Eur Biophys J (2008) 37:369–379
DOI 10.1007/s00249-007-0220-9

ORIGINAL PAPER

Ligand-escape pathways from the ligand-binding domain of PPARγ receptor as probed by molecular dynamics simulations

D. Genest · N. Garnier · A. Arrault · C. Marot · L. Morin-Allory · M. Genest

Received: 25 May 2007 / Revised: 30 August 2007 / Accepted: 14 September 2007 / Published online: 11 October 2007

Abstract Conformational rearrangements of peroxysome proliferator activated receptor (PPARγ) ligand-binding domain (LBD) that accompany the release and binding of ligands are not well understood. To determine the major events associated with the escape of the partial agonist GW0072, molecular dynamic (MD) simulations were performed using two different methods: reversed targeted molecular dynamics (TMD^{-1}) and time-dependent distance restraints (TDR) using as restraints either the root mean square deviation from a reference structure (TMD^{-1}) or the distance between the geometrical centers of the binding pocket and of the ligand (TDR). Both methods do not assume any a priori route for ligand extraction. To avoid artifacts, different initial simulation conditions were used and particular attention was paid for giving time to the protein to relax during the extraction process by running 10–12 ns simulations within explicit water. Two distinct exit gates A and B were found, independently of initial conditions and method. During the exit process no interaction between GW0072 and the transactivation AF-2 helix was observed. Our results suggest that the ligand uses the intrinsic flexibility of the protein to move within the receptor. Paths A and B are very similar to those found for other nuclear receptors, suggesting that these routes are a common characteristics of nuclear receptors that are used by different kinds of ligands. Finally, the knowledge of entry/exit pathways of a receptor should be very useful in discriminating between different ligands that could have been favorably docked in the binding pocket by introducing docking along these pathways into computational drug design protocols.

Keywords PPAR · Ligand-binding domain · MD simulation · Unbinding pathway · GW0072

D. Genest (✉) · N. Garnier · M. Genest
Centre de Biophysique Moléculaire (affiliated to the University of Orléans and to INSERM), CNRS, Rue Charles Sadron, 45071 Orléans cedex 2, France
e-mail: genest@cnrs-orleans.fr

A. Arrault · C. Marot · L. Morin-Allory
Institut de Chimie Organique et Analytique, CNRS-Université d'Orléans, BP 6759, Rue de Chartres, 45067 Orléans cedex 2, France

Introduction

The knowledge of detailed pathways for small organic compounds to enter into a protein receptor and reach a binding site or to escape from this site should bring interesting information for rational drug design. Indeed the docking of a particular ligand into a binding pocket or a cavity could be energetically favorable while the existence of a high energetic barrier along the entering pathway could prevent its progression toward the pocket. Therefore comparing the interaction energy of different ligands along the known entering pathway could help in discriminating between equally well-docked molecules. Up to now this point is not routinely taken into account in structure-based drug development. Experimental search for exit/entry pathways is generally difficult to achieve, but computational methods such as molecular dynamic (MD) simulations using appropriate restraints are well suited for this task. The most widely used technique for this purpose is steered MDs (Isralewitz et al. 2001) (and references therein) in which a force, dependent on time, is applied to the ligand to force its displacement along a reaction coordinate. Unfortunately the pathway of the reaction coordinate has to be known or hypothesized, thus creating a bias. To remove the bias,

random expulsion MD was developed (Ludemann et al. 2000) in which the direction of an additional force with constant magnitude is randomly changed during the time course of the simulation. In order to save computational time allowing many runs, each run is limited to a few hundreds of picoseconds, by imposing strong restraints so that the system is far from equilibrium during the simulation. This methodology is still in development as the resulting pathways may depend on parameters used in the simulation (Carlsson et al. 2006) such as the atom(s) to which the artificial force is added, the tolerance for energy change at each step, the magnitude of the force, etc... Locally enhanced sampling MDs (Elber and Karplus 1990) is another technique that has been used by Blondel et al. (Blondel et al. 1999) for the study of the retinoic acid receptor and by Martinez et al. (2005) for exploring unbinding pathways of several ligands of thyroid hormone receptors. These studies were performed in the gas phase approximation in order to run a sufficiently large number of simulations and multiple pathways were found.

The aim of the present study is to evidence, at the atomic level, pathways for a ligand to move into a protein receptor, using MD techniques as close as possible to equilibrium, with the idea of limiting the bias induced by the restraints. This imposes the use of sufficiently slow variation of the restraints during the simulation of the ligand displacement and a realistic description of the environment with explicit water molecules and counter ions.

Molecular dynamic simulations of large systems are time consuming thus preventing a complete exploration of possible pathways. Nevertheless, it is expected that the most probable paths can be evidenced. Here, two types of restraints acting on selected atoms of the ligand and of the binding pocket are used without imposing a predetermined direction for the ligand migration. These restraints consist in adding a harmonic potential to the physical force field that progressively increases a specific quantity. The targeted molecular dynamic (TMD) method (Schlitter et al. 1994) uses as restraints the root mean square deviation (RMSD) from the initial configuration of the complex. The other method, named time-dependence distance restraint method (TDR) uses the distance between the geometrical centers of the atoms of the binding pocket and of the atoms of the ligand.

These two approaches have been applied to the dissociation of the partial agonist GW0072 (Fig. 1 top) from the ligand-binding domain (LBD) of the PPARγ receptor, a member of the super family of nuclear receptors (Issemann and Green 1990). Peroxysome proliferator-activated receptors (PPARs) play a central role in the regulation of genes involved in the biosynthesis, the storage and catabolism of dietary fats. They are found as three homologous isoforms: alpha, delta and gamma (Willson et al. 2000). The thiazolidinediones (TZDs), which improve insulin sensitivity bind to the γ sub-type (Inzucchi et al. 1998) and are clinically used in the treatment of type II diabetes and obesity diseases. Consequently PPARγ has become a target receptor for the discovery of new drugs. Ligand effects on PPARγ are mediated through the LBD, which contains 270 amino acid residues. The main region of LBD is the transactivation region AF-2 located on helix H12 (Fig. 1 bottom). Several ligands have been co-crystallized in PPARγ with different activities: agonists such as rosiglitazone (Nolte et al. 1998) and partial agonists such as GW0072 (Oberfield et al. 1999). While an agonist induces an active conformation of the receptor (Tyr473, localized on AF-2, points towards the centre of the active site), the partial agonist GW0072 leads to a mixture of active and inactive conformer (Tyr473 points its hydroxyl group out of the pocket). Recent studies highlighted the huge proportion of binding "hot spots" within the protein. More precisely, they identified four "hot spots" in the ligand-binding pocket, two in the co-activator-binding region, one in the dimerization domain, two around the supposed ligand entrance site, and one minor site without a known function (Sheu et al. 2005). It was shown that conformational changes can occur in most of these sites and in some cases the ligand can stabilize the LBD and the co-activator binding region.

