (0.27)	3 (0.62)	2 (0.64)	2 (0.56)	5 (0.80)	2 (0.52)	5 (0.38)
	3	P3	1 (8.84)	2 (7.65)		1 (7.49)	2 (9.69)	3 (8.10)	1 (8.67)
	4	B	5 (14.46)		4 (13.89)	4 (13.09)		5 (14.69)	4 (13.46)
	5	F		5 (8.76)	5 (9.55)	5 (8.78)	4 (8.32)		3 (7.48)
	6	C1		4 (7.64)	3 (8.22)	3 (7.92)		4 (8.72)	
1k74 (d)	1	P1	3 (0.70)	1 (0.92)	2 (0.27)	1 (0.38)	2 (0.64)	1 (0.60)	4 (0.72)
				5 (0.76)		2 (0.50)			
	2	E1	1 (8.46)		1 (8.94)	3 (8.83)	4 (8.96)	4 (8.91)	3 (8.45)
	3	C1	4 (7.10)	3 (7.95)	4 (6.97)		3 (7.79)		5 (6.28)
	4	P4		2 (0.66)	5 (0.85)	5 (0.53)	5 (2.64)	2 (1.13)	
	5	P3	2 (7.62)				3 (9.41)		2 (8.41)
	6	C2		4 (10.15)	3 (10.27)			5 (10.88)	
	7	E2					1 (8.23)		1 (7.06)
1knu (a)	1	P4	1 (4.45)	1 (3.55)	1 (0.47)	3 (0.82)	4 (1.38)	2 (3.21)	
			4 (1.48)		5 (0.63)			3 (0.75)	
	2	P1	3 (1.07)	2 (0.85)		1 (0.67)	1 (0.67)	1 (1.05)	1 (0.79)
	3	C1	5 (7.74)	4 (6.77)	2 (6.72)	5 (7.97)		5 (7.98)	4 (6.13)
	4	C2		3 (8.46)	3 (8.73)	4 (10.04)	3 (10.87)	4 (9.66)	
	5	P3	2 (4.73)		4 (3.85)	2 (4.64)			2 (4.49)
	6	E2		5 (8.93)			2 (9.09)		3 (8.24)
	7	F					5 (8.35)		5 (7.43)
1nyx (a)	1	P1	2 (0.27)	1 (1.28)	4 (1.00)	1 (0.72)	1 (1.64)	1 (0.92)	2 (1.15)
				5 (1.72)				3 (1.44)	
	2	C2	4 (8.16)	3 (8.43)	1 (7.87)	2 (8.99)	4 (9.85)	2 (9.03)	5 (8.52)
	3	C1	5 (7.58)	4 (8.27)	2 (7.69)	3 (8.28)		4 (7.92)	4 (8.12)
	4	B	1 (11.85)		3 (11.88)		2 (11.37)	5 (11.54)	
	5	P3	3 (4.77)			4 (4.95)	5 (6.18)		3 (5.75)
	6	E2		2 (10.02)			3 (10.42)		
	7	P4			5 (4.52)	5 (0.91)			
							1 (3.77)	1 (2.14)	2 (3.25)

[a] Free-energy rank of the probe cluster in the consensus site, where 1 denotes the lowest free-energy cluster. The number in parentheses shows the distances (in angstroms) between the cluster center (the lowest free-energy ligand conformation in the cluster) and the nearest atom of the bound ligand (agonist or partial agonist). A cell with no entry indicates that the particular probe does not form a cluster in the consensus site. [b] Chain identifier is given in parentheses. [c] Consensus site ranked according to the number of clusters contained. [d] See Table 2 and Figure 2B for description and location, respectively, of the "hot spots".

but on the surface rather than in the ligand-binding site. Sites C1 and C2 are near helix H12, close to the putative region of coactivator/corepressor binding. Site C2, located between helices H12 and H5, overlaps with the binding site of the SRC-1 coactivator peptide, which is present in several of the PPARγ X-ray structures. C1 is on the opposite side of H12, between H11 and H12. Sites E1 and E2 are surface pockets at the openings of the ligand-binding site. Site E2, defined by helices H2′ and H3 and the β sheet, is located in a position that is frequently mentioned as the putative entrance to the binding site. Site E1 is located between the lower ends of H3 and H7.

Because the results of targeted mapping describe the relative importance of the individual sites better than those from uniform mapping, the latter are not discussed further. However, we emphasize that the 10 "hot spots" are consistently determined by the uniform mapping method without any a priori assumptions about their locations. We recall that to determine consensus sites we consider the centers (i.e., the lowest energy conformations) of the five lowest free-energy clusters of each probe, i.e., a total of 35 probe positions. In all 12 mapped PPARγ structures, any 2 or more of these probe positions overlap in 1 of the 10 "hot spots". Thus, we conclude that these 10 locations are the primary sites accessible for the binding of small probes and that they are likely to play some roles in the recognition of other molecules by PPARγ.

Targeted Mapping. To further explore the 10 sites identified by the uniform mapping, additional mapping calculations were carried out from initial points placed in each of the sites. Our goal was to determine which sites are capable of retaining most of the probe clusters when the probes are moved toward their respective free-energy minima using a nonlinear simplex method (*19*). The five lowest-free-energy clusters for each probe were superimposed to find the consensus sites shown in Table 3. The integers in this table represent the ranking of probe clusters in term of their free energy. For example, the first row for 1prg (a) shows that the largest consensus site 1 found by the mapping is located in the P3 site, which includes the first and fifth lowest free-energy clusters of acetonitrile, the first lowest free-energy cluster of acetone, and so on. The numbers in parentheses show the distances between the cluster center (the lowest free-energy ligand conformation in the cluster) and the nearest atom of the bound ligand (agonist or partial agonist) in the particular structure. Because 1prg is ligand-free, we placed the agonist GI262570 in its binding site by overlapping the structures 1prg (a) and 1fm9 (d). Thus, the entry "1 (4.73)" in the top line for 1prg (a) in Table 3 indicates that the center of the lowest free-energy cluster for aceto-

