Supporting materials

SI methods

1. Convergence analysis

Both the RMS average correlations (RAC) and Kullback–Leibler divergence (KLD) methods [18-19] were used to access whether the individual simulations are reasonably converged.

The RAC functionality can be thought of as a pseudo-autocorrelation function for RMSD values. This essentially measures the convergence of the overall average structure at different time intervals within a single trajectory. For a given time interval or lag (τ) a straight coordinate running average over that time interval is performed over the entire trajectory [24]; each sliding averaged structure over the time interval τ is then either fit to the first averaged structure (time 0 − τ) or a reference structure specified by the user, and finally the average RMSD value of all averaged structures of length τ is calculated according to:

$$RAC(\tau) = \frac{\sum_{t=0}^{N} RMSD(AvgCrd(t, t+\tau))}{N - \tau + 1} \quad (1)$$

where N is the total number of frames. At time τ=1, this is the standard average RMSD over the whole trajectory. When τ approaches the end of the trajectory length, the value approaches zero and loses meaning. The averaged protein structure along the whole trajectory was used in this study as a reference (see Figure S7).

For the KLD analysis, to ensure that the eigenvectors obtained from each simulation being compared match, the coordinate covariance matrix (using only backbone atoms) was calculated using a combined trajectory from both simulations. Each frame of the trajectory was RMS-fit to the overall average coordinates in order to remove global rotational and translational motions. Next, the projection along these eigenvectors of each coordinate frame from the first simulation trajectory was calculated; this was then repeated for the second simulation trajectory. Finally, at each frame t a histogram for each simulation of the PC projection values for a given PC was constructed, and the overlap of these histograms was calculated by KLD using the following term [18-19]:

$$KLD(t) = \sum_{i=o}^{M} \ln\left(\frac{hPC1_N(t,i)}{hPC2_N(t,i)}\right) hPC1_N(t,i) \quad (2)$$

where $hPCX_N(t,i)$ denotes bin *i* of the histogram from trajectory X for the projection of PC N using data from frames 0 to t, and M is the total number of histogram bins (400 in this case).
The KDL methods histogram was also constructed using a Gaussian kernel density estimator with a bandwidth obtained via the normal distribution approximation of the PC data. The above analyses were carried out by AmberTools 15 package [21] and are present on Figure S8. See references [18-19] for more details.
It should be noted that the cross-correlation approach and other similar analyses are mainly developed for the cMD simulations convergence description where one should expect a gradual decrease of the observed structure deviations with an increase of the simulation time. However, in the aMD runs it is reasonable for more local energy minimums to be sampled and a continuous jumping from one to another minimum to be observed.
The intra simulations convergence, i.e. the reproducibility between the different simulations is also important, were investigated by KLD technique and present on Figures S9 and S10. As it is evident from these graphics there was a good convergence of the whole receptor structure between the aMD runs were aMD7 and aMD8 simulations on a dimer (see Figure S9), whereas an example for poor convergence was found between aMD2 and aMD10 runs (see Figure S10).

2. aMD Reweighting scheme

Reweighting of biased aMD frames is an important procedure and was performed using a previously described protocol [12-13,24-25]. For an aMD simulation of a N_{atom} protein system, the probability distribution along a selected reaction coordinates *A(r)* is written as *p*(A)*, where r denotes the atomic positions $\{r_1,\cdots,r_N\}$. Given the boost potential *ΔV(r)* of each frame, *p*(A)* can be reweighted to recover the canonical ensemble distribution, *p (A)*, as:

$$p(A)_j = p^*(A_j)\frac{\langle e^{\beta\Delta V(r)}\rangle j}{\sum_{j=1}^{M}\langle e^{\beta\Delta V(r)}\rangle j}, \qquad j=1,...,M \tag{3}$$

where M is the number of bins, $\beta = 1/kBT$ and $\langle e\beta\Delta V(r)\rangle j$ is the ensemble-averaged Boltzmann factor of *ΔV(r)* for simulation frames found in the *jth* bin. The above equation provides an exponential average algorithm for the reweighting of aMD simulations. Further, the exponential term was approximated here by the summation of the Maclaurin series of boost potential *ΔV(r)* with the reweighting factor rewritten as:

$$\langle e^{\beta \Delta V} \rangle = \sum_{k=0}^{\infty} \frac{\beta^k}{k!} \langle \Delta V^k \rangle \quad (4)$$

where the subscript *j* has been suppressed. The Maclaurin series expansion to the 10th order has been used here.

A toolkit of Python scripts "PyReweighting" [12] was used to reweight the aMD simulations to calculate free energy profiles and also in the house scripts developed. In many cases, when the number of simulation frames within a bin was lower than a certain limit, i.e. cutoff, then the bin was not sufficiently sampled and thus excluded for reweighing [12]. To overcome this problem, the cutoff was determined by iteratively increasing it until the minimum position of the PMF profile does not change. The bin size was set to 500 for reweighting the aMD simulations here.

Supporting information figures

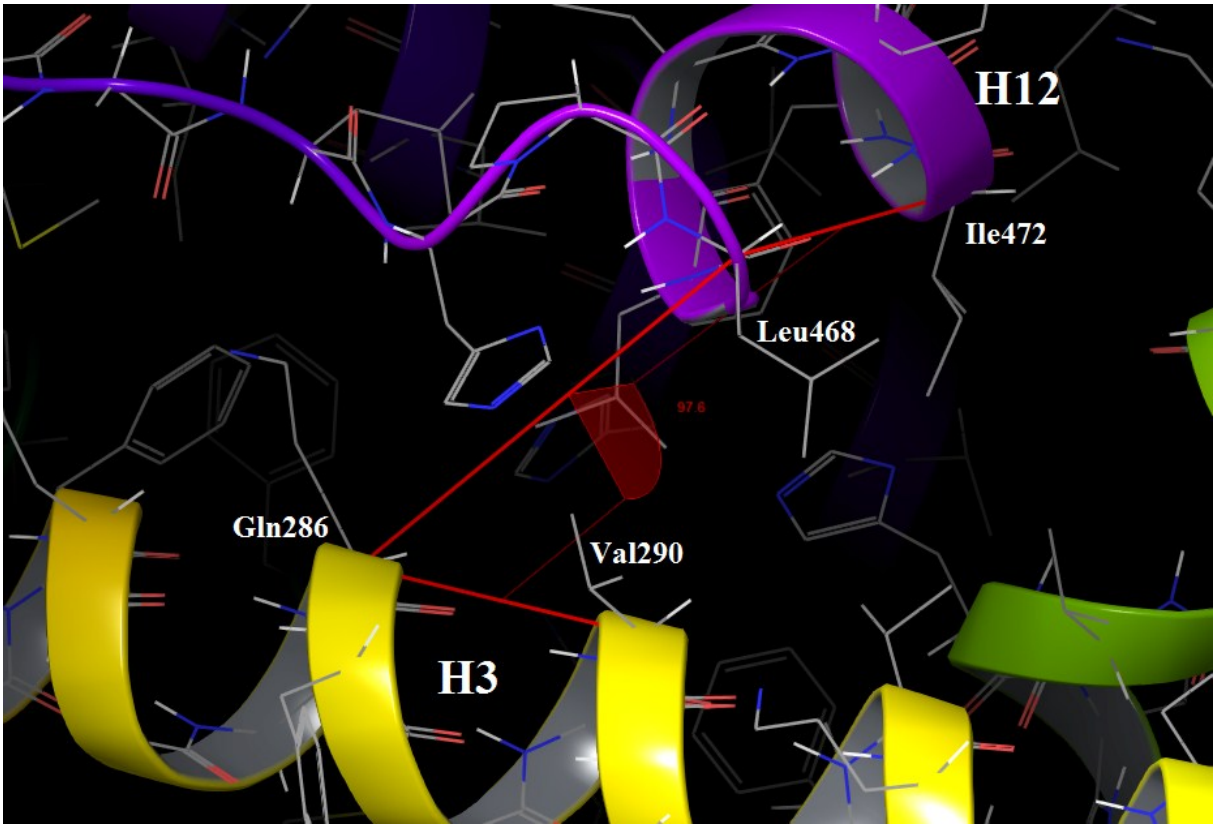

Figure S1A. Selected for the description of the activation helix conformations angle 1 (in red color). An agonist H12 conformation was used for this representation.

