Electronic Supplementary Material (ESI) for RSC Advances.

Supplementary material

Section S1. *Description of in-house Python script.*

The algorithm for selecting the representative protein-ligand complexes is as follows (flowchart in Figure S1): having protein conformations p_1, p_2, …, p_m, and ligands l_1, l_2, …, l_n, there are a total of $m \times n$ complexes, where $p_i l_j$ denotes ligand l_j docked into protein structure p_i. Subsets of the total m protein structures were defined with a reduced dimension k, where $k \leq m$. For each k value, $m!/(k!(m-k)!)$ subsets of k distinct conformations were generated. Subsets were tidily explored from the ones grouped in the smallest dimension ($k = 1$) to the longest one ($k = m$).

Each subset grouped in a dimension k has k complexes per ligand from which a single representative is chosen. This selection is performed using a naive approach. First, the scoring values computed either by docking or MM/GBSA of all $k \times n$ complexes in the subset are fitted to the experimental activities using the linear least squares (LLS) fitting technique. Next, the best $p_i l_j$ combinations are determined by optimizing the overall fitting (higher R^2) within the subset until all ligands are represented by a single complex. For the best combination, the higher R^2 is denoted as R^2_k. When $k = 1$, R^2_1 is defined as R^2_{max1}; when $k > 1$, R^2_{maxk} is $R^2_{max(k-1)}$ if $R^2_{max(k-1)} > R^2_k - 0.05$, otherwise, R^2_{maxk} is R^2_k (Figure S1). These conditions guarantee that higher values of R^2_{max} with small k dimension are kept; i.e., correlation models that contain few X-ray crystallographic structures are prioritized, and more X-ray crystals are considered if they considerably increase R^2_{max}.

Figure S1. Flowchart of *in-house* Python script

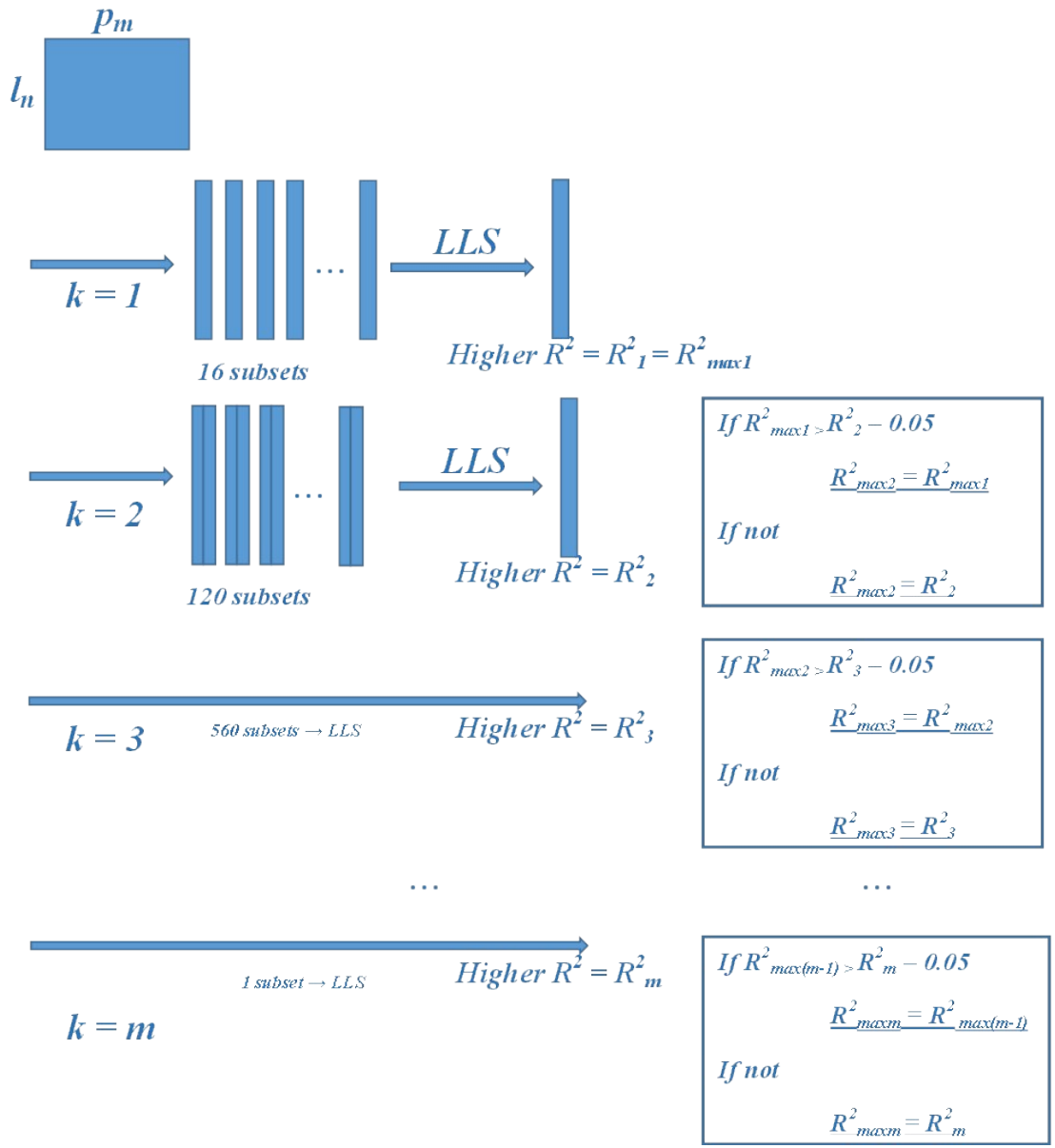
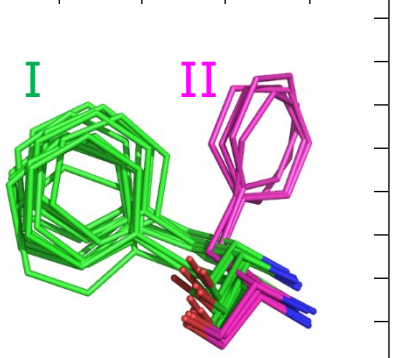