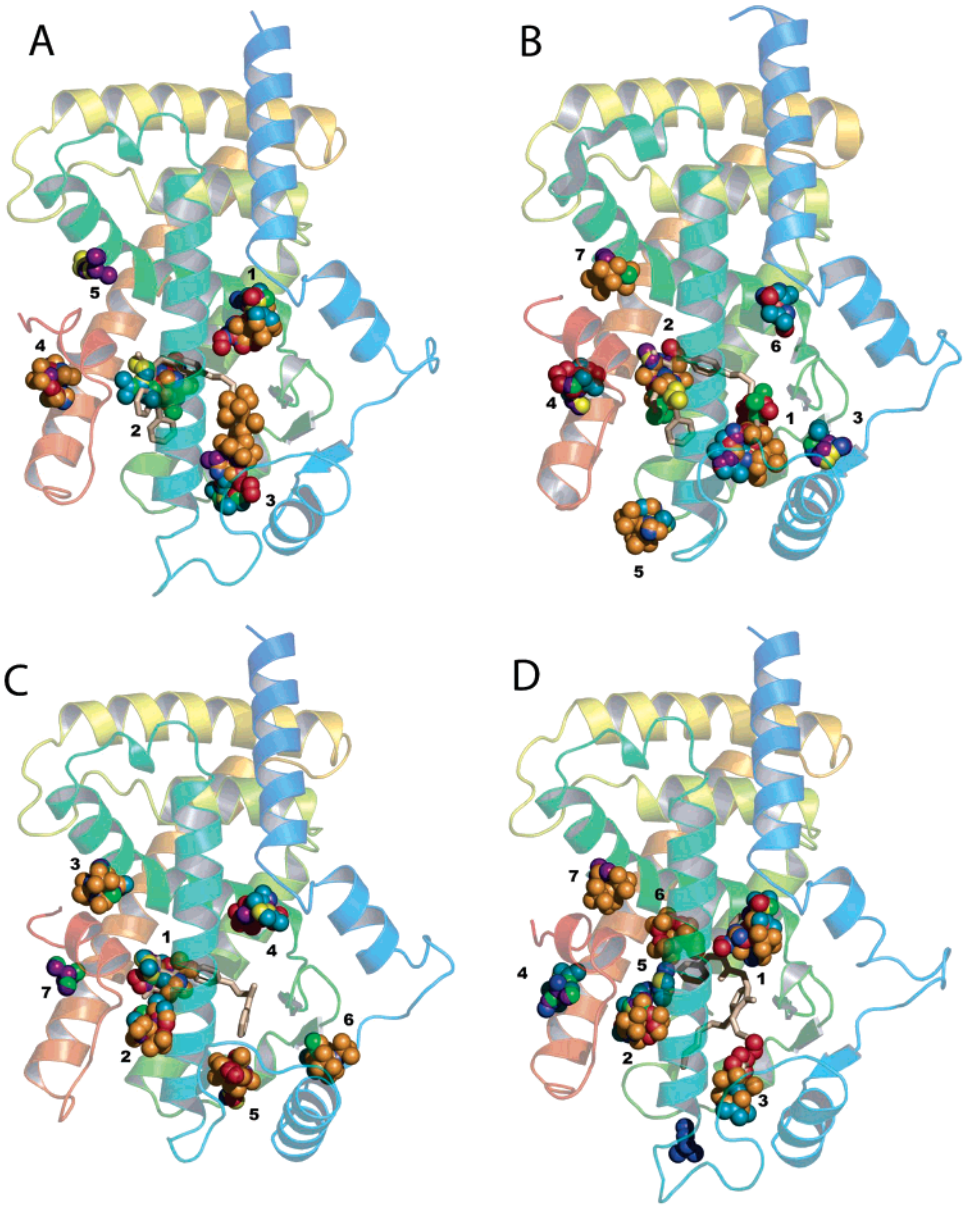


FIGURE 3: Largest consensus sites in selected PPARγ LBD structures based on the 5 lowest free-energy clusters identified by targeted mapping. The color scheme used for the ligands is violet, 2-propanol; green, acetone; blue, acetonitrile; orange, *tert*-butyl alcohol; red, phenol; yellow, methanol; and teal, urea. The protein and ligands are shown in cartoon and space-filling representations, respectively. Targeted mapping was performed on the apoprotein; the bound agonists are shown superimposed for reference only. From the consensus site with the largest number of probe clusters (consensus site 1 in Table 3), 5−7 of the largest consensus sites are shown. (A) 5 largest consensus sites for the active-like conformation of the ligand-free PPARγ LBD structure (PDB code 1prg, chain a). The agonist farglitazar (GI262570, from PDB structure 1fm9) is shown for reference. As shown in Table 3, the consensus sites numbered on the LBD structure correspond to the following sites: 1, site P3; 2, site P2; 3, site P4; 4, site C1; and 5, site C2 (cf. Figure 2B). (B) 7 largest consensus sites for the PPARγ LBD, cocrystallized with the agonist farglitazar (PDB code 1fm9). As shown in Table 3, the locations of the consensus sites are 1, site P4; 2, site P1; 3, site E2; 4, site C1; 5, site E1; 6, site P3; and 7, site C2. (C) 7 largest consensus sites for the PPARγ LBD, cocrystallized with the agonist rosiglitazone (PDB code 2prg). The locations of the consensus sites are 1, P1; 2, F; 3, C2; 4, P3; 5, P4; 6, E2; and 7, C1. (D) 7 largest consensus sites for the active-like conformation of the PPARγ LBD structure, cocrystallized with the partial agonist GW0072 (PDB code 4prg for chain a). The locations of the consensus sites are 1, site P3; 2, site F; 3, site P4; 4, site C1; 5, site P2; 6, site B; and 7, site C2 (see Table 3).

nitrile is 4.73 Å from the closest atom of the GI262570 ligand. For all other structures, we use the native ligand in the binding site for reference.

Figure 3A shows the results of mapping for chain a of the apo structure 1prg. Mapping was performed on the ligand-free structure, but the figure also includes the agonist farglitazar (GI262570) in the binding site for reference. Each consensus site is ranked on the basis of the number of probe clusters that it contains. As shown in Table 3, the largest consensus sites are P3, P2, and P4, with 9, 9, and 8 clusters of probes, respectively. Thus, a large fraction of the binding site is readily accessible even in the ligand-free structure. However, the probe molecules cannot reach site P1, on the left end of the binding site, in contact with H12. In fact, site P2 is located at the benzophenone moiety of the agonist, rather than at its carboxyl headgroup. In chain a of 1prg, helix H12 is in the active conformation, and Figure 3A shows that sites C1 and C2 in the coactivator-binding region are

FIGURE 4: Distribution of intermolecular interactions among ligands and individual residues of PPARγ LBD. The interactions were determined from two sources: extracted from the PDB structure using the HBPLUS program (on the left side) and by computational mapping. Computational mapping results are based on the interactions found between various PPARγ residues and the probes in the 10 pockets P1−E2. The color scheme used is as follows: white, no interaction; wheat, less than 2% of the total; coral, between 2 and 10%; and crimson, >10%. (A) Nonbonded interactions. (B) Hydrogen bonds.

already formed (consensus sites 4 and 5, respectively, in Figure 3A), albeit site C2 contains only two probe clusters. We note that for 1prg most probe clusters with the lowest or second lowest free energy are in site P3 (see Table 3), further supporting that P3 is a well-formed pocket prior to ligand binding.

Figure 3B shows results from the mapping of the LBD in the structure 1fm9, cocrystallized with the agonist farglitazar (GI262570). As noted above, the agonist was removed during the mapping calculations and is shown only for reference. In this case, the two largest consensus sites (ranked as 1 and 2) are P4 and P1, with 8 and 6 clusters, respectively. The probe clusters in site P1 cover both the carboxyl headgroup and part of the benzophenone moiety. A comparison of parts A and B of Figure 3 reveals that, in the agonist-bound LBD, the clustering of the probes occurs deeper in the binding site than it does in the ligand-free structure. These differences will be examined in detail in the Discussion. The ligand-bound structure also contains site P3, ranked 6, as well as sites C1 and C2 in the coactivator-binding region (consensus sites 4 and 7, respectively, in Figure 3B). However, P3 is smaller, and C1 and C2 are both larger than the corresponding sites in the ligand-free structure. Agonist binding results in two new sites, E1 and E2 (consensus sites 5 and 3, respectively, in Figure 3B), whose potential origin will be discussed further in this paper. Note that five of the six probe clusters in Site P1 have the lowest or second lowest free energy (Table 3). Table 3 shows that this is the case for almost all agonist-bound structures, indicating that the binding of a strong agonist results in a

Table 4.

rmsd Values for R280 (Upper Triangle) and I281 (Lower Triangle)[a]

	1prg (a)	1prg (b)	2prg (a)	4prg (a)	1fm9 (d)	1k74 (d)	1i7i (a)	1nyx (a)
1prg (a)	0	0.98	0.94	0.84	1.77	3.89	1.13	0.88
1prg (b)	0.50	0	1.31	1.03	1.38	3.69	1.41	1.37
2prg (a)	0.56	0.99	0	1.39	1.58	3.75	0.84	0.85
4prg (a)	0.49	0.32	0.95	0	1.94	3.64	1.47	1.39
1fm9 (d)	1.35	1.64	0.95	1.70	0	0.87	1.33	1.82
1k74 (d)	1.00	1.21	0.75	1.26	0.77	0	3.57	4.35
1i7i (a)	1.17	1.38	0.90	1.38	1.12	0.99	0	1.25
1nyx (a)	0.37	0.80	0.50	0.76	1.29	1.10	1.14	0

rmsd Values for F282 (Upper Triangle) and Q286 (Lower Triangle)[a]