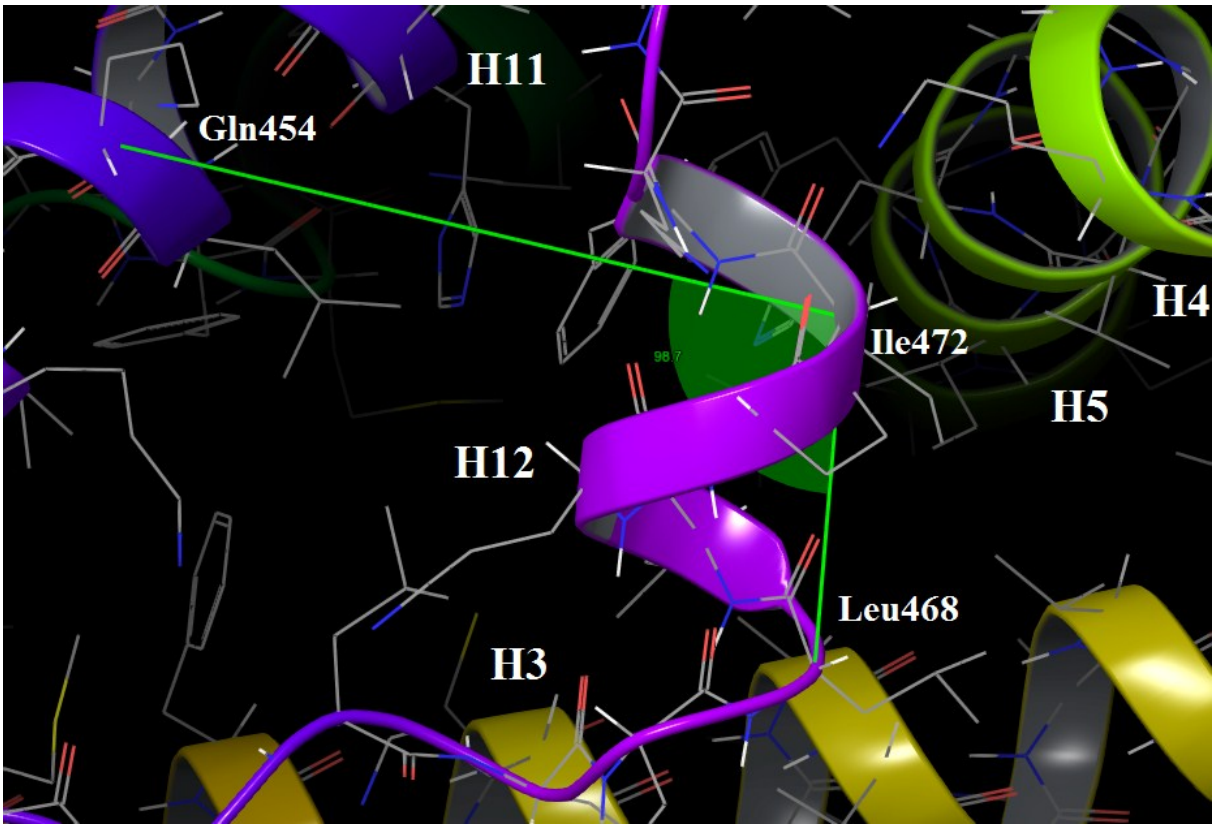

Figure S1B. Selected for the description of the activation helix conformations angle 2 (in red color). An agonist H12 conformation was used for this representation.

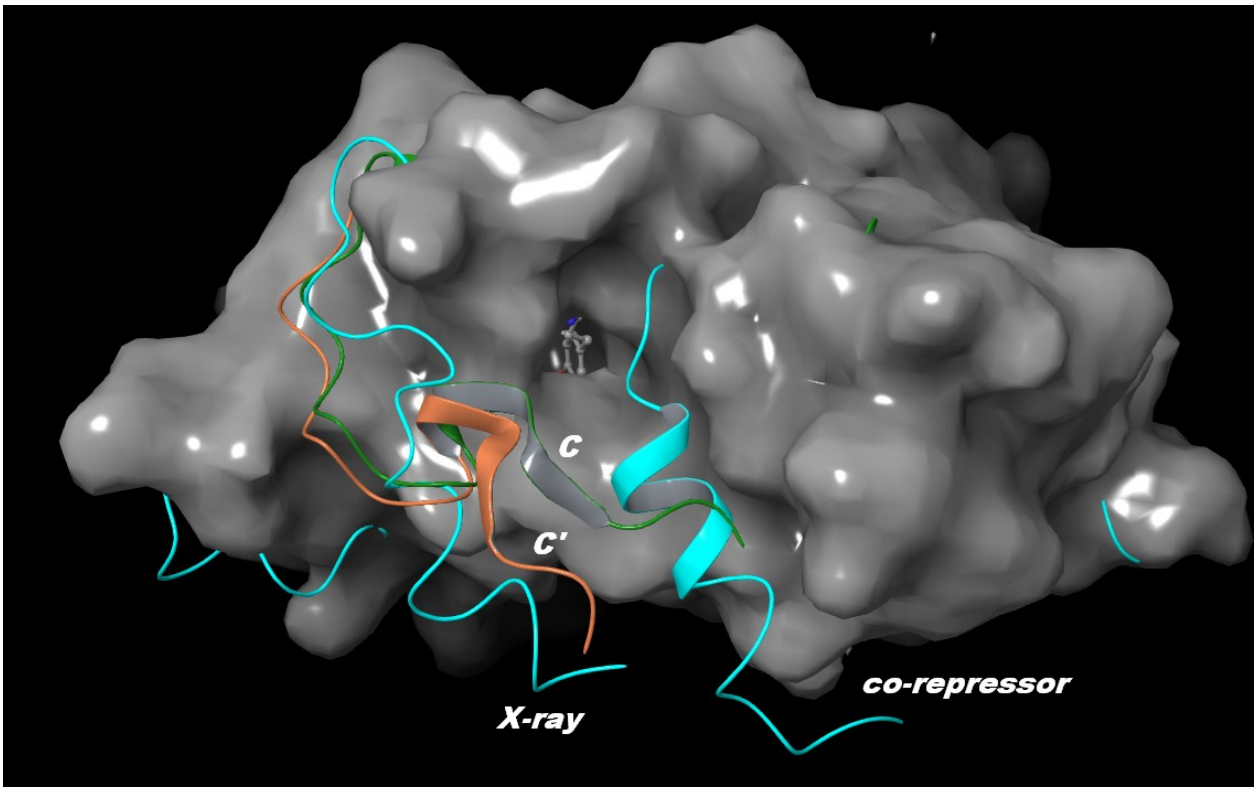

Figure S2. Structural representation of the detected by metadynamics H12 antagonist configurations. Conformations C, C' and the PPARα H12 X-ray structure (pdb id 1kkq) are in a green, ochre and sky blue colors, respectively. Note that the metadynamics indicates a non helical H12 antagonist structure which opens the gate for co-repressors binding in position C'.