The first step of this study was the choice of the 3D structure of the receptor. Numerous X-ray structures of the LBD of PPARγ are available in the Protein Data Bank (Berman et al. 2000) (http://www.rcsb.org/pdb/home/home.do) either co-crystallized with a ligand or depleted from ligand (Cronet et al. 2001; Kuhn et al. 2006; Li et al. 2005; Lu et al. 2006; Mahindroo et al. 2005; Ostberg et al. 2004; Uppenberg et al. 1998), and some of them correspond to the active form of the receptor while others correspond to the inactive form. However, in many X-ray structures all residues are not resolved so that they cannot be used for MD simulations. Among the fully solved structures of the LBD of PPARγ, we retained the PDB entry # 4PRG, which contains four monomers. We considered the first one (chain A) which exhibits the active form structure (Oberfield et al. 1999).

Eight 10–12 ns simulations under different conditions were performed and were complemented with sampling of particular intermediate conformational states along the exit pathways for energetic analysis. A total of about 130 ns of simulation were carried out for this study.

Materials and methods

The RMSD between two structures of a set of N atoms is defined by

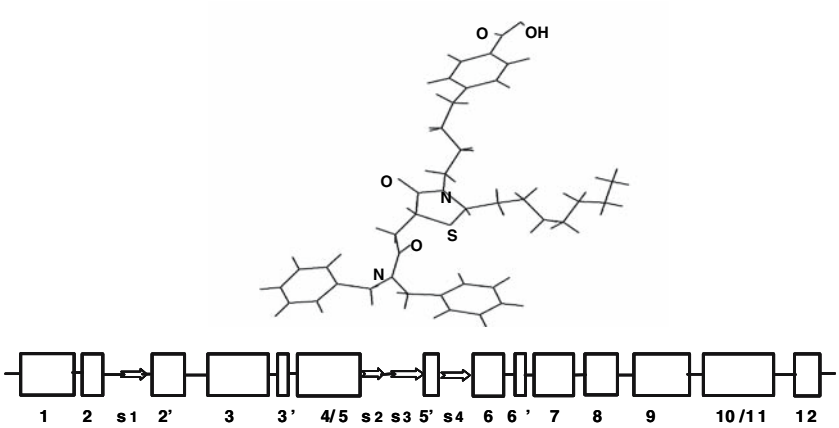

Fig. 1 *Top* Ligand GW0072. *Bottom* X-ray secondary structure of the LBD of PPARγ showing the label of helices (*rectangles*) and of the β-strands (*arrows*)

$$\text{RMSD} = \{1/N \sum (\mathbf{r}_{i1} - \mathbf{r}_{i2})^2\}^{1/2}$$

where $\mathbf{r}_{i1}$ and $\mathbf{r}_{i2}$ are the coordinate vectors of atom i in structures 1 and 2, respectively, the sum being performed from atom $i = 1$ to atom $i = N$.

All simulations were performed in the presence of explicit water and counterions with PMEMD 4.0 (Particle Mesh Ewald Molecular Dynamics) (Duke and Pedersen 2003), an improved version of the SANDER program from the AMBER package (Case et al. 2002), in which we have implemented the TMD algorithm. The ff99 and gaff force fields were applied to the atoms of the protein and of the ligand, respectively. Atomic charges on GW0072 were obtained with the ANTECHAMBER program of the AMBER 9 distribution with the bond charge correction (BCC) method. The length of covalent bonds involving a proton was maintained constant with the SHAKE algorithm (Ryckaert et al. 1977) and the time step for integrating the Newton equations was 2 fs.

Targeted molecular dynamics (TMD)

Considering a set of atoms, the TMD method (Aci et al. 2004, 2005; Schlitter et al. 1994) constrains the RMSD between the current structure and a reference structure to a value ρ_0, which is slowly varied from an initial value to a targeted value during a simulation. If the reference structure is a targeted structure (direct TMD), ρ_0 is decreased from the RMSD between the initial and targeted (final) structures to a value close to 0. If the reference structure is the initial structure, the RMSD is increased from 0 to a predefined value. In that case, the method referred to reverse targeted molecular dynamics (TMD^{-1}) is much less constraining than the direct TMD for searching how a structure is destabilized, since no indication of the search direction is given. In AMBER (and in our home version of PMEDM) the TMD^{-1} technique is introduced as a harmonic restraining potential, and atomic coordinates are mass weighted in the calculation of RMSD.

Model, thermalization and free MD simulation of the PPARγ-GW0072 complex

The crystallographic atomic coordinates of the complex were taken from Oberfield et al. (1999) deposited at the PDB (Berman et al. 2000) (code 4PRG chain A). Since the carboxyl groups of Glu259 and of GW0072 face each other in 4PRG, they are likely not simultaneously charged. As the p*K*a of Glu in solution is 4.3 and that of GW0072 is probably on the order of 4–5 (4.2 for benzoic acid) it is not straightforward to determine which one is deprotonated in the protein. Studies in molecular mechanics usually assume Glu residues to be deprotonated in proteins, and we adopted this state for all Glu residues including Glu259 so that GW0072 was assumed to be in protonated state.

Missing H atoms and 5 Na^+ counter-ions for neutralizing the system were added and the complex was inserted in a rectangular box containing 10,317 water molecules. The whole system was energy minimized with 200 steps of steepest descent and 1,800 steps of conjugate gradient methods for removing possible bad contacts. This was followed with a 40 ps MD heating period from 0 to 300 K in the NVT ensemble during which the position of protein atoms was restrained. Ions were restrained only during the first 20 ps. Then a 60-ps MD equilibration period was performed with progressive removing of restraints on the protein during the first 40 ps. Equilibration was continued for 400 ps in the NPT ensemble without restraining atomic positions except those of the C and N atoms of the peptide bond linking Thr328 and Met329. This bond was roughly parallel to the longest dimension of the box so that this protocol limits the overall motions of the complex within the box.

Finally, 11.5 ns of free MD simulation (MD0) of the thermalized X-ray structure (Oberfield et al. 1999) were performed at 300 K, during which atomic coordinates were stored at every picosecond.