	1prg (a)	1prg (b)	2prg (a)	4prg (a)	1fm9 (d)	1k74 (d)	1i7i (a)	1nyx (a)
1prg (a)	0	0.46	0.66	0.59	2.36	2.29	1.21	0.58
1prg (b)	2.41	0	0.78	0.67	2.11	2.09	1.16	0.60
2prg (a)	1.58	1.54	0	1.06	2.18	2.16	0.93	0.48
4prg (a)	1.39	1.90	1.07	0	2.39	2.24	1.65	0.87
1fm9 (d)	1.23	2.47	2.14	1.83	0	0.65	2.18	2.00
1k74 (d)	1.01	2.31	1.92	1.57	0.87	0	2.29	2.02
1i7i (a)	1.34	1.94	1.10	1.42	1.71	1.49	0	1.00
1nyx (a)	0.73	2.25	1.50	1.60	1.30	1.06	0.96	0

rmsd Values for Q288 (Upper Triangle) and R363 (Lower Triangle)[a]

	1prg (a)	1prg (b)	2prg (a)	4prg (a)	1fm9 (d)	1k74 (d)	1i7i (b)	1nyx (a)
1prg (a)	0	5.60	4.45	2.34	4.11	4.24	3.76	4.39
1prg (b)	3.22	0	3.20	5.81	3.69	3.78	3.60	3.35
2prg (a)	3.78	2.19	0	4.58	0.69	0.82	0.93	0.38
4prg (a)	2.58	1.70	2.52	0	4.28	4.34	4.01	4.53
1fm9 (d)	0.84	3.20	3.52	2.45	0	0.56	0.56	0.55
1k74 (d)	0.64	2.96	3.42	2.35	0.48	0	0.80	0.67
1i7i (a)	3.15	2.13	1.51	2.02	2.81	2.73	0	0.72
1nyx (a)	0.93	2.62	3.26	2.15	0.99	0.64	2.67	0

rmsd Values for H449 (Upper Triangle) and Y473 (Lower Triangle)[a]

	1prg (a)	1prg (b)	2prg (a)	4prg (a)	1fm9 (d)	1k74 (d)	1i7i (a)	1nyx (a)
1prg (a)	0	0.46	1.28	0.89	1.56	1.46	1.63	1.37
1prg (b)	10.93	0	1.21	0.97	1.34	1.27	1,47	1.23
2prg (a)	0.94	10.12	0	1.63	0.73	0.64	0.74	0.44
4prg (a)	2.31	10.65	1.82	0	1.95	1.87	2.08	1.69
1fm9 (d)	0.84	10.62	0.91	2.57	0	0.21	0.41	0.47
1k74 (d)	0.86	10.39	0.26	1.92	0.73	0	0.51	0.44
1i7i (a)	1.19	11.23	1.51	1.51	0.64	1.33	0	0.50
1nyx (a)	1.11	11.86	1.67	3.25	0.96	1.51	0.74	0

[a] rmsd values, in angstroms, between the positions of a specified side chain in pairs of selected structures. The first structure is given in the corresponding row of column 1, and the second is given in the corresponding column of row 1. Because the rmsd values are symmetric, the upper and lower triangles of the table are used to list the rmsd values for two different side chains. By definition, any entry in the diagonal is 0.

high-affinity pocket at site P1. The only exception is chain b of 1i71, which is an anomaly, probably because of crystal contacts.

For the rosiglitazone-bound structure 2prg, the largest clusters are at sites P1, F, and C2, respectively (consensus sites 1, 2, and 3, respectively, in Figure 3C). As for 1fm9, most low-free-energy clusters are in pocket P1. Sites P3 and P4 are the next largest clusters. Note that rosiglitazone is too short to reach the P4 site, which is substantially smaller in this structure than in 1fm9 (5 probe clusters rather than 8). The TZD group of rosiglitazone is also smaller than the corresponding hydroxyl and benzophenone headgroups in farglitazar, and the extra space leads to the formation of pocket F, not present in 1fm9. As in 1fm9, the coactivator-binding region contains sites C1 and C2, but the latter is now substantially more dominant than in the farglitazar-bound structure.

Figure 3D shows the mapping of chain a structure 4prg, in which PPARγ is cocrystallyzed with the weak partial agonist GW00720. The binding mode of GW00720 differs substantially from that of the other agonists. First, one of the benzyl rings of GW00720's dibenzylamide group is present at site P3, which is not occupied by any of the strong agonists. The second major difference is that the carboxyl group of GW0072 binds at site P4 rather than at site P1 as for all other agonists. The mapping identifies sites P3, F, and P4 (consensus sites 1, 2, and 3, respectively, in Figure 3D). The next largest clusters were found at sites C1, P2, B, and C2. Most of the sites in this structure are very similar to those seen in the ligand-free LBD (compare parts A and D of Figure 3). Indeed, sites P3 and P4 are similarly large, and site P1 does not exist. Because the AF-2 motif in chain a of 4prg has an active-like conformation, we find sites C1 and C2 in the coactivator-binding region, although the latter contains only 3 probe clusters. Changes in the conformations of the pockets are also shown by the fact that the low-free-energy probe clusters now distribute between pockets P3 and P4 rather than binding to the P1 site as in the other agonist-

bound structures. P4 is clearly dominating in chain b, suggesting that the binding of the carboxyl group of GW0072 at this site rearranges the residues lining the pocket, resulting in higher affinity for the probes.

Interactions with Individual Residues. To gain insight into the roles of individual amino acid residues in the binding site of the PPARγ LBD, the probe clusters in the 10 sites P1−E2 were divided into subclusters based on probe orientations to find the residues interacting with the probes (see the Materials and Methods). Subclustering showed that each probe molecule binds in a number of rotational states, with the nonpolar moiety generally located in a hydrophobic pocket and the polar part pointing toward various polar patches on the protein, in some cases forming a hydrogen bond. After a representative conformation was selected from each subcluster, we counted the nonbonded interactions and hydrogen bonds between the probes and the protein using the HBPLUS program (*39*, *40*) and determined their distribution among the amino acid residues. Using a color scheme to represent interaction frequency, Figure 4 compares these distributions to the ones based on the interactions extracted from the 10 PPARγ complexes in the PDB (*41*), with the experimentally observed interaction shown on the left and the mapping results shown on the right of each column. The color scheme used is as follows: white, no interaction; wheat, less than 2% of the total; coral, between 2 and 10%; and crimson, over 10% of the total interactions found for the entire LBD structure.

We note that Figure 4A show only residues that interact with the probes in at least 3 structures; i.e., we assume that interactions occurring only in 1 or 2 structures of the 12 studied are not robust enough to provide reliable prediction of ligand-binding residues. Indeed, after the removal of the residues that have only two or fewer interactions with probes, mapping predicts the ligand-binding residues very well, with all important residues correctly identified (Figure 4A). Note that the partial-agonist-bound structures 4prg (a) and 4prg (b) do not have site P1, and hence, the probes for these proteins do not interact with residues Q286, H323, and Y473. Apart from a weak false-positive interaction with Y473 in 4prg (a), the mapping clearly discriminates the two partial agonist-bound structures from the rest, all cocrystallized with agonists. However, we predict some interactions with H449 in both 4prg (a) and 4prg (b), indicating that this residue reaches into site P2, although the partial agonist does not interact with it. There are a few other false positives, e.g., for E259, R280, and F287, and a few false negatives, e.g., for F282 and C330. However, considering that our calculations have been carried out with no a priori information on the binding site, the agreement with experimental data is excellent.

Table 4 shows all residues that form hydrogen bonds with the probes. There are no interactions that occur only in one or two structures, and hence, we do not need the filter used for the nonbonded interactions. However, we predict more hydrogen bonds than observed experimentally. In fact, the small polar probes tend to form hydrogen bonds with all potential hydrogen-bond donor and acceptor groups in the pockets, although those group may not form any hydrogen bonds with the particular agonist. According to Figure 4B, false-positive hydrogen bonds occur for residues E259, Q271, R280, I281, G284, C285, and Y327, in some cases involving the backbone. Although these residues are in hydrogen-bonding positions, they do not actually form hydrogen bonds with the agonist in any of the structures. The residues that hydrogen bond with the agonists (S289, H323, H449, and Y473) or with the partial agonist (R288 and S342) are correctly identified by the mapping.