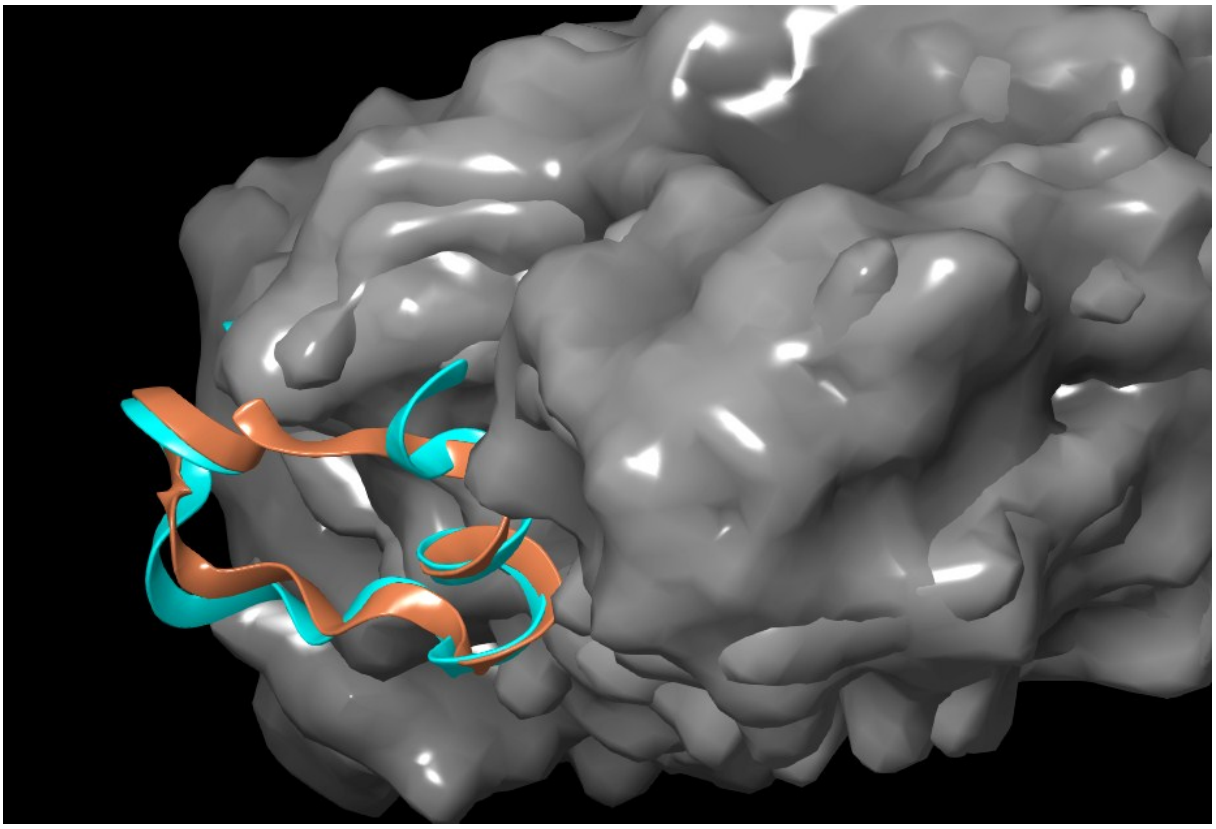

Figure S3. Comparison between averaged H12 conformations based on all aMD trajectories in a chain A of the dimer (in ochre color) and the X-ray structure of the activation helix in the same chain (in sky blue color).

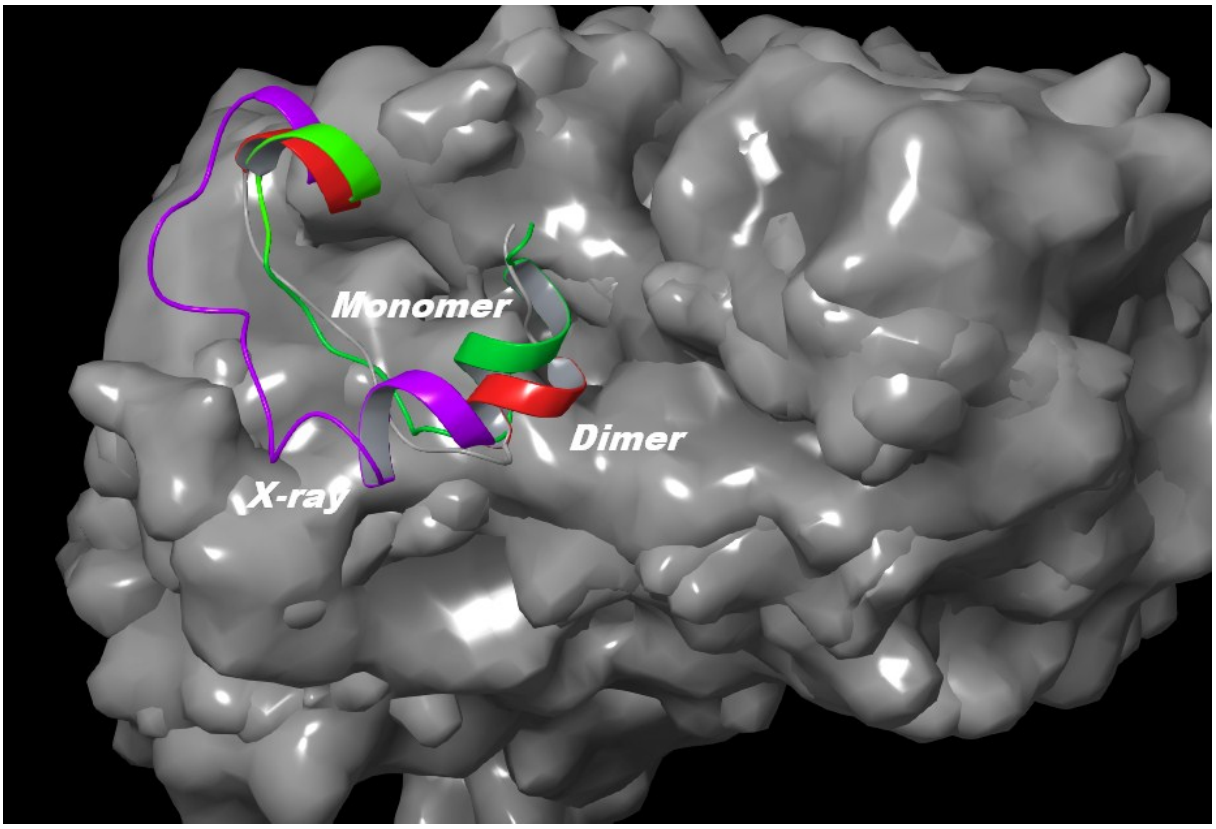

Figure S4. Averaged H12 conformations based on all aMD trajectories in a monomer (in green color) in a chain B of the dimer (in red color). The X-ray activation helix configuration in chain B (pdb id 3vso), which is presumably a crystallographic artifact, is also provided for a comparison (in violet color).