Simulation of ligand extraction

In order to investigate a possible influence of initial conditions and methodology on the search for ligand extraction pathways, eight restrained MD runs were performed under different conditions as follows (see Table 1 for details):

Initial structures

Three conformations were selected at time 0, 5.5 and 11.5 ns, respectively, of the production period of MD0 and used as initial structures for restrained MD simulations. They will be named I1, I2 and I3, respectively, in the following. Each starting structure was simulated at least twice, using a different restraining method.

Restraints

Two methods are used for removing the ligand from its binding site: one is the time-dependent distance-restraint method (TDR), the other is TMD^{-1}, a time-dependent RMSD-restraint method. In the TDR method the distance between the geometrical centers of the ligand and of the atoms belonging to the binding site is continuously increased from its initial value to about 25 Å at a rate of 0.002 Å/ps. A flat 1 Å wide restraint potential flanked with two half-harmonic walls (force constant 10 kcal/Å^2) is used in order to allow the distance to temporarily decrease if necessary (MD1, MD2, MD3 and MD8).

In TMD^{-1} a unique set of atoms belonging to the ligand and to the binding site is considered. The RMSD (ρ_0) between the initial conformation and the current conformation of this set of atoms is constrained to increase progressively from 0 Å to approximately 18 Å by varying ρ_0 by steps of 0.09 Å every 50 ps (MD4, MD5, MD6 and MD7). The remainder of the protein is not restrained.

The length of the simulations (or equally the rate of varying the restraints) is guided by a balance between a reasonable computational time and the necessity to give time to the protein to be reactive to the time-dependent restraints. At the end of the simulations the ligand is outside the protein.

In six simulations (MD1–MD6) all heavy atoms of the ligand and all backbone-heavy atoms (N, Cα, C and O) of the binding site are taken into account in calculating the RMSD or the distance between geometrical centers. In MD7 and MD8, all heavy atoms of the ligand and all heavy atoms of the binding-site amino acid (backbone and side chains) are used in the restraints.

Table 1 Overview of the conditions of simulations for the nine MD simulations and exit routes. Definition of the binding site atoms are given as footnotes

MD simulation	Initial structure	Restraining method	Initial distance[d] (Å)	Simulation length (ns)	Exit route
MD0	I1	–	–	11.5	–
MD1[a]	I1	TDR	0.8	12	A1
MD2[a]	I2	TDR	1.2	12	A1
MD3[a]	I3	TDR	1.	12	A1
MD4[a]	I1	TMD^{-1}	–	10	B1
MD5[a]	I2	TMD^{-1}	–	10	A1
MD6[a]	I3	TMD^{-1}	–	10	A2
MD7[b]	I1	TMD^{-1}	–	12	A1
MD8[c]	I2	TDR	0.4	10	B2

[a] N, Cα, C and O atoms of the peptide chain of the nine residues Glu259, Arg280, Ile281, Arg288 , Ile326, Tyr327, Ile341, Phe363 and Met364

[b] All heavy atoms of the same residues as in (a)

[c] All heavy atoms of residues Leu255, Glu259, Arg280, Gly284, Cys 285, Ile326, Tyr327, Met329 and Leu333

[d] The initial distance corresponds to the distance between the geometrical center of the ligand and the geometrical center of the binding site calculated in the initial structures I1, I2, or I3 (see text for definition). TDR is for time-dependant distance restraints and TMD^{-1} is for targeted molecular dynamics (see text for details). The last column gives the exit routes found in this study

Residues defining the binding site

In seven simulations (see Table 1) the residue set that defines the binding site is composed of nine amino acids having at least one atom separated by less than 3.5 Å from any atom of the ligand, at once in I1, I2, I3 and the X-ray structure. All the selected residues are also found to belong to the binding site by the online program CASTp (Liang et al. 1998) (http://cast.engr.uic.edu/cast). In MD8, nine residues were randomly picked up from the set of amino acids given by CASTp for defining the pocket, and used in the restraints.

Energetics

Additional 500 ps MD runs were performed for each restrained MD, starting from snapshots selected every 2 ns of the trajectories, by keeping constant either the distance between the geometrical centers (MD1-3 and MD8) or the RMSD of the atom set described above (MD4-7). Atomic coordinates were stored every 2 ps. Using these trajectories

the average intermolecular energy between the ligand and the protein was calculated with the MM_PBSA program of AMBER.

Molecular display

All molecular graphic representations were displayed with the Visual Molecular Dynamics VMD software (Humphrey et al. 1996).

Results

Description of the X-ray structure

The PDB file in 4PRG contains four monomers named chain A, B, C and D, respectively. We describe here the conformational characteristics of the first one, which was used as starting structure for the present study. The authors (Oberfield et al. 1999) report the presence of 14 helices and of a four-strand β sheet. To be in agreement with usual nomenclature (Nolte et al. 1998), helices are labeled as shown on Fig. 1 bottom. However, three helices (H3′, H5′ and H6′) are very short, implying only three residues each, and are not detected as α helices with the MolMol program (Koradi et al. 1996) but rather as 3_{10} helices. It is worth noting that H5′ is mentioned only in seven structures by the authors among a set of 35 structures extracted from the PDB; H4′ is mentioned in 24 of them, and H6′ in 26 of them. Thus the structure of these three small helices is certainly labile. Furthermore, helices H4 and H5 on one hand and H10 and H11 on the other hand appear as long single α helices in 4PRG, chain A.

Internal dynamics of the complex

During the 11.5 ns unrestrained simulation of the LBD-ligand complex (MD0), the main deviations from the X-ray secondary structure occur at the level of H3′, H6′, H10, H12 and in the H2′–H3 loop (Fig. 2). The structure of H3′ and H6′ becomes α helical, H6′ and H7 frequently collapse as a single continuous helix, H10/11 exhibits a kink in the middle and H12 appears as a 3_{10} helix most of the time. In the H2′–H3 loop a short 3_{10} helix is formed including residues 270–273 during the last half of the simulation. Analysis of the set of 35 X-ray structures reveals that H10/11 is curved or kinked in five structures and that the additional helix in the H2′–H3 loop is also present in the structure of Hopkins et al. (2006) (PDB # 2HFP). Furthermore H12 is not characterized as α helix in PDB # 2F4B chain B (Mahindroo et al. 2006). Therefore all