Residues with Substantive Conformational Change. The mapping results shown in parts A and B of Figure 4 enable us to identify amino acid residues that frequently interact with the probes and hence play important roles in defining the 10 specific sites. One of the goals of this study is to elucidate how and why the importance of these sites changes upon the binding of various PPARγ ligands. To address this question, we have considered the residues listed in Figure 4A and calculated their root-mean-square deviation (rmsd) values for each pair of structures. Because these results constitute a very large file, in Table 4, we list the pairwise rmsd values for 8 selected structures and for the amino acid residues that both play important roles in defining the pockets and are subject to considerable conformational change. Thus, some residues that are very important, e.g., S289 and H323, are not considered, because they occupy an almost invariant position in the different structures. The significance of these distances is discussed in the next section.

DISCUSSION

Table 5 summarizes the results extracted from the detailed consensus site in Table 3. For each of the 12 PPARγ structures analyzed by computational solvent mapping, Table 5 shows the number of clusters found in each of the 10 sites. The value in the parentheses indicates the rank of the consensus site based on the number of clusters found. For example, for 1prg (a), sites P2 and P3 both have 9 different probe clusters. Because no other site has so many clusters, sites P2 and P3 are ranked 1 and 2. The ranking of sites with the same number of clusters is arbitrary and was primarily used to identify the different sites in parts A−D of Figure 3. In this section, we show that the mapping results summarized in Table 5 provide a substantial amount of information on the 10 solvent-binding "hot spots" of PPARγ.

Robustness of Mapping Results. Before discussing the individual sites, we note that the 12 PPARγ structures studied in this paper come in pairs, and this is clearly expressed in the results of Table 5. In chains a and b of the ligand-free structure 1prg, helix H12 is in active and inactive conformations, respectively, and therefore, we expect differences in the numbers of probes bound in the coactivator-binding region. Interestingly, this difference is seen at site C1 rather than at site C2, which binds the SRC-1 coactivator peptide. All other sites are essentially the same in the two chains, as emphasized by the bold font in Table 5. Similarly, chains a and b of the partial agonist-bound structure 4prg have H12 in active and inactive conformations, respectively. In this case, both C1 and C2 are missing in the inactive chain b, but all other sites are again similar. 2prg and 1fm6 are both roziglitazone-bound structures (see Table 1). Although 2prg is crystallized as a homodimer and 1fm6 is crystallized as a PPARγ/RXRα heterodimer, the mapping results for the two structures are very similar. The only important difference is at site F, and we will show that this is due to a small change in the conformation of the Q286 side chain. Structures 1fm9

Table 5: Number of Clusters and Ranking of the Binding Sites, Determined by Targeted Mapping[a]

	PPARγ chain												
site	1prg (a)	1prg (b)	4prg (a)	4prg (b)	2prg (a)	1fm6 (d)	1fm9 (d)	1k74 (d)	1i7i (a)	1i7i (b)	1knu (a)	1nyx (a)	description
P1					**7 (1)**	**9 (1)**	**6 (2)**	**9 (1)**	**8 (2)**	**7 (2)**	**6 (2)**	**9 (1)**	site for TZD and carboxyl groups of agonists
P2	**9 (2)**	**11 (1)**	**3 (5)**	**5 (4)**									buried between H3 and H5
P3	**9 (1)**	**9 (2)**	**8 (1)**	**6 (2)**	**5 (4)**	**7 (2)**	**3 (6)**	**3 (5)**	**5 (3)**	**6 (3)**	**4 (5)**	**4 (5)**	upper distal end of the binding site
P4	**8 (3)**	**9 (3)**	5 (3)	7 (1)	**5 (5)**	**5 (4)**	8 (1)	5 (4)	**10 (1)**	**8 (1)**	9 (1)	2 (7)	hydrophobic pocket toward entrance of site
B		2 (5)	3 (6)	4 (5)					3 (6)	5 (4)		4 (4)	surface pocket in the dimerization region
F			7 (2)	5 (3)	6 (2)	2 (7)				5 (5)	2 (7)		surface pocket at H3
C1	7 (4)		4 (4)		**2 (7)**	**3 (5)**	**5 (4)**	**5 (3)**	**5 (4)**	**4 (6)**	**6 (3)**	**6 (3)**	between H10 and H12 and may contribute to cofactor binding
C2	**2 (5)**	**2 (4)**	3 (7)		**6 (3)**	**6 (3)**	**3 (7)**	**3 (6)**	4 (5)		**5 (4)**	**6 (2)**	cofactor-binding site between H5 and H12
E1							4 (5)	6 (2)					at the lower ends of helices H3, H7, and H10
E2					4 (6)	2 (6)	5 (3)	2 (7)			3 (6)	3 (6)	putative ligand entrance site between H2′ and the β sheet

[a] For each PPARγ structure, the table shows the number of clusters found at each of the 10 sites, P1−E2. The numbers in parentheses indicate the rank of the corresponding consensus site among all consensus sites for that structure, ranked on the basis of the number of clusters. The bold font emphasize the similarity of mapping results for structures that share certain properties (see the text for details).

and 1k74 were both cocrystallized with L-tyrosine derivatives. Farglitazar in 1fm9 has a benzophenone group, which is replaced by the less bulky vinylogous amide substituent of the agonist GW409544 in structure 1k74. The latter may allow for better interactions between the carboxyl headgroup and the polar residues in site P1, which may explain why this site binds more probe clusters in 1k74 (Table 5). This difference is compensated for by the number of probes in site P3, and the important C1 and C2 sites are identical between the two structures. Structure 1i7i has two chains, but chain b is in an inactive conformation with some segments disordered. The agonists ragaglitazar in structure 1nyx and 3q in structure 1knu differ in their tail groups that bind close to site P4, affecting the number of probes bound to P4. However, the important coactivator binding sites C1 and C2 do not differ between the two structures, as expected for these closely related agonists. More generally, solvent mapping yielded similar results for similar molecules, and this may be regarded as an indication that the algorithm is fundamentally robust; i.e., it exhibits only moderate sensitivity to minor structural perturbations, but at the same, it is sensitive enough to reflect more substantial structural differences.

Binding at the AF-2 Motif: Sites P1 and P2. It is well-known that the binding of agonists introduces localized conformational changes. For example, superimposing the inactive and ragaglitazar-bound (active) structures, Ebdrup et al. (*11*) observed that H323 and H449 were situated in similar positions, while Y473 in the inactive receptor was not in place. They found it even more striking that the C terminus of the inactive receptor was positioned at approximately the same position as the carboxylic acid moiety of the ligand in the active receptor; i.e., the binding pocket was partly sterically blocked by its own C-terminal residue. However, because helix H12 is in the active conformation in chain a of the ligand-free structure 1prg, the above observation is clearly not general. Using the mapping results summarized in Table 5, we are able to distinguish between general and ligand-type specific conformational changes.

As shown in Table 5, site P1 is present only in structures cocrystallized with a strong agonist. In contrast, site P2 is already formed in the ligand-free PPARγ. In structure 4prg, which has the bound partial agonist GW0073, we also find site P2 rather than site P1. As noted above, sites P1 and P2 are very close to each other and share a number of residues, including F282, C285, and S289 on helix H3, Y327 on helix H5, M364 on helix H7, and H449 on helix H10. However, a number of residues are absent from site P2 but become accessible to the solvent probes upon agonist binding, thereby opening up the new binding site P1. These newly accessible residues include Q286, H323, and Y473. Y473 is particularly important, because this residue is located on helix H12, and access to it contributes to the stabilization of H12 in the active conformation. Hydrogen bonds are made between the TZD headgroup (in structures 2prg and 1fm6) or the carboxyl group (in the other structures with strong agonists) and the side chains of residues S289, H323, H449, and Y473. It appears that the strong nonbonded interactions and hydrogen bonds between the TZD or carboxyl group and the protein are necessary for creating a pocket at site P1. Indeed, as shown in Table 5, neither the ligand-free structure 1prg nor the partial agonist-bound structure 4prg has any consensus site in the vicinity of site P1. However, once the hydrophilic, high-affinity pocket at P1 becomes accessible, it siphons away the probes from P2 and the latter site is lost.