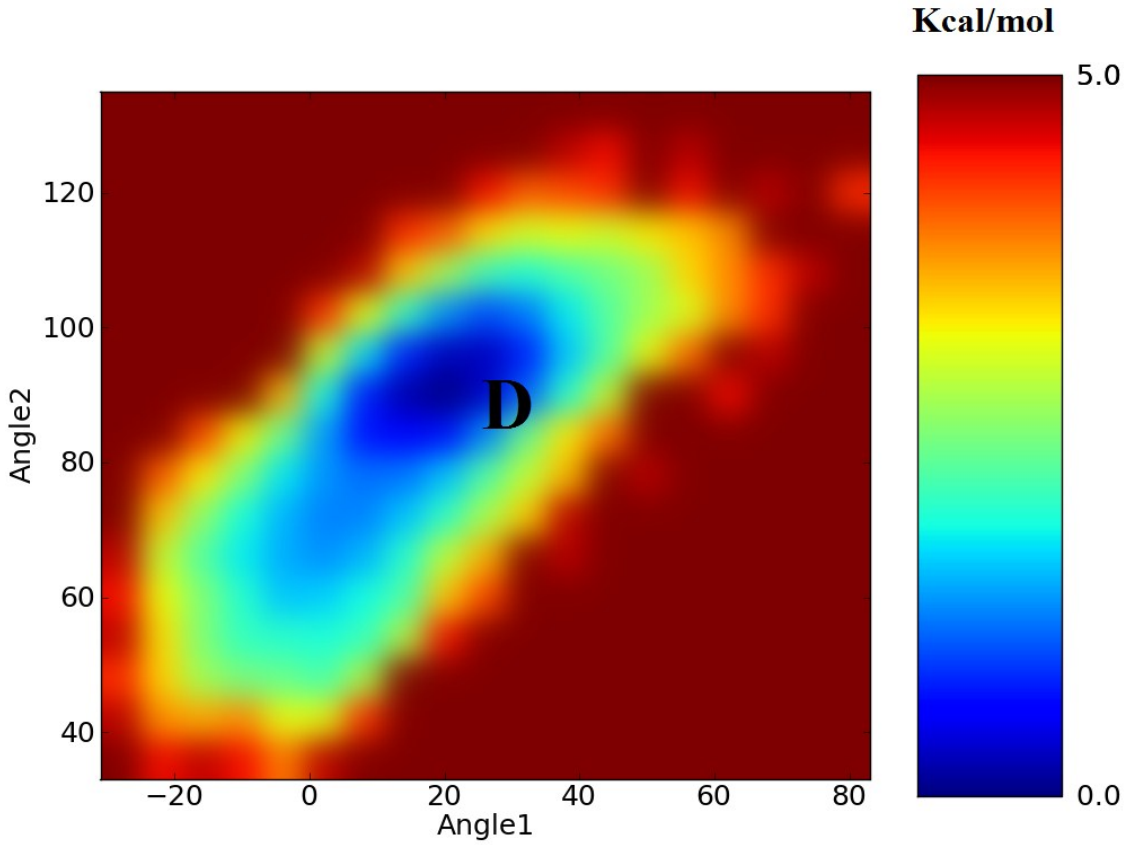

Figure S5. Reweighted free energy plot created by the combined values from two 200-ns-long cMD simulations in a monomer, using the angles 1 and 2 as coordinates**.** Note that only one H12 cluster (point D, see Figure 2A for comparison) was detected.

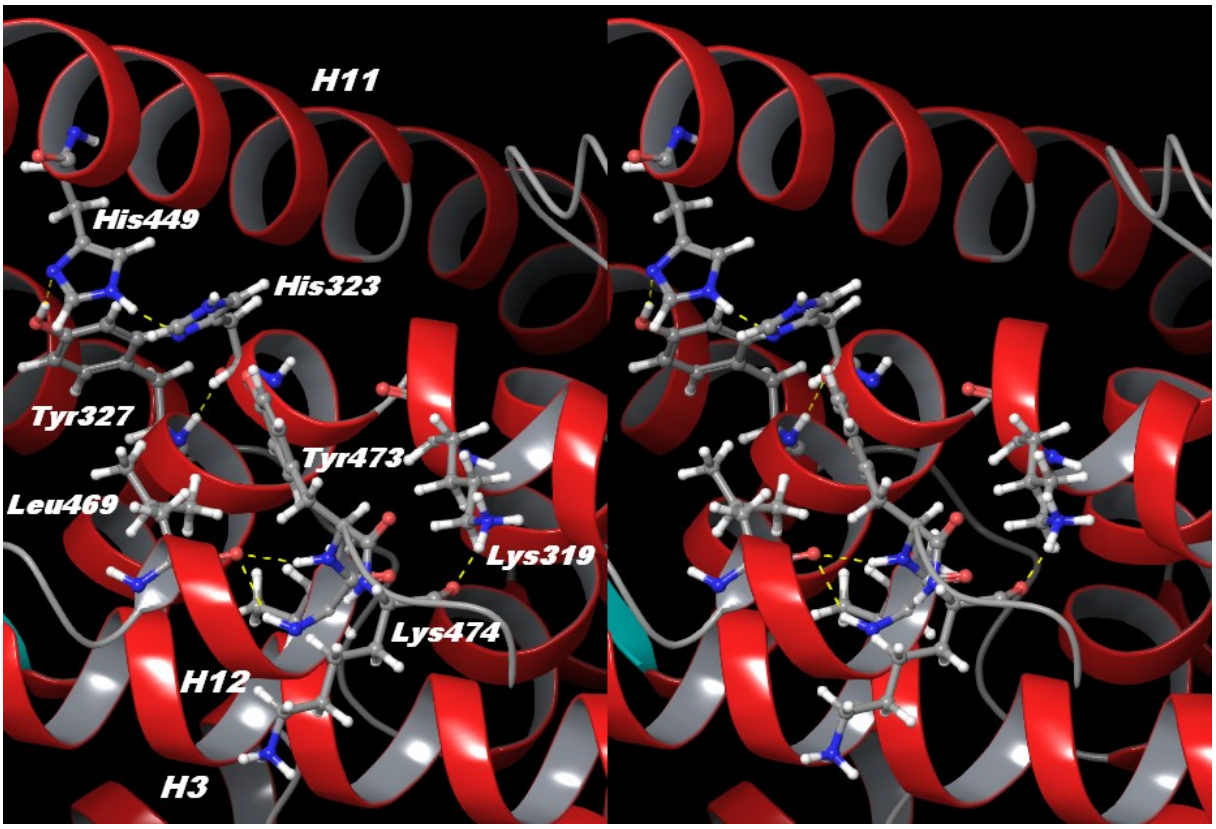

Figure S6A. Stereo view of the identified interactions in LBD of PPARγ in an antagonist state as detected by the aMD simulations (point C).