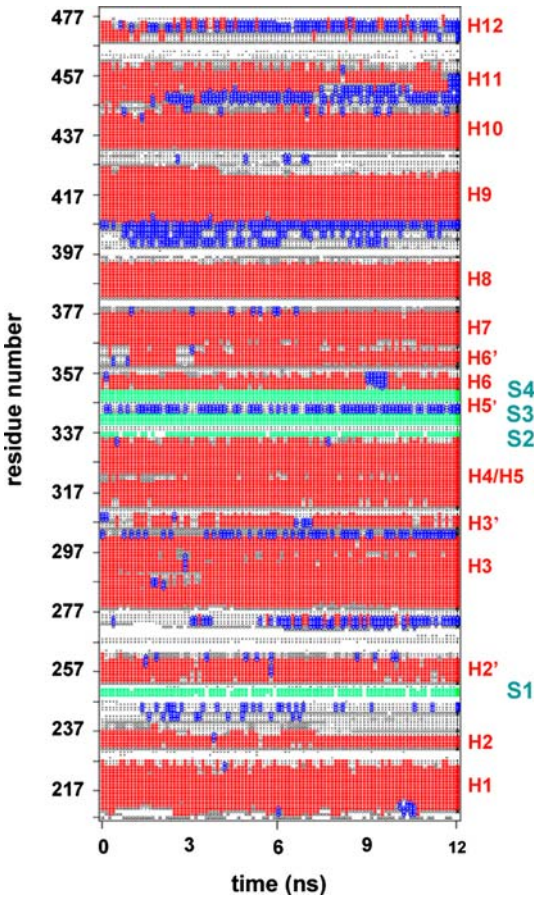

Fig. 2 Time evolution of the secondary structure of the LBD of PPARγ during the 11.5 ns of the free simulation MD0 (obtained with MolMol and represented with a home made program). The helices and the strands of the β-sheet s_i (i = 1–4) are indicated on the *right ordinate*. *Red* α helix structure; *green* β-strands; *blue* 3_{10} helix structure

deviations from the initial structure (PDB # 4PRG chain A) found in the present simulation correspond to the variability observed within the set of experimental structures.

The internal dynamics of the receptor is revealed by the fluctuations of the C_α positions as illustrated in Fig. 3a. The base line is about 0.8–0.9 Å and seven peaks of large atomic fluctuations are observed (excluding end effects) corresponding to helices H3′, H6′ and H12 with adjacent residues, the loop joining H2 and the first strand of the β sheet, and the loops H2′–H3, H8–H9 and H9–H10. The most flexible region is the H2′–H3 loop with standard deviation of 4 Å. The flexibility of this region is in agreement with the experimental temperature factors in 4PRG and with the observation that this loop is unresolved in a number of X-ray structures. Amino acids located between H2 and the first β-strand s1 on one hand and the set of residues joining H9 and H10 on the other hand are also largely agitated with a standard deviation of 3 Å. Four

other lower peaks with amplitude ranging around 1.5–2 Å are observed at the level of H3′ and the flanking loops, the residues from the middle of H6 to the end of H6′, the H8–H9 loop and the C-terminus of the two helices H11 and H12. During all the time course of the MD0 simulation, the ligand remains H-bonded to Glu 259 as in the X-ray structure.

RMSD and restraint energy

The eight restrained MD simulations are successful in bringing the ligand out of the receptor. The average value of the restraint energy is 3.5 kcal mol^{-1} when applying the TDR method and 0.55 kcal mol^{-1} when using TMD^{-1} with, in both cases, a standard deviation of 1.7 kcal mol^{-1}. These values are smaller than the average fluctuation of the potential energy of the complex (103 kcal mol^{-1}) and of the ligand (6.8 kcal mol^{-1}) at equilibrium as calculated from MD0. However, for two simulations, MD7 and MD8, a long 3-ns relaxation period during which the restrained quantities were maintained constant was necessary after 2 ns of simulation, because of the H2′–H3 loop that stopped the progression of the ligand. The necessity for such a longer relaxation time results from the inclusion of the amino acid side chains in the restraints. The number of degrees of freedom to be restrained is considerably increased and, as a consequence, the time spent by the binding pocket and the ligand to explore the different conformational possibilities is much longer.

The time evolution of the protein structure has been monitored by calculating the RMSD time series for each simulation. In the case of MD0 a plateau at 2.6 Å is observed after 2.5 ns. In six restrained simulations a plateau is also reached in the range 2.2–3.5 Å after 3–6 ns depending on simulation conditions (see Table 2 and figures in supplemental materials). This indicates that the overall structure of the protein is not much perturbed during the ligand-extraction process. This is confirmed by the graphic visualization of the different trajectories. However, for two simulations, MD4 and particularly MD7, the plateau value is significantly higher.

Gates and pathways

Several pathways are found for the expulsion of GW0072 from LBD. In six restrained simulations (MD1-3 and MD5-7), GW0072 uses roughly the same gate for leaving the protein. This gate (gate A) corresponds to a groove delimited by H3 on one side, H2 and the H1–H2 loop on the other side, the C-terminal extremity of H1 on the bottom and the last β-strand s4, helix H2′ and the H2′–H3 loop on the top (Fig. 4a). In the X-ray structure of the complex, the short helix H5′ lies across the gate, locking partially the groove as a safety belt, in the middle of the groove. For five pathways over the six corresponding to this gate the behavior of the ligand is approximately the same. The exit point is located between helix H5' and the H1–H2 loop, due to the sliding of GW0072. The two phenyl groups are first extracted, followed by the aliphatic arm and finally by

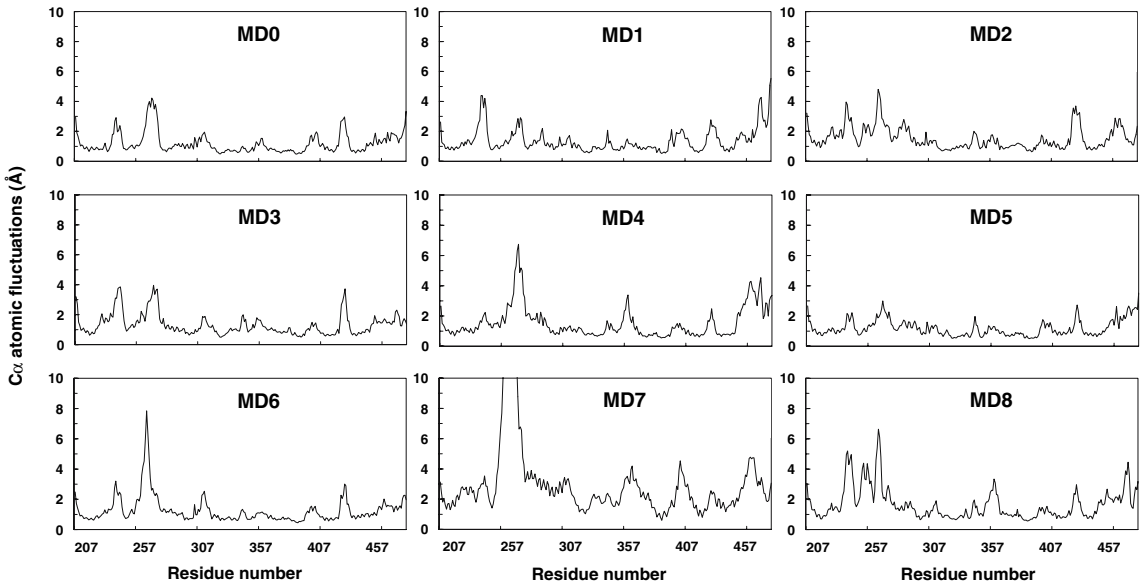

Fig. 3 Fluctuations of Cα positions of LBD during the different simulations

Table 2 Plateau values of the RMSD (in Å) calculated for the different MD simulations