Table 5 further indicates that site P1 binds about the same number of probes in each of the structures with strong agonists bound, suggesting that the latter are likely to induce similar conformational changes. Figure 4 and inspection of the structures further suggest that the side chains primarily affected are R288, Q286, F363, and H449. The conformational change is particularly large for side chain R288. In the active-like conformation of the ligand-free structure (chain a of 1prg), R288 protrudes into the binding site, overlapping with the position of any bound ligand. The same side chain is turned outward in the inactive apo structure (chain b of 1prg). Upon binding of a strong agonist, the R288 side chain turns upward such that its extends into site P3,

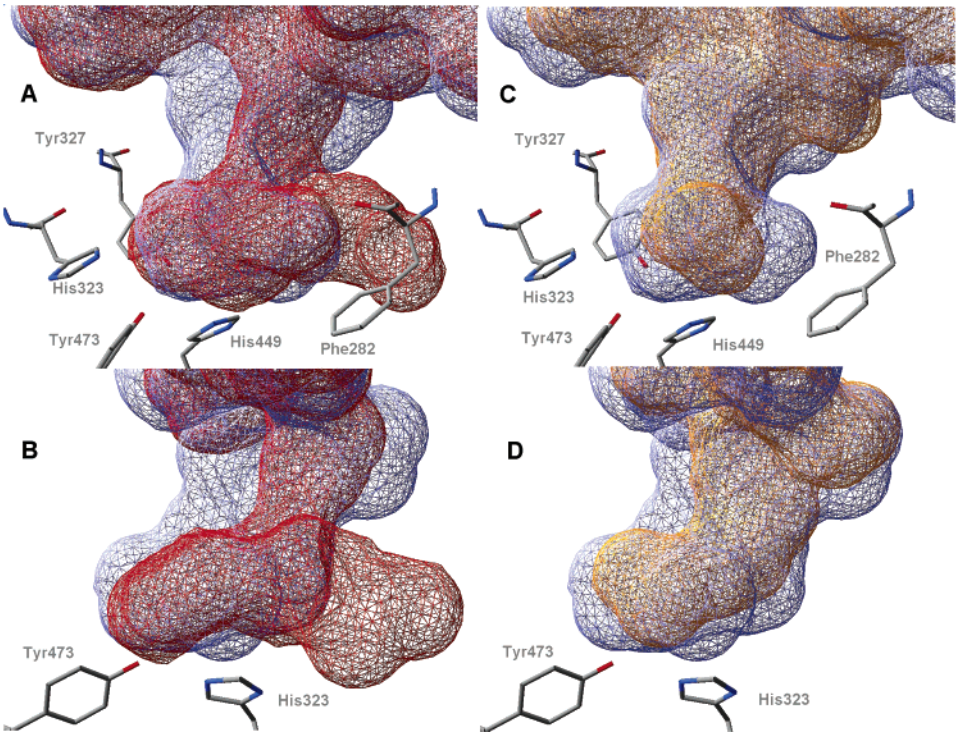


FIGURE 5: Changes in the shape of the P1 and P2 pockets. (A) Comparison of the pockets in the active-like chain of the ligand-free PPARγ (PDB code 1prg for chain a) and in the rosiglitazone-bound structure 2prg. The pockets in chain a of 1prg and 2prg are shown in red and blue, respectively. (B) Same as in A, viewed from the side. (C) Comparison of the pockets in the rosiglitazone-bound structure (PDB code 2prg) and in the active-like conformation of the PPARγ LBD with the bound partial agonist GW0072 (PDB code 4prg for chain a). The pockets in 2prg and chain a of 4prg are shown in blue and yellow, respectively. (D) Same as in C, viewed from the side.

more than 4 Å from its original position (see Table 4). The R288 side chain also has to move to accommodate the partial agonist in 4prg, but the conformational change is substantially smaller. Indeed, in 4prg, the R288 side chain occupies a "lying-down" position at the side of the cavity, forming a hydrogen bond with the partial agonist (see Figure 4B).

Because the position of R288 in the ligand-free structure 1prg (a) would clash with any bound agonist, it is tempting to assume that the large conformational change of this side chain creates the pocket at site P1 and provides access to residues Q286, H323, and Y473, which are not accessible to the probes in the ligand-free structure. However, this is not likely to be the case, because the small probes move into the pocket past R288 even in 1prg (a), still not contacting Q286, H323, and Y473. It appears that pocket P1 is the result of a complex set of cooperative conformational changes in a number of residues, primarily Q286, F363, H449, Y473, H323, S289, and I341, somewhat coordinated by the movement of R288. The small conformational change in S289 is particularly telling, because it is directly controlled by the conformation of the R288 side chain. Q286 moves inward upon ligand binding, forming direct interactions with all ligands, including the partial agonist. Thus, while it contributes to forming a tighter pocket, this conformational change itself again is not sufficient for creating site P1. The F363 side chain also moves toward the ligand, creating a more restricted pocket. However, similar conformational changes occur in both agonist-bound and partial agonist-bound structures. H449 moves out of the way to accommodate the ligand and allows for direct interaction between the ligand and Y473. Although this movement is not very large, the position of the H449 side chain in the agonist-bound 2prg is distinctly different from that in the other two structures. Parts A and B of Figure 5 show (in red) two different views of the molecular surface surrounding the cavity in chain a of 1prg and reveal that in the ligand-free structure site P1 is separated from the rest of the cavity by a narrow "neck". The blue surface in parts A and B of Figure 5 represents the molecular surface around the cavity in the structure with bound rosiglitazone (2prg) and reveals that, upon agonist binding, the collective motion of the nearby side chains substantially widens this "neck" region and site P1 site becomes accessible to the probes. Parts C and D of Figure 5 show the same two views of the cavity in 2prg but this time superimposed with the cavity in chain a of the structure 4prg, which includes the partial agonist GW0072 (yellow surface). As seen in these figures, the binding of GW0072 also widens the "neck" region but the pocket extends much less toward the side chains of H323 and Y473 than in 2prg, in agreement with the finding that the probes do not reach these residues (see Figure 4).

While the above changes are generic, the conformations of several side chains depend on the shape of the agonist. In particular, the inward motion of residue Q286 creates a surface cavity, which we identified as site F (Table 2 and Figure 2); thus, the latter has its origin in the conformational change of a single side chain. However, this change is small for agonists that have a relatively large moiety protruding downward from site P1. This is the case for farglitazar (1fm9), which has a bulky benzophenone group close to its carboxyl head, or for the related agonist GW409544 (1k74) and, to a lesser degree, tesaglitazar in 1i7i and ragaglitazar in 1nyx (Figure 1). In these structures, the Q286 side chain remains closer to the surface of the LBD, and we do not find a pocket at site F (Table 5). Another side chain whose conformation depends on the shape of the agonist is F282,

which defines the lower end of site P1. The change in the position of the F282 ring creates a surface cavity at site E1 but only when the agonist has a long headgroup, i.e., in structures 1fm9 and 1k74 (see Table 5).

Site P3: Partial Agonist Binding and Domain Stabilization. As we described, site P3 is a pocket on the right arm of the Y-shaped-binding site, which is out of reach of the known strong agonists. The only ligand known to get into this pocket is the partial agonist GW0072, which has one of its two benzylamide groups in site P3. According to Table 5, in the apo structure, the P3 site is wide open and available to the probes. On the basis of the number of probe clusters bound, the binding of the partial agonist yields a smaller P3 site, and the pocket is even smaller in structures with strong agonists. Indeed, we have already mentioned that, in all structures with a strong agonist, the R288 side chain moves upward into the P3 site. In this position, it interacts with the side chains of residues E291 and E295 and closes down part of the pocket. A less open pocket at site P3 in all agonist-bound structures agrees with the observations that ligand binding globally stabilizes the LBD domain (*2*, *10*, *13*, *14*). It is generally assumed that the stabilizing effects are concentrated around helix H12. However, our results indicate that the conformational changes extend to other regions of the LBD domain, which is fully supported by recent data from H/D exchange experiments (Patrick Griffin, ExSAR Corp., personal communication). Thus, R288 appears to play an important role in contributing to the increased stability of the ligand-bound PPARγ. However, we note that ligand binding also stabilizes the structure of PPARα (*10*), although that PPAR has a threonine residue (T279) at the corresponding position.