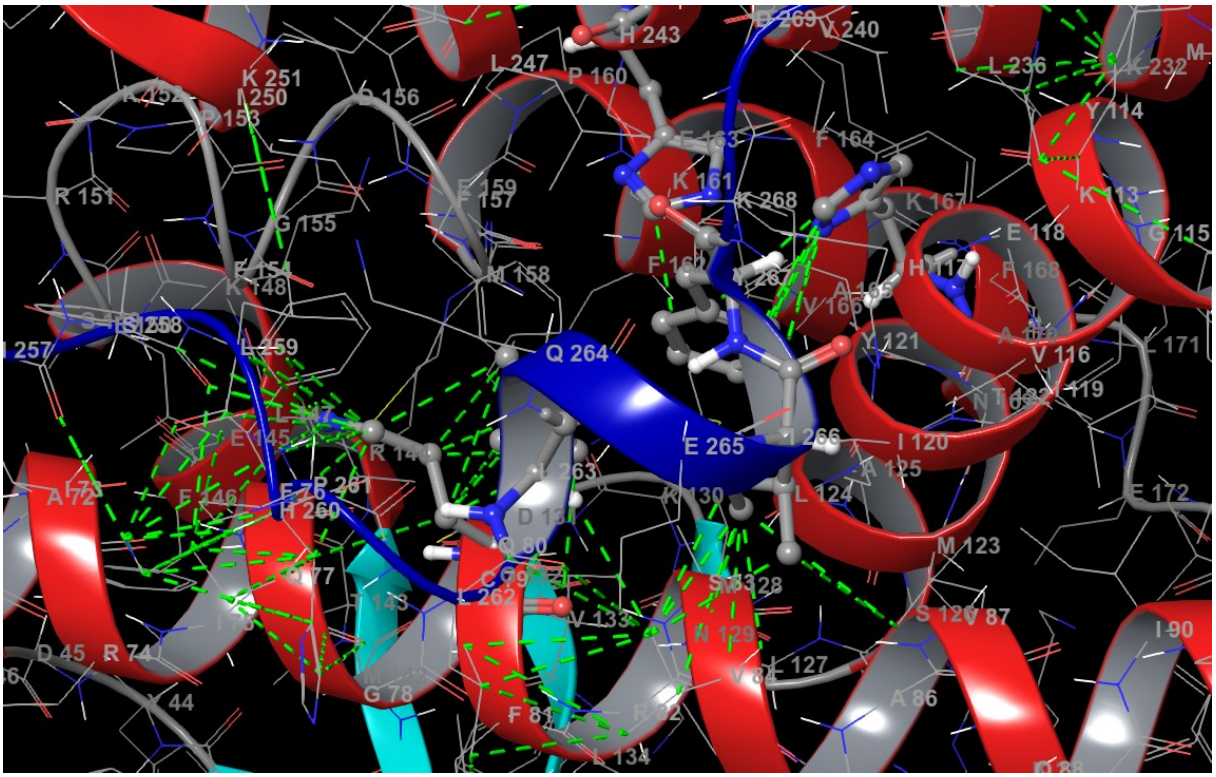

Figure S6B. The good contacts observed between the H12 conformational state B (in blue color) and surrounding PPARγ residues. The "good contact" is defined as an atomic contact at a particular distance. The contacts are good if C=D12/ (R1 + R2) > 1.30 A, where D12 is the distance between the atomic centres 1 and 2, and R1 and R2 are the radii of the atomic centres 1 and 2.

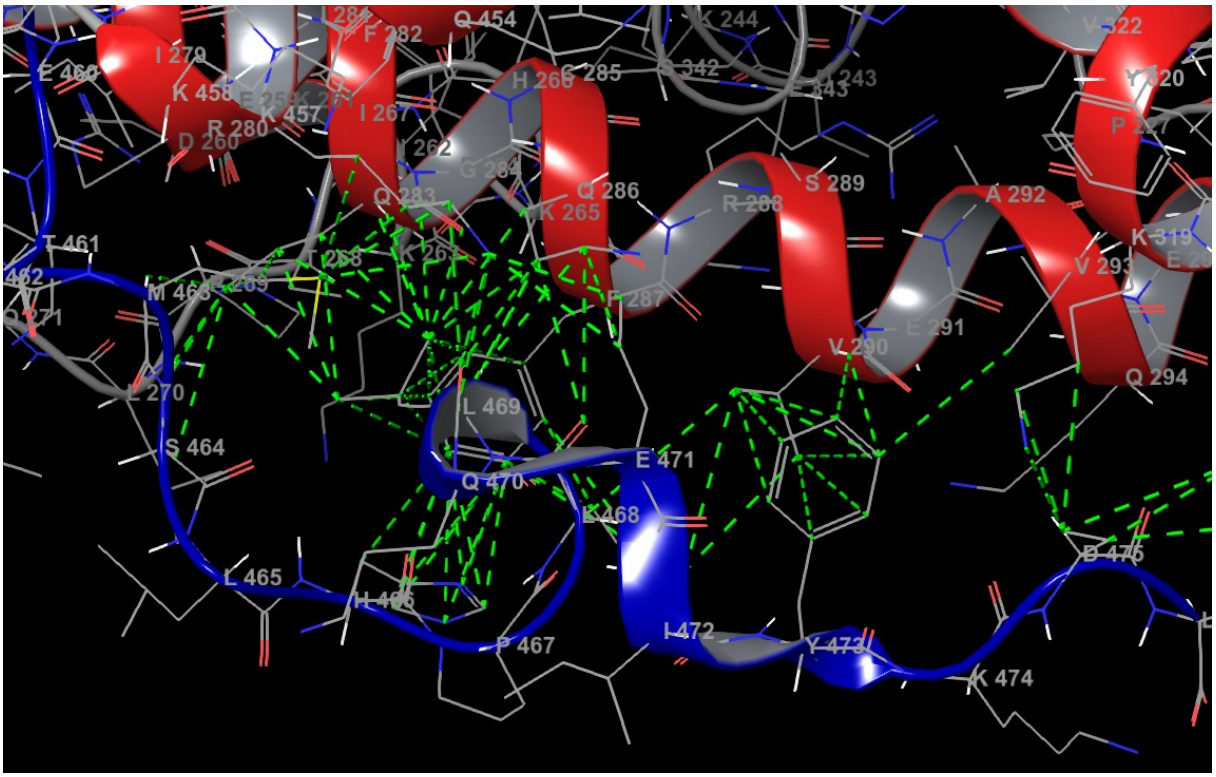

Figure S6C. The good contacts observed between the H12 antagonist conformation (in blue color) and surrounding PPARγ residues. The "good contact" is defined as an atomic contact at a particular distance. The contacts are good if C=D12/ (R1 + R2) > 1.30 A, where D12 is the distance between the atomic centres 1 and 2, and R1 and R2 are the radii of the atomic centres 1 and 2.