Figure S2. RMSD values (in Å) of the residues A) Phe282, B) Arg288, C) Phe363, D) Tyr473, for different PPARγ crystallographic structures. RMSD < 2.0 Å are represented in blue, RMSD ≥ 2.0 Å and < 3.0 Å are represented in yellow, and RMSD ≥ 3.0 Å are represented in red. Conformations of each residue are represented to the right, colored according to their orientation.

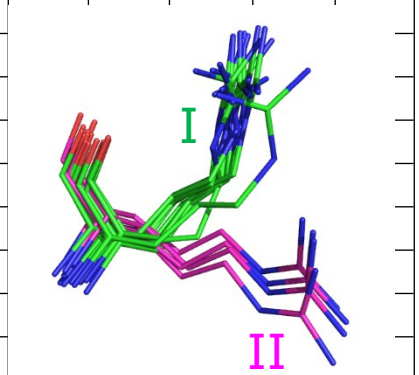
A (Phe282)

	2ATH	2F4B	2I4J	2PRG	2Q59	2VV0	2VV1	2VV2	2XKW	3B3K	3CDS	3GBK	3HO0	3HOD	3NOA	3QT0
2ATH	0.00															
2F4B	0.92	0.00														
2I4J	0.94	0.74	0.00													
2PRG	0.90	0.31	0.75	0.00												
2Q59	1.10	0.72	0.35	0.78	0.00											
2VV0	0.88	0.82	0.35	0.88	0.51	0.00										
2VV1	1.28	0.68	0.77	0.76	0.83	0.77	0.00									
2VV2	0.75	1.05	0.78	1.04	0.81	0.60	1.20	0.00								
2XKW	0.33	0.95	0.81	0.92	0.95	0.71	1.24	0.46	0.00							
3B3K	4.11	4.59	4.25	4.54	4.35	4.10	4.55	3.83	3.95	0.00						
3CDS	3.92	4.45	4.08	4.40	4.16	3.93	4.45	3.60	3.74	0.67	0.00					
3GBK	0.55	0.73	0.70	0.67	0.73	0.69	1.11	0.50	0.41	4.14	3.93	0.00				
3HO0	3.96	4.47	4.09	4.40	4.17	3.94	4.43	3.62	3.78	0.68	0.45	3.95	0.00			
3HOD	3.84	4.36	3.99	4.29	4.07	3.84	4.34	3.51	3.66	0.78	0.53	3.83	0.19	0.00		
3NOA	0.43	0.66	0.99	0.67	1.08	0.98	1.16	0.93	0.62	4.41	4.24	0.60	4.27	4.15	0.00	
3QT0	0.76	0.65	1.11	0.72	1.18	1.05	0.98	1.12	0.92	4.56	4.42	0.86	4.43	4.31	0.46	0.00



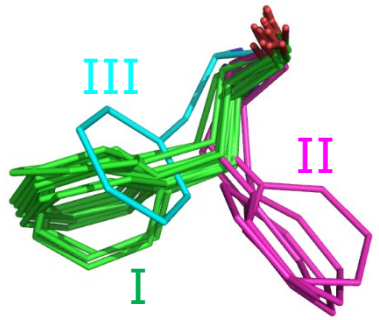
B (Arg288)