Site P4: Hydrophobic Part of the Ligand-Binding Pocket. Site P4 is close to the lower end of the ligand-binding site, defined by H3 from the left and by the β sheet and helix H2′ from the right. As seen in Table 5, site P4 is accessible to probes in the ligand-free PPARγ structures. The pocket at site P4 is slightly smaller after binding the partial agonist GW00720 (4prg), which places its carboxyl group in this region. P4 is also reduced in structures with short agonists such as rosiglitazone (2prg and 1fm6), in which the distal end does not reach site P4. P4 generally remains large for longer agonists that have their distal end group bound at P4. For example, the methylsulfoxy group of AZ242 (1i7i) occupies site P4. The comparison of agonists ragaglitazar in 1nyx and 3q in 1knu is particularly interesting. Ragaglitazar has a bulky phenoxazine end group and is too short to reach site P4, which is small and binds only two probe clusters (Table 5). In 3q, the phenoxazine group is replaced by a carbazole ring, which is rotated downward into the P4 pocket, resulting in its larger size (Table 5).

Site B: Potential Hot Spot in the Dimerization Region. Site B is located in the back of the LBD, between helices H7 and H10/H11, and close to residue Q444 on the latter. This site is part of the protein surface involved in forming both the PPARγ/RXRα heterodimer and the PPARγ homodimer. Table 5 suggests that the size of site B may depend both on the dimeric partner and on the ligand. On one hand, we find a pocket at this site only in the homodimeric PPARγ structures. On the other hand, this binding site is either missing or weak in ligand-free structures but is present in structures with the partial agonist GW00720, as well as in structures with the L-tyrosine derivative agonists tesaglitazar and ragaglitazar that apart from their distal ends are fairly similar. Nettles et al. (*15*) demonstrated that, because of an allosteric effect, H11 conveys structural information between the ligand and H12, thereby affecting receptor activation. Because the conformation of H11 is affected by dimerization partners, these authors also suggested that H11 may contribute to phenomena such as the "phantom ligand effect", in which a ligand bound to one receptor partner alters the activity of the other partner. Because the presence of site B depends on both ligand and dimerization partner, it may play a role in this communication. Indeed, residue E448 on H11 contributes to site B, whereas H449 is one of the major hydrogen-bonding residues in site P1, suggesting that there is relatively direct communication between the two sites.

Site F is a surface pocket between H3 and H12. As we have discussed, this site is created almost exclusively by the inward motion of the Q286 side chain upon binding of the partial agonist in 4prg or by the binding of agonists that do not have a bulky group protruding downward from the site P1 pocket. The latter condition excludes farglitazar (1fm9), which has a benzophenone moiety, its vinylogous amide analogue GW409544 (1k74), and to a lesser degree, tesaglitazar in 1i7i and ragaglitazar in 1nyx (Figure 1). The significance, if any, of this pocket for coactivator recognition is unknown.

Sites in the Coactivator-Binding Region: C1 and C2. Sites C1 and C2 are around helix H12, close to the putative region of the coactivator/corepressor binding. The mapping results for site C2 are easy to interpret. This site is located between helices H12 and the H4/H5 boundary, with the closest residues on the latter being V315 and K319. The pocket overlaps with the binding site of the SRC-1 coactivator peptide, which is present in several X-ray structures. As shown in Table 5, site C2 is present in the ligand-free and partial-agonist-bound structures but only when H12 is in the active state, i.e., in structures 1prg (a) and 4prg (a). Agonist binding stabilizes site C2, which accommodates about the same number (5 or 6) clusters in all structures with agonists. Site C1 is located on the opposite side of H12, close to the center of a triangle formed by H10, H12, and the loop between the two helices. The phenyl ring of Y473 contributes to this pocket, whereas its hydroxyl group forms a hydrogen bond with the agonists. Although there is no direct evidence that this site is involved in coactivator binding, the coactivators of PPARγ are large proteins and hence are likely to extend beyond the known SRC-1 peptide-binding site, also covering the site C1. The C1 site is very weak or absent in structures in which H12 is in the inactive position, i.e., 1prg (b), 4prg (b), and 1i7i (b). The latter finding is particularly interesting, because it has been observed that chain b of structure 1i7i does not adopt the active conformation despite the presence of the agonist tesaglitazar, possibly because of crystal contacts (*10*). As shown in Table 5, this conformational difference does not affect the mapping results for the C2 site.

The numbers of probe clusters in C1 and C2 appear to provide information on the coactivator-binding specificity of a particular PPARγ structure. Without a ligand, site C2 is very small in both chains of 1prg (Table 5), whereas site C1 discriminates between the active-like and inactive forms. The binding of the partial agonist creates both C1 and C2 in

the active chain a of 4prg, but both sites are of modest size. In addition, none of the sites is present in chain b; thus, the binding of the partial agonist GW0072 is unable to stabilize the coactivator recognition pockets. Apart from the already discussed chain b in 1i7i, all agonist-bound structures exhibit stable C1- and C2-binding sites, which therefore appear to be prerequisites for coactivator binding. The relative strengths of the two pockets seem to depend on the properties of the agonist. In both rosiglitazone-bound structures 2prg and 1fm6, we see a modest C1 and very strong C2, the latter with 6 clusters (Table 5). The mapping results for the two structures are similar, although 2prg is a homodimer and 1fm6 is a PPARγ/RXRα heterodimer. The binding of the L-tyrosine derivatives in 1fm9 and 1k74 creates a relatively large C1 site (5 clusters) and a smaller C2 site (3 clusters). It is notable that the results are identical, despite coming from the mapping of two different structures. Sites C1 and C2 are both relatively large in chain a of 1i7i, with 5 and 4 probe clusters, respectively, and even larger in the structures with ragaglitazar (1nyx) and its 3q analogue (1knu). It is reassuring that, with the already mentioned exception of chain b in 1i7i, while similar structures yield similar mapping results, the sizes of C1 and C2 (i.e., the numbers of clusters at these sites) vary between different classes of agonists.

Putative Entrance Sites: E1 and E2. Sites E1 and E2 are surface pockets at the openings on the two ends of the PPARγ LBD-binding site. Site E2 is located between helix H2′ and the β sheet. This position is frequently mentioned as the putative entrance to the LBD-binding site (*10*). As shown in Table 5, E2 is present only in the agonist-bound structures, with the exception of 1i7i (a) and 1i7i (b). Inspection of these structures shows that the methylsulfoxy group on the distal end of the agonist tesaglitazar (AZ242) in these structures reaches further down into pocket P4 than in the other agonist-bound structures. The relatively large P4 site in 1i7i (see Table 5) draws the probes from the entrance further into the binding site. In view of the lack of any alternative site at the distal end of the binding site, our results support the hypothesis that the ligands are likely to enter through the pocket at site E2. Indeed, the loop connecting H2′ and H3 is very flexible. For example, the positions of residues 264 and 266 in this region can differ more the 10 Å among different structures (results not shown), suggesting that the loop does not prevent the entrance of large ligands into the binding site. The conformational change is the result of agonist binding, because the rmsd values for these residues are much larger between apo and agonist-bound structures, as well as between the partial agonist-bound structure and any agonist-bound structure, than between any two agonist-bound structures. Inspection of structures suggests that the loop connecting H2′ to H3 thrusts forward more in the agonist-bound structures than in the ligand-free one or with the partial agonist, partly resulting in the E2 pocket. Thus, while agonist binding results in the stabilization and increased compactness of some structural elements (e.g., at the AF-2 motif and the P3 site), this does not apply to this region.

Site E1 is located between the lower ends of helices H3 and H7. Originally, we thought that this could also be a putative ligand entrance, directed to the other (i.e., H12-contacting) end of the ligand. However, further inspection revealed that site E1 is primarily created by the movement of the F282 side chain and that it exists only in structures with the L-tyrosine-based antagonists farglitazar (1fm9) and GW409544 (1k74). These agonists contain either a benzophenone or a vinylogous amide as the L-tyrosine N substituent, which goes downward from the site P1 pocket. A line along this cavity connects to site E1, but because E1 is not seen in any other structures, the site is most likely produced specifically by L-tyrosine agonists, primarily through their effects on the orientation of the F182 side chain.