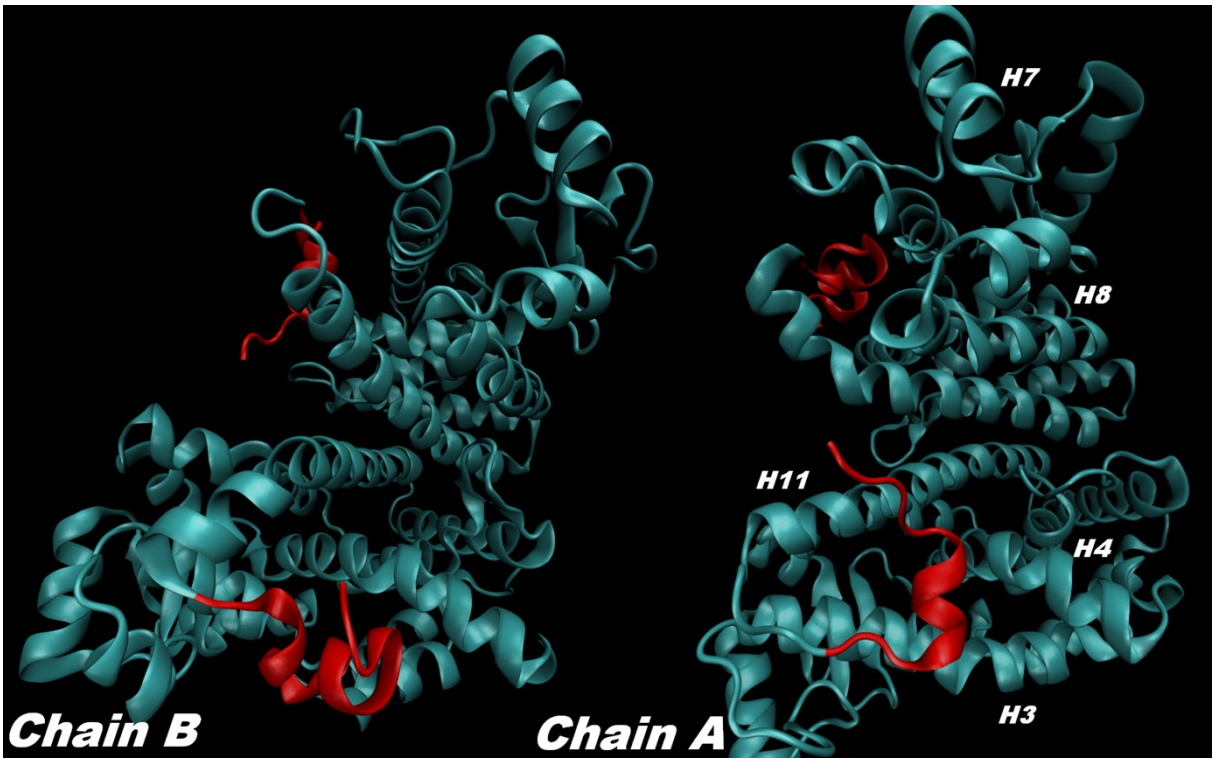

Figure S6D. Structural representation of the H12 (in red color) in a dimer. For a better view the structure was rotated and the Chains A and B are shown separately on the right and left plane, respectively.

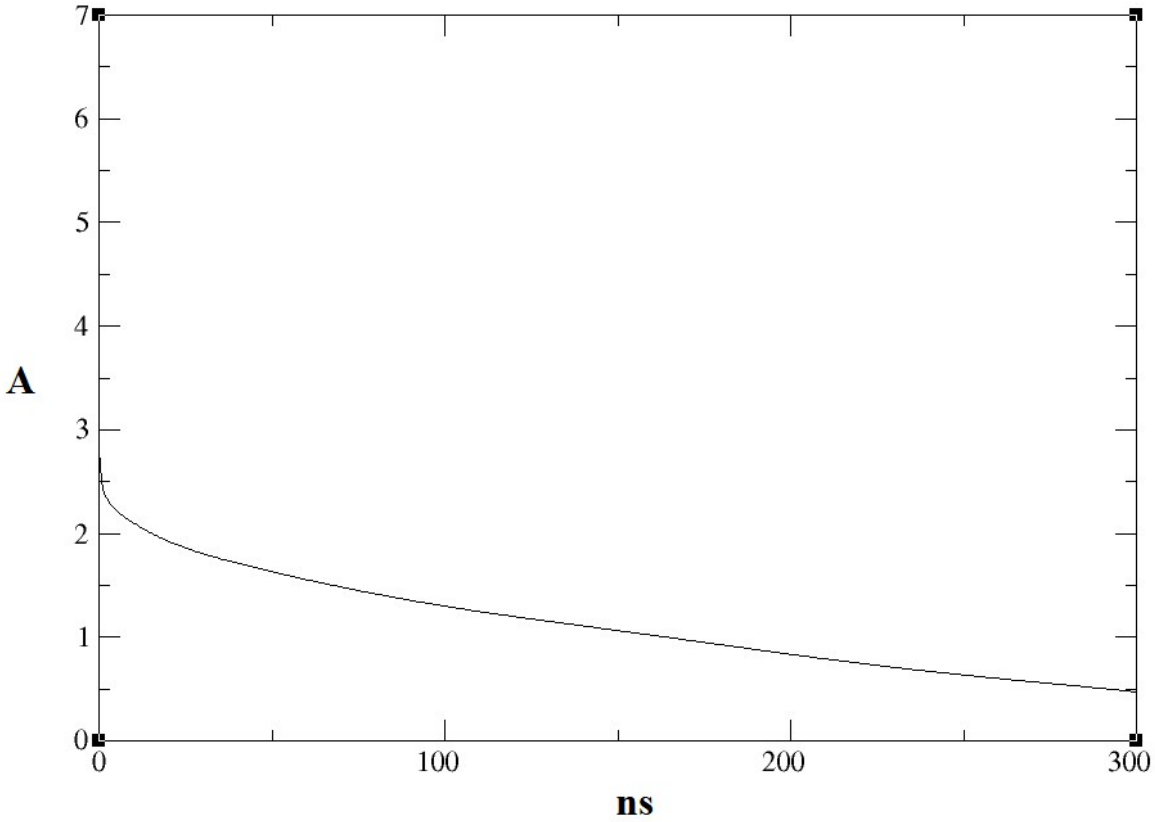

Figure S7. RMS average correlation (RAC) computed at different time intervals of the PPARγ backbone atoms in a dimer of running average structures over the full aMD7 trajectory, with frames spaced at 100 ps intervals, calculated at each time interval from 100 ps to 500 ns, with an offset of 100 frames. The aMD7 run in a dimer was showed up to 300 ns for better representation of the plateau observed after 200 ns of the simulation time. Note that for the interval 200-300ns the structural deviations are less than 0.5 Å.