Simulation	MD0	MD1	MD2	MD3	MD4	MD5	MD6	MD7	MD8
RMSD	2.6	2.8	3.5	2.4	4.2	2.2	3.0	6.5	3.0

RMSD values were calculated on the Cα of the LBD of PPARγ

the carboxyl branch. This pathway will be designated pathway A1 in the following. For the last simulation of this family (MD6), the behavior of the ligand is more complex. The orientation of GW0072 is early inverted within the pocket so that the carboxyl branch oscillates, making H-bonds with residues pertaining either to H5′ or to H3. The aliphatic branch is first expulsed, followed by the two phenyl rings and then the carboxyl branch, between H5′, H2′, the H2′–H3 loop and the N-terminal extremity of H3. This path named A2 is made possible because of a large displacement of the H2′–H3 loop. The exit zones of paths A1 and A2 are located in the same groove but are separated by H5′.

The two other simulations, MD4 and MD8, lead to a different exit gate (gate B) delimited by H2′, H6, the N-terminal end of H3 and the loops H2′–H3, H6–H7 and H11–H12 (Fig. 4b). As in the case described above, this gate is also divided into two parts by the small helix H6′. In MD4, defining pathway B1, the exit point is bordered by the two loops H11–H12 and H6–H7 (including H6′) and by the N-terminus of H3 and the C-terminus of H11. In MD8 that follows pathway B2, the exit point is located on the other half of the gate, near the H2′–H3 loop. The different pathways are shown in Fig. 5, with the ligand trajectory illustrated by the positions of the geometrical center of GW0072 during the course of the different simulations. An example of the main route A1 is given in Fig. 6 for MD1. It is interesting to note that in MD7, GW0072 first follows pathway B1/B2, but is stopped by the H2′–H3 loop, and finally adopts pathway A1.

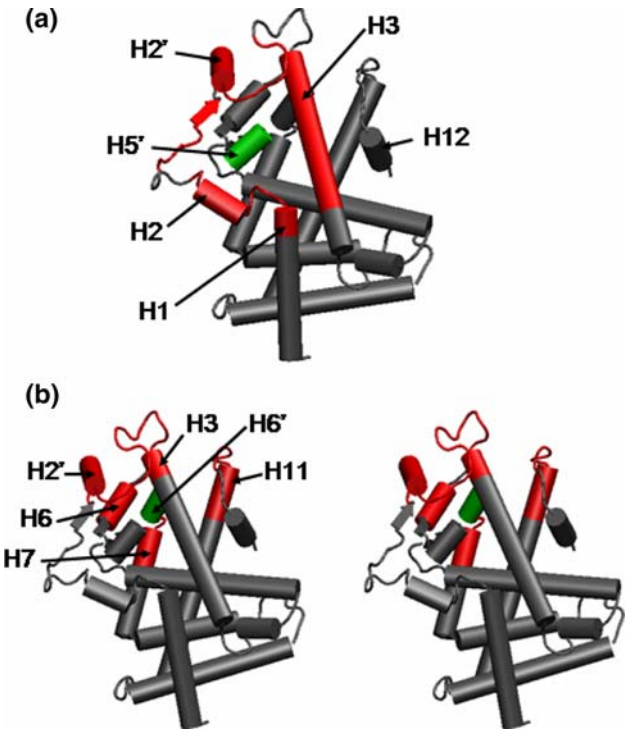

Fig. 4 Views of the exit gates: **a** gate A; **b** stereo view of gate B. The boarder of the gates is colored in *red*. The short H5′ and H6′ helices are colored in *green*. The structure of the receptor corresponds to a schematic representation of the X-ray structure of the LBD of PPARγ (Oberfield et al. 1999)

Hydrogen bonds and pathways

The different extraction pathways are determined by the succession of erratic H-bond (HB) interactions between GW0072 and the receptor. One of the strongest HBs implies the carboxyl group of GW0072. Indeed the oxygen atoms of this group are always accessible, while the nitrogen atoms are not. The other oxygen atoms of GW0072 are more or less hindered, depending on ligand conformation. During the early stage of the extraction process, up to about 3 ns, the carboxyl group remains H-bonded to the Glu259 side chain and, depending on the simulations, different H-bond patterns are observed. The

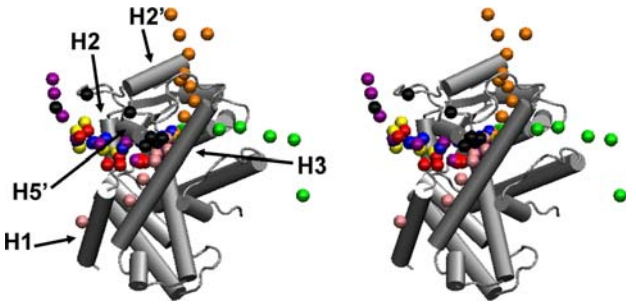

Fig. 5 Stereo view of the different exiting routes of the ligand. In *gray* the X-ray structure of the LBD of PPARγ (Oberfield et al. 1999). The *spheres* represent the position of the centers of geometry of GW0072 calculated every ns for each simulation. *Blue* MD1; *red* MD2; *yellow* MD3; *green* MD4; *purple* MD5; *black* MD6; *pink* MD7; *orange* MD8