Conclusions. The clustering of probes in the uniform solvent mapping of 12 PPARγ structures identifies 10 sites as "hot spots" for the recognition of interaction partners by the LBD. A total of 4 of these sites (P1−P4) are in the ligand-binding site, 2 (C1 and C2) in the coactivator-binding area, 1 in the dimerization region (B), 2 (E1 and E2) in the lower half of LBD, where E2 is at the putative ligand-entrance site, and an additional site is on the surface (F). These sites were further explored by targeted mapping to determine, with higher accuracy, the number of clusters that remain in each pocket after the probes have been moved around by a search. We find that large fractions of the PPARγ-binding site are already formed in the ligand-free structure, but the important site P1 near the AF-2 motif is accessible only in structures that are cocrystallized with strong agonists. This suggests that PPARγ can bind a large variety of ligands, but only a selected class of molecules with the appropriate polar headgroup can achieve the required conformational changes at the AF-2 motif and thus serve as strong agonists. Our results also show that binding of such agonists moves one side chain (R288) more than 4 Å to the distal P3 site that is wide-open in the ligand-free structure. The newly formed salt bridges restrict the size of the pocket at the P3 site and contribute to the stabilization of the LBD. It is interesting that the weak partial agonist GW0072, which binds in a very different mode from the strong agonists, places one of its two benzylamide groups in P3.

We emphasize that the analysis described in this paper was made possible by the availability of 12 X-ray structures of the PPARγ LBD, 2 without a ligand, 2 with a partial agonist, and 8 with 6 different agonists bound. As we described, conformational solvent mapping identified 10 binding sites and showed how the size of these pockets differ among the different structures. Restricting considerations only to ligand-free conformations would have provided much less information. Indeed, in the apo structures (chains a and b of 1prg), the important site P1 is not present and the "entrance sites" E1 and E2 as well as site F are also missing, i.e., clearly created in the process of ligand binding. The binding of strong agonists introduces conformational changes at a number of further sites, including C1 and C2 in the region of coactivator binding. We consider it very important that the numbers of probe clusters at C1 and C2 appear to provide information on the coactivator-binding specificity of a particular PPARγ structure. Apart from a structure with missing fragments (chain b of 1i71), all agonist-bound structures exhibit stable pockets at both C1 and C2 sites. However, the number of probe clusters, retained at these sites, depends on the properties of the agonists, suggesting that mapping may help to understand the effect on the ligand on coactivator binding and transactivation profiles.

Binding Sites Recognized by Solvent Mapping in General. Solvent mapping by X-ray crystallography has been applied

to a limited number of enzymes (*22*−*26*). Results show that the probes cluster in the active site, forming a "consensus" site that delineates the binding pocket. All other binding sites are either in crystal contact, occur only at high ligand concentrations, or are in small, buried pockets that bind only a subset of the solvent molecules rather than all of them. The preferential binding of organic molecules to the active site of the lysozyme has also been shown in aqueous solution by NMR methods (*42*). Using computational methods, we have mapped some 50 enzymes (*19*−*21*) and a few non-enzyme proteins. These studies helped to elucidate three properties of binding sites that attract organic molecules and hence are found by mapping. First, the probes cluster in fairly large pockets that can provide a substantial number of ligand−protein interactions. Indeed, the intermolecular van der Waals energy is generally the largest contribution to the calculated binding free energy (*24*). Second, the pockets are partially nonpolar, with the hydrophobic interactions also contributing to the binding free energy. Third, the presence of several polar patches in the binding site is very important, because it enables the probes to bind in a number of rotational/translational states. In each conformation, the polar parts of the probe form favorable electrostatic interactions or possibly a hydrogen bond with one of the polar groups of the protein (*19*). Because of the multiplicity of the bound conformations, probes binding in relatively large pockets retain more of their rotational/translational entropy than the ones that bind elsewhere in single conformational states, and the resulting difference in the free energy makes tight binding in small crevices less favorable.

Thus, the solvent-mapping methods used in the present study generally identify binding sites that are relatively large and include both nonpolar and polar patches. These conditions are satisfied by the active site of most enzymes and also by several subsites of the large PPARγ ligand-binding pocket. In previous applications, we also identified ligand entrance sites (*21*) similar to site E2 of the PPARγ. However, it was not at all obvious that these methods would identify interaction sites in the AF-2 and dimerization regions of a receptor protein such as PPARγ. In fact, earlier calculations showed that mapping may fail to recognize protein−protein interactions sites, which are substantially more planar than the sites that bind relatively small ligands. Thus, the present identification of sites C1 and C2 in the coactivator-binding region of PPARγ was somewhat unexpected. The fact that C1 and C2 are well-defined consensus sites suggests that the coactivator may partially unfold before binding to PPARγ. This assumption is supported by the fact that coactivator recognition is dominated by a single LXXLL peptide motif, although the selection of the coactivator must be modulated by other parts of the interface (*16*, *17*). More generally, the results of this paper suggest that computational solvent mapping is a powerful tool to study the binding sites of nuclear receptor ligand-binding domains. It will be particularly interesting to apply the method to orphan nuclear receptors, for which the binding site is less well-understood.

ACKNOWLEDGMENT

We thank Karl Clodfelter, Spencer Thiel, and Michael Silberstein for checking some of the results and for discussions and Dr. Scott Mohr for reading and improving the manuscript.

SUPPORTING INFORMATION AVAILABLE

Mapping results for the 12 structures studied in the paper are given as PDB files. Each file contains the coordinates of the protein and the coordinates of five conformations for each of the seven different ligands used in our study (35 small molecules for each protein). The ligand conformations are the centers of the five lowest free-energy clusters for each ligand. In addition to heavy atoms, the PDB files also include the hydrogen coordinates. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Berger, J., and Moller, D. E. (2002) The mechanism of action of PPARs, *Annu. Rev. Med. 53*, 409−435.
2. Nagy, L., and Schwabe, J. W. R. (2004) Mechanism of the nuclear receptor molecular switch, *Trends Biochem. Sci. 29*, 317−324.
3. Li, Y., Lambert, M. H., and Xu, H. E. (2003) Activation of nuclear receptors: A perspective from structural genomics, *Structure 11*, 741−746.
4. Willson, T. M., Brown, P. J., Sternbach, D. D., and Henke, B. R. (2000) PPARs: From orphan receptors to drug discovery, *J. Med. Chem. 43*, 527−550.
5. 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. (1998) Ligand binding and co-activator assembly of the peroxisome proliferators-activated receptor-γ, *Nature 395*, 137−144.
6. Uppenberg, J., Svensson, C., Jaki, M., Bertilsson, G., Jendeberg, L., and Berkenstam, A. (1998) Crystal structure of the ligand binding domain of the human nuclear receptor PPARγ, *J. Biol. Chem. 273*, 31108−31112.
7. Oberfield, J. L., Collins, J. L., Holmes, C. P., Goreham, D. M., Cooper, J. P., Cobb, J. E., Lenhard, J. M., Hull-Ryde, E. A., Mohr, C. P., Blanchard, S. G., Parks, D. J., Moore, L. B., Lehmann, J. M., Plunket, K., Miller, A. B., Milburn, M. V., Kliewer, S. A., and Willson, T. M. (1999) A peroxisome proliferators-activated receptor γ inhibits adipocyte differentiation, *Proc. Natl. Acad. Sci. U.S.A. 96*, 6102−6106.
8. Gampe, R. T., Montana, V. G., Lambert, M. H., Miller, A. B., Bledsoe, R. K., Milburn, M. V., Kliewer, S. A., Willson, T. M., and Xu, H. E. (2000) Asymmetry in the PPARγ/RXRα crystal structure reveals the molecular basis of heterodimerization among nuclear receptors, *Mol. Cell 5*, 545−555.
9. 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. (2001) Determinants of ligand binding selectivity between the peroxisome proliferator-activated receptors, *Proc. Natl. Acad. Sci. U.S.A. 98*, 13919−13924.
10. Cronet, P., Petersen, J. F., Folmer, R., Blomberg, N., Sjoblom, K., Karlsson, U., Lindstedt, E. L., and Bamberg, K. (2001) Structure of the PPARα and -γ ligand binding domain in complex with AZ242; Ligand selectivity and agonist activation in the PPAR family, *Structure 9*, 699−706.
11. Ebdrup, S., Pettersson, I., Rasmussen, H. B., Deussen, H. J., Frost Jensen, A., Mortensen, S. B., Fleckner, J., Pridal, L., Nygaard, L., and Sauerberg, P. (2003) Synthesis and biological and structural characterization of the dual-acting peroxisome proliferator-activated receptor α/γ agonist ragaglitazar, *J. Med. Chem. 46*, 1306−1317.
12. Sauerberg, P., Pettersson, I., Jeppesen, L., Bury, P. S., Mogensen, J. P., Wassermann, K., Brand, C. L., Sturis, J., Woldike, H. F., Fleckner, J., Andersen, A. S., Mortensen, S. B., Svensson, L. A., Rasmussen, H. B., Lehmann, S. V., Polivka, Z., Sindelar, K., Panajotova, V., Ynddal, L., and Wulff, E. M. (2002) Novel tricyclic-α-alkyloxyphenylpropionic acids: Dual PPARα/γ agonists with hypolipidemic and antidiabetic activity, *J. Med. Chem. 145*, 789−804.
13. Kallenberger, B. C., Love, J. D., Chatterjee, V. K., and Schwabe, J. W. (2003) A dynamic mechanism of nuclear receptor activation and its perturbation in a human disease, *Nat. Struct. Biol. 10*, 136−140.