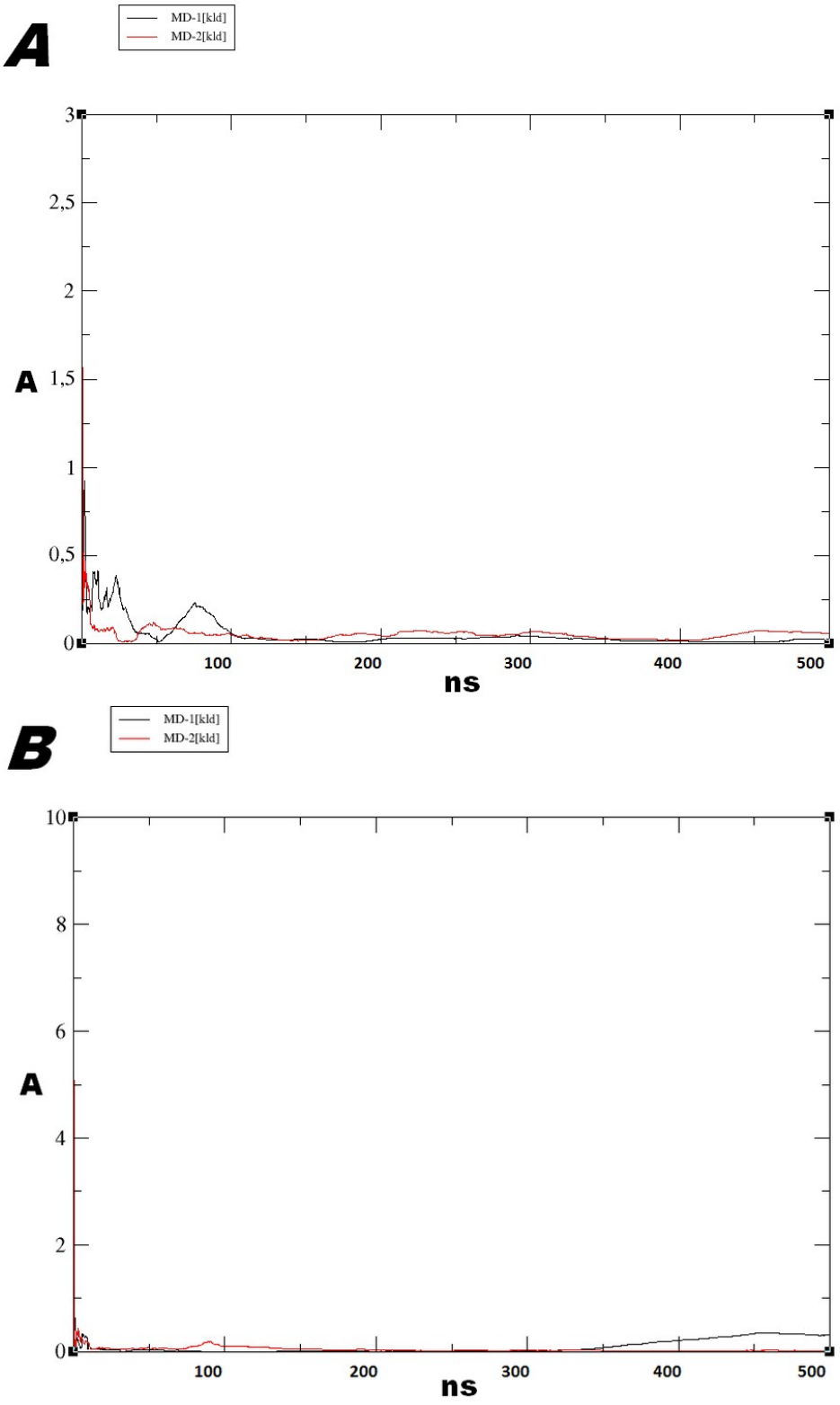

Figure S8. The Kullback−Leibler divergence (KLD) of the first two principal component projection histograms from the independent runs aMD7 and aMD8 in a dimer versus simulations time for: ***(A)*** the whole receptor and ***(B)*** helix 12. PC1 is in black, whereas PC2 is in red.

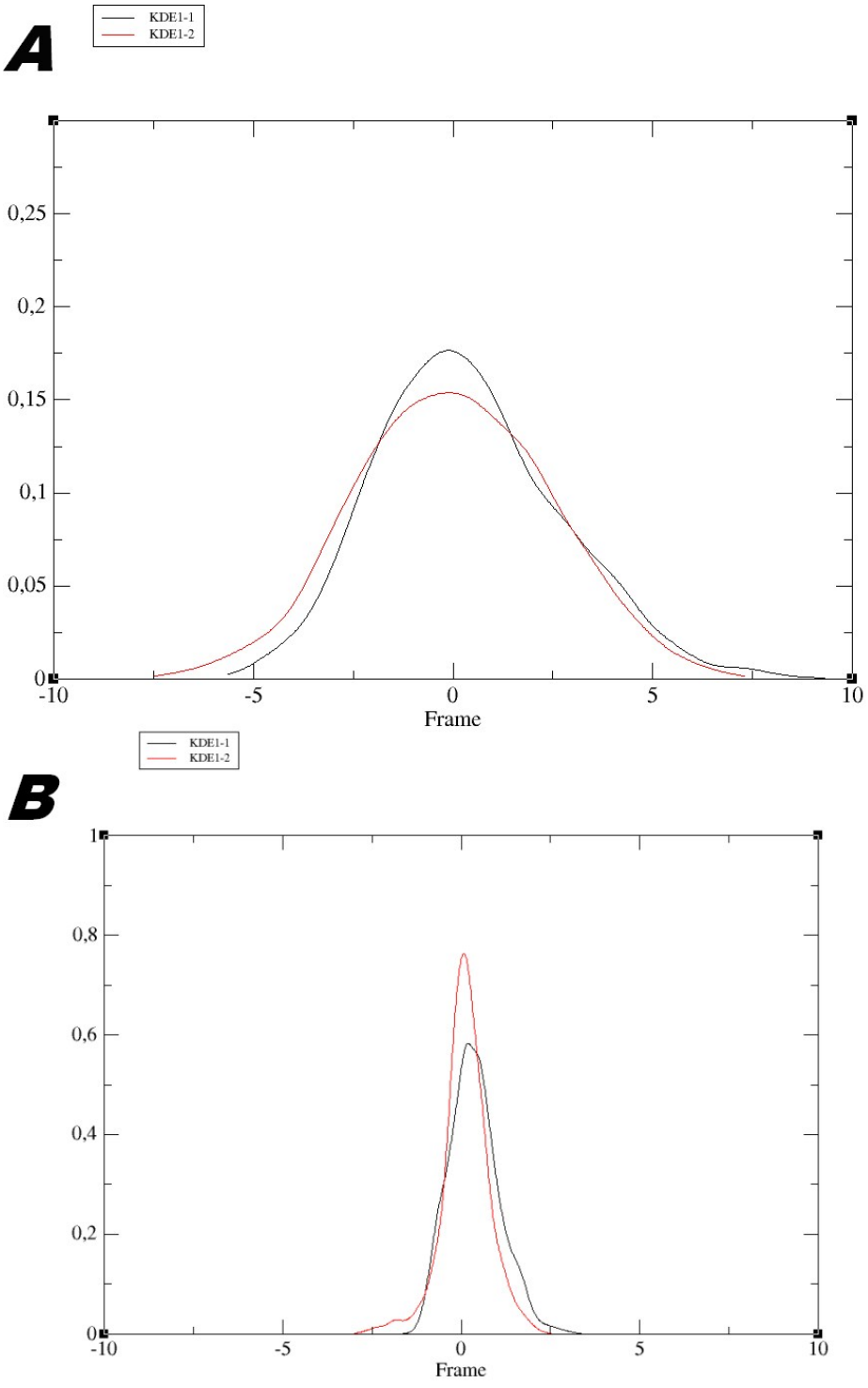

Figure S9. The Kullback−Leibler divergence (KLD) PC histogram from PC analysis in the Cartesian space, calculated from the combined aMD7 and aMD8 simulation trajectories in a dimer for: ***(A)*** the whole receptor and ***(B)*** helix 12.