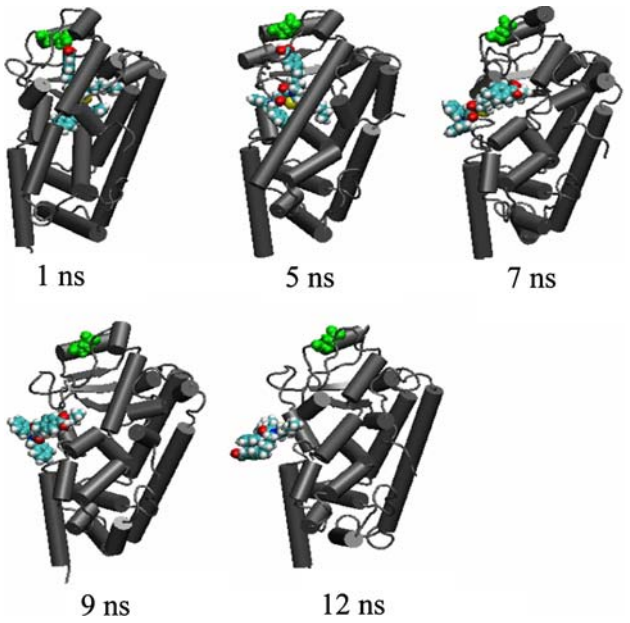

Fig. 6 Snapshots showing the main exit route A1.during the release of GW0072 in.MD1 The ligand is in vdW representation. Glu 259 is colored in *green*

two other oxygen atoms of GW0072 are also frequently involved in HBs, especially with Ser342, Arg288 or Arg280 side chains. It is important to note that HBs between the ligand and the amino acids of the terminal helix H12 containing AF-2 are never observed during the time course of the different simulations.

Internal dynamics of the protein during extraction

The same base line at 0.8–0.9 Å and the same regions of high flexibility (peaks) as for MD0 are observed for the fluctuations of atomic positions during all the restrained simulations (Fig. 3). The amplitude of the different peaks reflects the particular pathway followed by the ligand. As an example, pathway A2 detected in MD6, passing close to the H2′–H3 loop, increases the atomic motions at this level. Very few new peaks are created by the restraints; only MD8 exhibits an additional fluctuation peak at the level of H2′.

The analysis of the secondary structure evolution along the trajectories resulting from restrained simulations exhibits approximately the same deviations from the X-ray structure (not shown) as those found in MD0 (Fig. 2). Only two new deviations are observed corresponding to breaks or curvatures in H3 and/or H4/H5, while they occur only temporarily at a much less extent in MD0. The existence of breaks in H4/H5 is also found in three X-ray structures among the set of 35 PDB that we have analyzed.

All these observations on the dynamical behavior of the complex suggest that the ligand takes advantage of the intrinsic flexible zones of the protein to move inside the receptor.

Energetics

Figure 7 shows the time evolution of the interaction energy between GW0072 and the receptor during the extraction process for the different simulations. A common behavior is the continuous increase of the interaction energy from –110 kcal mol^{-1} to a value tending toward 0 kcal mol^{-1}, although some differences exist between the different simulations. No activation energy barrier is detected. The convergence is faster with TMD^{-1} simulations. However it is clear that during the first 2–4 ns of all simulations, the interaction energy remains stable. This corresponds to a phase during which both the binding pocket and the ligand explore the conformational space compatible with the restraints without expulsion of the ligand.

Discussion

The present study aims to get definite information on pathway(s) followed by the partial agonist GW0072, and more generally by other ligands, to unbind from the LBD of PPARγ. Restrained MD simulations under soft conditions were performed and particular emphasis was made on limiting the consequences of too strong restraints during simulations. Indeed strong restraints or simplified force fields may disrupt protein structure leading to artifacts in describing the extraction pathway. We used slow-time variation of the restraints for removing away GW0072 from its binding site, applied either on the distance between the geometric centers of the binding pocket and of the ligand or on the RMSD of a set of atoms belonging to the pocket and the ligand. The use of these restraints does not necessitate the assumption of an *a priori* direction for the displacement of GW0072. The rate of variation of the RMSD used in the TMD^{-1} method (1.8 Å/ns) was on the same order of magnitude as that found for the RMSD of the complex to reach a plateau during a free simulation (MD0), and a similar value was retained for the restrained distance method. Therefore it is expected that sufficient time is given to the different degrees of freedom for relaxing during the restrained simulations.

The choice of the atoms used in both types of restraints was also dictated by the will to limit artifacts: (1) for the receptor, only the atoms belonging to the same set of nine residues which are the closest to the ligand at once in the three initial structures (I1, I2, I3) and in the X-ray structure are considered (MD1-7); (2) in order to treat equally these nine residues, the restraints are applied only to the

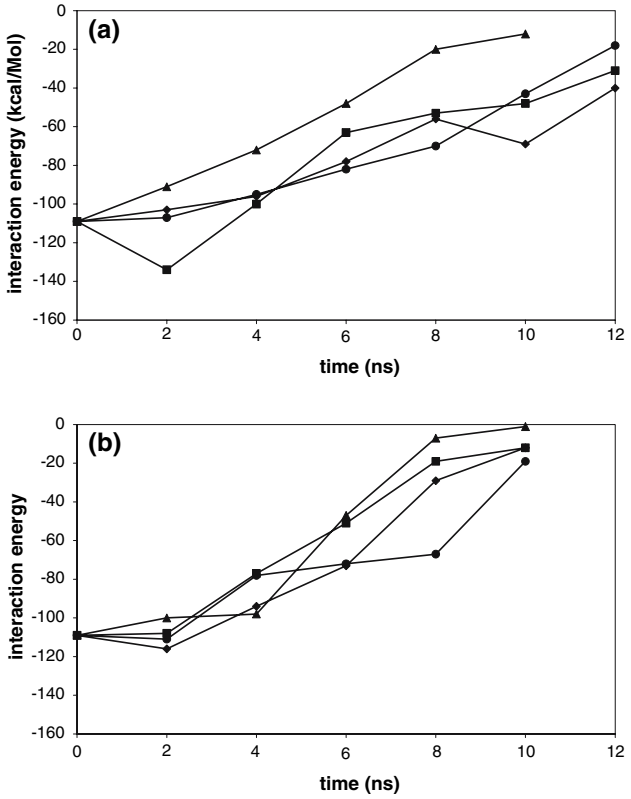

Fig. 7 Time evolution of the interaction energy between GW0072 and the LBD of PPARγ (Oberfield et al. 1999) during the different MD simulations using either TDR (**a**) or TMD^{-1} (**b**)

backbone atoms in six simulations (MD1-6); (3) in MD7 and MD8 all the heavy atoms were used to check the influence of side chains; (4) in MD8, a different set of residues was used for describing the binding pocket; (5) for the ligand the restraints are spread over all the heavy atoms.