14. Pissios, P., Tzameli, I., Kushner, P., and Moore, D. D. (2000) Dynamic stabilization of nuclear receptor ligand binding domains by hormone or corepressor binding, *Mol. Cell 6*, 245−253.
15. Nettles, K. W., Sun, J., Radek, J. T., Sheng, S., Rodriguez, A. L., Katzenellenbogen, J. A., Katzenellenbogen, B. S., and Greene, G. L. (2004) Allosteric control of ligand selectivity between estrogen receptors α and β: Implications for other nuclear receptors, *Mol. Cell 13*, 317−327.
16. Wu, Y., Chin, W. W., Wang, Y., and Burris, T. P. (2003) Ligand and coactivator identity determines the requirement of the charge clamp for coactivation of the peroxisome proliferator-activated receptor γ, *J. Biol. Chem. 278*, 8637−8644.
17. Kodera, Y., Takeyama, K., Murayama, A., Suzawa, M., Masuhiro, Y., and Kato, S. (2000) Ligand type-specific interactions of peroxisome proliferator-activated receptor γ with transcriptional coactivators, *J. Biol. Chem. 275*, 33201−3324.
18. Stanley, T. B., Leesnitzer, L. M., Montana, V. G., Galardi, C. M., Lambert, M. H., Holt, J. A., Xu, H. E., Moore, L. B., Blanchard, S. G., and Stimmel, J. B. (2003) Subtype specific effects of peroxisome proliferator-activated receptor ligands on corepressor affinity, *Biochemistry 42*, 9278−9287.
19. Dennis, S., Kortvelyesi, T., and Vajda, S. (2002) Computational mapping identifies the binding sites of organic solvents on proteins, *Proc. Natl. Acad. Sci. U.S.A. 99*, 4290−4295.
20. Kortvelyesi, T., Dennis, S., Silberstein, M., Brown, L., III, and Vajda, S. (2003) Algorithms for computational solvent mapping of proteins, *Proteins 51*, 340−351.
21. Silberstein, M., Dennis, S., Brown, L., III, Kortvelyesi, T., Clodfelter, K., and Vajda, S. (2003) Identification of substrate binding sites in enzymes by computational solvent mapping, *J. Mol. Biol. 332*, 1095−1113.
22. Mattos, C., and Ringe, D. (1996) Locating and characterizing binding sites on proteins, *Nat. Biotechnol. 14*, 595−599.
23. Ringe, D., and Mattos, C. (1999) Analysis of the binding surfaces of proteins, *Med. Res. Rev. 19*, 321−331.
24. Allen, K. N., Bellamacina, C. R., Ding, X., Jeffery, C. J., Mattos, C., Petsko, G. A., and Ringe, D. (1996) An experimental approach to mapping the binding surfaces of crystalline proteins, *J. Phys. Chem. 100*, 2605−2611.
25. English, A. C., Done, S. H., Caves, L. S., Groom, C. R., and Hubbard, R. E. (1999) Locating interaction sites on proteins: The crystal structure of thermolysin soaked in 2% to 100% isopropanol, *Proteins 37*, 628−640.
26. English, A. C., Groom, C. R., and Hubbard, R. E. (2001) Experimental and computational mapping of the binding surface of a crystalline protein, *Protein Eng. 14*, 47−59.
27. Laskowski, R. A., Luscombe, N. M., Swindells, M. H., and Thornton, J. M. (1996) Protein clefts in molecular recognition and function, *Protein Sci. 5*, 2438−2452.
28. Hendlich, M., Rippmann, F., and Barnickel, G. (1997) LIGSITE: Automatic and efficient detection of potential small molecule-binding sites in enzymes, *J. Mol. Graphics Modell. 15*, 359−363.
29. Brady, G. P., Jr., and Stouten, P. F. W. (2000) Fast prediction and visualization of protein binding pockets with PASS, *J. Comput.-Aided Mol. Des. 14*, 383−401.
30. Gilson, M. K., and Honig, B. (1988) Calculation of the total electrostatic energy of a macromolecular system: Solvation energies, binding energies, and conformational analysis, *Proteins 4*, 7−18.
31. Honig, B., and Nicholls, A. (1995) Classical electrostatics in biology and chemistry, *Science 268*, 1144−1149.
32. Bruccoleri, R. E. (1993) Grid positioning independence and the reduction of self−Energy in the solution of the Poisson−Boltzmann equation, *J. Comput. Chem. 14*, 1417−1422.
33. *QUANTA/CHARMM Program*, Molecular Simulations Inc., Waltham, MA.
34. Zhang, C., Vasmatzis, G., Cornette, J. L., and DeLisi, C. (1996) Determination of atomic desolvation energies from the structures of crystallized proteins, *J. Mol. Biol. 267*, 707−726.
35. Miyazava, S., and Jernigan, R. (1985) Estimation of effective interresidue contact energies from protein crystal structures: Quasi-chemical approximation, *Macromolecules 18*, 534−552.
36. Schaefer, M., and Karplus, M. A. (1996) A comprehensive analytical treatment of continuum electrostatics, *J. Phys. Chem. 100*, 1578−1599.
37. Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., and Karplus, M. (1983) CHARMM: A program for macromolecular energy, minimization, and dynamics calculations, *J. Comput. Chem. 4*, 187−217.
38. MacKerell, A. D., Jr., Brooks, B., Brooks, C. L., III, Nilsson, L., Roux, B., Won, Y., and Karplus, M. (1998) CHARMM: The energy function and its parameterization with an overview of the program, in *The Encyclopedia of Computational Chemistry* (Schleyer, P. v. R. et al., Eds.) Vol. 1, pp 271−277, John Wiley and Sons, Chichester, U.K.
39. Wallace, A. C., Laskowski, R. A., and Thornton, J. M. (1995) LIGPLOT: A program to generate schematic diagrams of protein−ligand interactions, *Protein Eng. 8*, 127−134.
40. McDonald, I. K., and Thornton, J. M. (1994) Satisfying hydrogen bonding potential in proteins, *J. Mol. Biol. 238*, 777−793.
41. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000) The Protein Data Bank, *Nucleic Acids Res. 28*, 235−242.
42. Liepinsh, E., and Otting, G. (1997) Organic solvents identify specific ligand binding sites on protein surfaces, *Nat. Biotechnol. 15*, 264−268.
43. DeLano, W. L. (2002) *The PyMol User's Manual*, DeLano Scientific, San Carlos, CA.

BI048032C