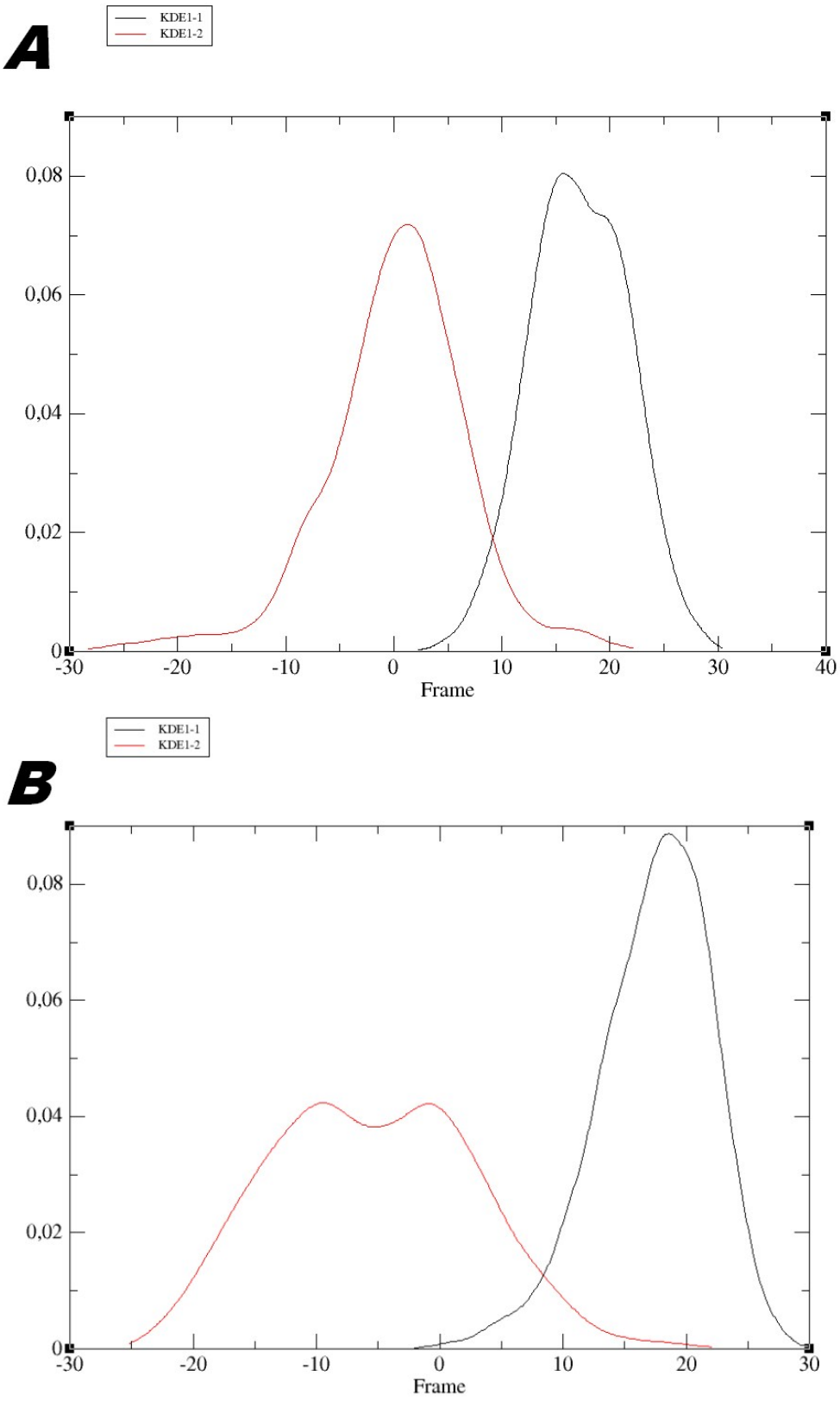

Figure S10. The Kullback−Leibler divergence (KLD) PC histogram from PC analysis in the Cartesian space, calculated from the combined aMD2 and aMD10 simulation trajectories in a dimer for: ***(A)*** the whole receptor and ***(B)*** helix 12.

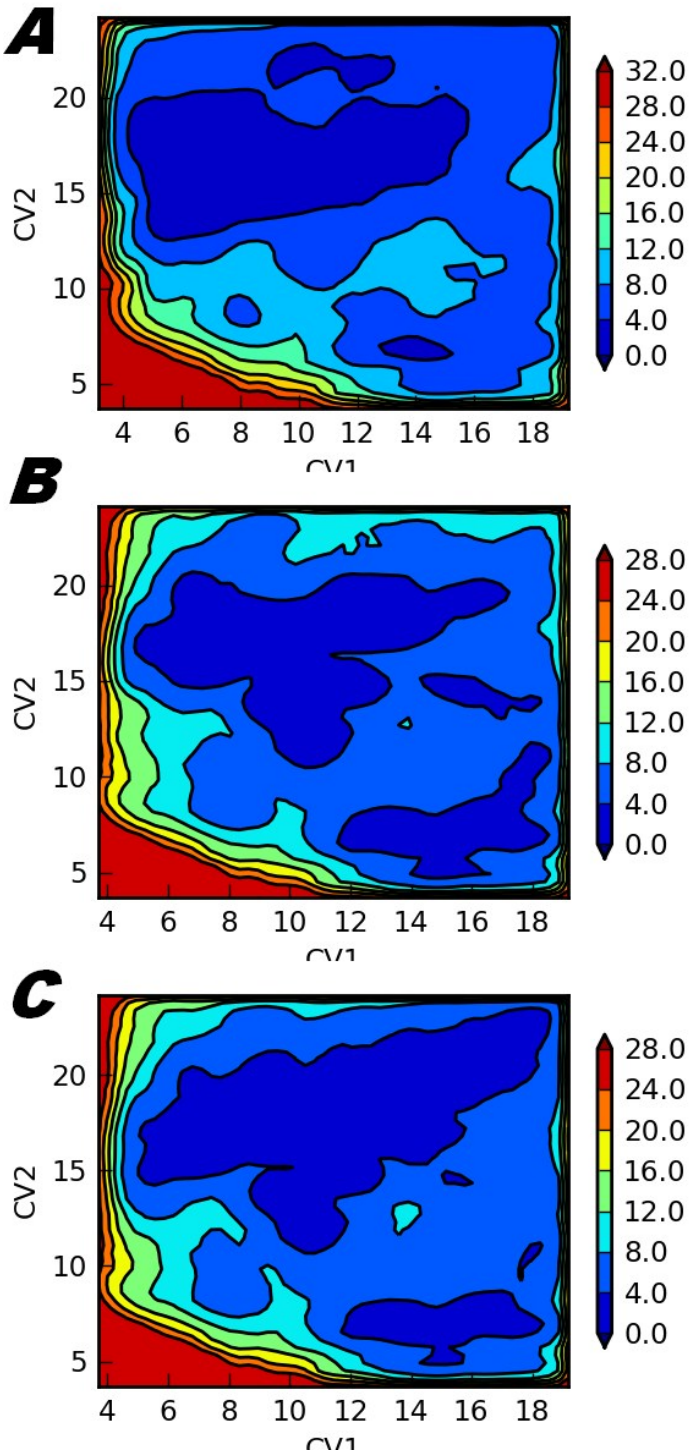

Figure S11. Free energy plots obtained by well tempered metadynamics simulations in a monomer using distances 1 and 2 as CVs after ***(A)*** 40 ns ***(B)*** 70 ns and ***(C)*** 100 ns of simulation time, respectively. Note that positions and values of the energy minimums are similar after 70 ns of simulation time.

Video S1. Detected mutable binding modes of the studied ligand during the aMD simulations.

Video S2. Detected an antagonist conformation of activation helix, helix 12, in PPARγ. Note the exchange of an agonist to antagonist conformation at about the 30th second.