The low level of restraints is controlled by the very small value of the restraint energy, much smaller than the energy fluctuations observed during a free simulation. In the case of the TDR method, one could argue that a continuous increase of the distance between the geometrical centers of the pocket and of the ligand can induce a bias and preclude from a temporary distance decrease before the ligand finds another lower energy pathway. We paid careful attention to this possibility by first allowing the distance to evolve freely within a 1 Å window about the targeted distances during simulation and second by performing a simulation (MD8) in which the initial distance was fortuitously very short. Furthermore the coordinates of the geometrical centers being dependent on the fluctuating conformation of the pocket and of the ligand, the system maintains its ability to adjust its conformation for satisfying at the same time the restraints and a low energy level. This last remark is also valid for the TMD^{-1} method. Therefore we expect that the results of our simulations are unbiased by the restraints. This was additionally tested by simulating three different starting structures and using each time two different types of restraints. These structures, extracted from a 11.5-ns free simulation, exhibit crossed RMSD values in the range 2.1–3.0 Å, attesting that they are sufficiently different for meaningful comparison.

Two gates were found but for each gate the accurate pathway may differ according to the evolution of the intermolecular H-bond pattern. Nevertheless five simulations give almost the same pathway, although they were performed under different conditions of initial structures and restraint methods. Furthermore, the side chains were restrained in in one of them (MD7), whereas they were not restrained in the four others. This reinforces our feeling that the pathways found here are really method independent. During the course of the different simulations, intermolecular interaction energy between the ligand and the protein does not reveal any activation energy barrier. On the contrary, a continuous 110-kcal mol^{-1} energy increase occurs from the equilibrium state of the protein–ligand complex to the unbound state. This is of course in favor of a high affinity for the binding of GW0072 to the LBD of PPARγ. The main exit gate, found in six simulations, corresponds roughly to the groove delimited by helices H1, H2, H3, and the H2-s1 and H2′–H3 loops. This gate was hypothesized as being the entry point for ligand binding in the LBD of PPARγ (Cronet et al. 2001; Nolte et al. 1998; Uppenberg et al. 1998; Willson et al. 2001; Xu et al. 1999). Although the present study does not completely validate this hypothesis, it shows at least that this gate is an exit point and that there are routes for the ligand to reach the binding site. These routes are very similar to pathways A and 3 described by Carlson et al. (2006) and by Kosztin et al. (1999), respectively, for the unbinding of retinoic acid from the retinoic acid receptor and to pathway III found by Martinez et al. (Martinez et al. 2005) for the extraction of different ligands of thyroid hormone receptors. These paths are located on the opposite side of H12 and it is to be noted that no interaction between the ligand and this last helix is observed during the extraction process. We have found that the short H5′ helix plays an important role in closing the gate, preventing GW0072 to escape easily from the binding pocket. The gate opening is achieved mainly by small temporary displacements of H5′ and adjacent residues comprising s3 and s4 permitted by the intrinsic flexibility of the receptor that exists at equilibrium. During the migration of GW0072 from its equilibrium position toward H5′, all routes using this gate are identical, but beyond H5′, different paths are possible. The fact that in several X-ray structures helix H5′ is replaced by a loop, certainly does not modify the dynamical role of the corresponding residues.

Another gate was found in two simulations, bordered by H2′, H6, the N-terminal end of H3 and the H2′–H3 and H6–H7 loops. The corresponding routes are very similar to pathways D of Carlsson et al. (2006), 2 of Kosztin et al. (1999) and II of Martinez et al. (2005). This path has also been described as the unique one by Blondel et al. (1999) for the unbinding of retinoic acid from the retinoic acid receptor. It is interesting to note that both gates are half-closed in their middle by a short labile helix (H5′or H6′), and that GW0072 can pass these helices on one side or on the other. Considering that GW0072 is a bulky ligand, it could be expected that the routes described here may also be used by other smaller molecules.

The present simulations also show that the ligand uses the intrinsic flexible regions of the receptor for escaping from its binding pocket without disturbing greatly the average structure of the protein, as if LBD were a bead of 7–8 nearly-rigid bodies linked by soft springs.

In conclusion, this study evidences several gates that are very similar to those described for the unbinding of various ligands from other nuclear receptors, suggesting that they are a common ligand-independent property of this superfamily. Only eight restrained simulations have been performed within reasonable delay, precluding from an accurate determination of the probability that each gate is used by the ligand. However, the present results could serve as a good reference for more extensive MD studies using more drastic conditions, such as implicit representation of the solvent or faster variation of the restraints. As an example, in two preliminary simulations performed under faster variation (by a factor 6) of the restraints, the ligand uses the major gate described in the present study for exiting from the receptor (results not shown). Finally we feel that the knowledge of entry/exit pathways of a receptor should be useful in discriminating between different ligands that could have been favorably docked in the binding pocket, by introducing into computational drug design protocols an additional element at the scoring step.

Acknowledgments The technical assistance of A. Boyer is greatly appreciated. This work was partly supported by the Region Centre.

References

Aci S, Ramstein J, Genest D (2004) Base pairing at the stem-loop junction in the SL1 kissing complex of HIV-1 RNA: a thermodynamic study probed by molecular dynamics simulation. J Biomol Struct Dyn 21:833–840

Aci S, Mazier S, Genest D (2005) Conformational pathway for the kissing complex–extended dimer transition of the SL1 stem-loop from genomic HIV-1 RNA as monitored by targeted molecular dynamics techniques. J Mol Biol 351:520–530

Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000) The Protein Data Bank. Nucleic Acids Res 28:235–242

Blondel A, Renaud JP, Fischer S, Moras D, Karplus M (1999) Retinoic acid receptor: a simulation analysis of retinoic acid binding and the resulting conformational changes. J Mol Biol 291:101–115

Carlsson P, Burendahl S, Nilsson L (2006) Unbinding of retinoic acid from the retinoic acid receptor by random expulsion molecular dynamics. Biophys J 91:3151–3161

Case DA, Pearlman DA, Caldwell JW, Cheatham ITE, Wang J, Ross WS, Simmerling CL, Darden TA, Merz KM, Stanton RV, Cheng AL, Vincent JJ, Crowley M, Tsui V, Gohlke H, Radmer RJ, Duan Y, Pitera J, Mascova I, Seibel GL, Singh UC, Weiner P, Kollman PA (2002) AMBER 7. University of California, San Francisco

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

Duke RE, Pedersen LG (2003) PMEMD, University of North Carolina, Chapel Hill

Elber R, Karplus M (1990) Enhanced sampling in molecular dynamics: use of the time-dependent Hartree approximation for a simulation of carbon monoxide diffusion through myoglobin. J Am Chem Soc 112:9161–9175

Hopkins CR, O'Neil S V, Laufersweiler MC, Wang Y, Pokross M, Mekel M, Evdokimov A, Walter R, Kontoyianni M, Petrey ME, Sabatakos G, Roesgen JT, Richardson E, Demuth TP Jr (2006) Design and synthesis of novel *N*-sulfonyl-2-indole carboxamides as potent PPAR-gamma binding agents with potential application to the treatment of osteoporosis. Bioorg Med Chem Lett 16:5659–5663

Humphrey W, Dalke A, Schulten K (1996) VMD: visual molecular dynamics. J Mol Graph 14(33–8):27–28

Inzucchi SE, Maggs DG, Spollett GR, Page SL, Rife FS, Walton V, Shulman GI (1998) Efficacy and metabolic effects of metformin and troglitazone in type II diabetes mellitus. N Engl J Med 338:867–872

Isralewitz B, Baudry J, Gullingsrud J, Kosztin D, Schulten K (2001) Steered molecular dynamics investigations of protein function. J Mol Graph Model 19:13–25

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

Koradi R, Billeter M, Wuthrich K (1996) MOLMOL: a program for display and analysis of macromolecular structures. J Mol Graph 14(51–5):29–32

Kosztin D, Izrailev S, Schulten K (1999) Unbinding of retinoic acid from its receptor studied by steered molecular dynamics. Biophys J 76:188–197

Kuhn B, Hilpert H, Benz J, Binggeli A, Grether U, Humm R, Marki HP, Meyer M, Mohr P (2006) Structure-based design of indole propionic acids as novel PPARalpha/gamma co-agonists. Bioorg Med Chem Lett 16:4016–4020

Li Y, Choi M, Suino K, Kovach A, Daugherty J, Kliewer SA, Xu HE (2005) Structural and biochemical basis for selective repression of the orphan nuclear receptor liver receptor homolog 1 by small heterodimer partner. Proc Natl Acad Sci USA 102:9505–9510

Liang J, Edelsbrunner H, Woodward C (1998) Anatomy of protein pockets and cavities: measurement of binding site geometry and implications for ligand design. Protein Sci 7:1884–1897

Lu IL, Huang CF, Peng YH, Lin YT, Hsieh HP, Chen CT, Lien TW, Lee HJ, Mahindroo N, Prakash E, Yueh A, Chen HY, Goparaju CM, Chen X, Liao CC, Chao YS, Hsu JT, Wu SY (2006) Structure-based drug design of a novel family of PPARgamma partial agonists: virtual screening, X-ray crystallography, and in vitro/in vivo biological activities. J Med Chem 49:2703–2712

Ludemann SK, Lounnas V, Wade RC (2000) How do substrates enter and products exit the buried active site of cytochrome P450cam? 2. Steered molecular dynamics and adiabatic mapping of substrate pathways. J Mol Biol 303:813–830

Mahindroo N, Huang CF, Peng YH, Wang CC, Liao CC, Lien TW, Chittimalla SK, Huang WJ, Chai CH, Prakash E, Chen CP, Hsu TA, Peng CH, Lu IL, Lee LH, Chang YW, Chen WC, Chou YC, Chen CT, Goparaju CM, Chen YS, Lan SJ, Yu MC, Chen X, Chao YS, Wu SY, Hsieh HP (2005) Novel indole-based peroxisome proliferator-activated receptor agonists: design, SAR, structural biology, and biological activities. J Med Chem 48:8194–8208

Mahindroo N, Wang CC, Liao CC, Huang CF, Lu IL, Lien TW, Peng YH, Huang WJ, Lin YT, Hsu MC, Lin CH, Tsai CH, Hsu JT, Chen X, Lyu PC, Chao YS, Wu SY, Hsieh HP (2006) Indol-1-yl acetic acids as peroxisome proliferator-activated receptor agonists: design, synthesis, structural biology, and molecular docking studies. J Med Chem 49:1212–1216

Martinez L, Sonoda MT, Webb P, Baxter JD, Skaf MS, Polikarpov I (2005) Molecular dynamics simulations reveal multiple pathways of ligand dissociation from thyroid hormone receptors. Biophys J 89:2011–2023

Nolte RT, Wisely GB, Westin S, Cobb JE, Lambert MH, Kurokawa R, Rosenfeld MG, Willson TM, Glass CK, Milburn MV (1998) Ligand binding and co-activator assembly of the peroxisome proliferator-activated receptor-gamma. Nature 395:137–143

Oberfield JL, Collins JL, Holmes CP, Goreham DM, Cooper JP, Cobb JE, Lenhard JM, Hull-Ryde EA, Mohr CP, Blanchard SG, Parks DJ, Moore LB, Lehmann JM, Plunket K, Miller AB, Milburn MV, Kliewer SA, Willson TM (1999) A peroxisome proliferator-activated receptor gamma ligand inhibits adipocyte differentiation. Proc Natl Acad Sci USA 96:6102–6106

Ostberg T, Svensson S, Selen G, Uppenberg J, Thor M, Sundbom M, Sydow-Backman M, Gustavsson AL, Jendeberg L (2004) A new class of peroxisome proliferator-activated receptor agonists with a novel binding epitope shows antidiabetic effects. J Biol Chem 279:41124–41130

Ryckaert JP, Ciccotti C, Berendsen HJC (1977) Numerical integration of the Cartesian equations of motion of a system with constraints: molecular. J Comput Phys 23:327–341

Schlitter J, Engels M, Kruger P (1994) Targeted molecular dynamics: a new approach for searching pathways of conformational transitions. J Mol Graph 12:84–89

Sheu SH, Kaya T, Waxman DJ, Vajda S (2005) Exploring the binding site structure of the PPAR gamma ligand-binding domain by computational solvent mapping. Biochemistry 44:1193–1209

Uppenberg J, Svensson C, Jaki M, Bertilsson G, Jendeberg L, Berkenstam A (1998) Crystal structure of the ligand binding domain of the human nuclear receptor PPARgamma. J Biol Chem 273:31108–31112

Willson TM, Brown PJ, Sternbach DD, Henke BR (2000) The PPARs: from orphan receptors to drug discovery. J Med Chem 43:527–550

Willson TM, Lambert MH, Kliewer SA (2001) Peroxisome proliferator-activated receptor gamma and metabolic disease. Annu Rev Biochem 70:341–367

Xu HE, Lambert MH, Montana VG, Parks DJ, Blanchard SG, Brown PJ, Sternbach DD, Lehmann JM, Wisely GB, Willson TM, Kliewer SA, Milburn MV (1999) Molecular recognition of fatty acids by peroxisome proliferator-activated receptors. Mol Cell 3:397–403