	2ATH	2F4B	2I4J	2PRG	2Q59	2VV0	2VV1	2VV2	2XKW	3B3K	3CDS	3GBK	3HO0	3HOD	3NOA	3QT0
2ATH	0.00															
2F4B	0.40	0.00														
2I4J	0.81	0.98	0.00													
2PRG	0.74	0.55	0.96	0.00												
2Q59	1.69	1.69	1.41	1.43	0.00											
2VV0	1.57	1.60	1.48	1.64	1.74	0.00										
2VV1	0.20	0.28	0.85	0.65	1.65	1.56	0.00									
2VV2	0.33	0.50	0.65	0.78	1.66	1.56	0.39	0.00								
2XKW	0.81	0.83	0.63	0.99	1.63	1.57	0.82	0.52	0.00							
3B3K	3.66	3.67	3.66	3.61	3.15	3.25	3.62	3.67	3.71	0.00						
3CDS	4.03	4.04	4.04	3.94	3.41	3.61	3.99	4.08	4.12	0.61	0.00					
3GBK	0.70	0.42	1.00	0.64	1.74	1.67	0.61	0.61	0.67	3.80	4.18	0.00				
3HO0	3.88	3.90	3.88	3.81	3.29	3.45	3.85	3.92	3.96	0.46	0.31	4.04	0.00			
3HOD	4.32	4.34	4.36	4.28	3.77	3.93	4.28	4.38	4.44	0.91	0.51	4.50	0.62	0.00		
3NOA	0.28	0.55	0.60	0.79	1.59	1.52	0.36	0.22	0.62	3.61	4.00	0.73	3.85	4.29	0.00	
3QT0	1.01	0.91	1.27	1.02	1.45	1.66	0.97	1.06	1.20	3.72	4.07	0.96	3.92	4.37	1.07	0.00



C (Phe363)

	2ATH	2F4B	2I4J	2PRG	2Q59	2VV0	2VV1	2VV2	2XKW	3B3K	3CDS	3GBK	3HO0	3HOD	3NOA	3QT0
2ATH	0.00															
2F4B	1.38	0.00														
2I4J	1.46	0.39	0.00													
2PRG	4.58	4.57	4.53	0.00												
2Q59	1.89	0.88	0.81	4.30	0.00											
2VV0	1.95	0.90	0.74	4.37	0.45	0.00										
2VV1	1.72	0.78	0.76	4.24	0.29	0.59	0.00									
2VV2	3.93	4.23	4.16	1.74	4.11	4.18	4.00	0.00								
2XKW	4.57	4.54	4.48	0.45	4.25	4.29	4.21	1.79	0.00							
3B3K	1.32	0.45	0.37	4.39	0.95	0.91	0.80	3.95	4.36	0.00						
3CDS	1.30	0.24	0.41	4.59	0.93	0.96	0.80	4.22	4.57	0.40	0.00					
3GBK	1.23	0.33	0.66	4.64	1.01	1.09	0.89	4.28	4.63	0.65	0.36	0.00				
3HO0	1.61	0.33	0.39	4.54	0.80	0.75	0.77	4.26	4.51	0.53	0.39	0.56	0.00			
3HOD	1.44	0.27	0.36	4.56	0.86	0.81	0.79	4.23	4.53	0.44	0.27	0.43	0.21	0.00		
3NOA	2.28	2.49	2.57	3.72	2.27	2.49	2.14	3.41	3.73	2.47	2.46	2.38	2.57	2.51	0.00	
3QT0	4.59	4.60	4.57	0.87	4.34	4.39	4.26	1.98	1.08	4.41	4.61	4.66	4.58	4.59	3.78	0.00



D (Tyr473)

	2ATH	2F4B	2I4J	2PRG	2Q59	2VV0	2VV1	2VV2	2XKW	3B3K	3CDS	3GBK	3HO0	3HOD	3NOA	3QT0
2ATH	0.00															
2F4B	1.85	0.00														
2I4J	1.70	1.22	0.00													
2PRG	2.07	0.54	1.53	0.00												
2Q59	1.96	0.94	0.96	1.26	0.00											
2VV0	1.54	1.31	0.87	1.69	0.73	0.00										
2VV1	1.52	1.27	1.00	1.64	0.70	0.20	0.00									
2VV2	1.44	1.32	1.02	1.67	0.83	0.34	0.29	0.00								
2XKW	1.55	1.33	1.51	1.63	1.01	0.89	0.71	0.74	0.00							
3B3K	1.72	1.21	0.51	1.51	0.67	0.72	0.80	0.88	1.29	0.00						
3CDS	1.93	1.39	0.43	1.68	0.86	0.91	1.02	1.07	1.55	0.35	0.00					
3GBK	1.77	0.35	0.94	0.79	0.83	1.10	1.11	1.17	1.33	1.01	1.15	0.00				
3HO0	1.41	1.29	0.47	1.57	0.99	0.78	0.87	0.82	1.35	0.52	0.61	1.07	0.00			
3HOD	1.57	1.33	0.44	1.66	0.84	0.62	0.72	0.73	1.24	0.39	0.44	1.10	0.36	0.00		
3NOA	1.92	2.57	2.25	2.76	2.49	2.12	2.14	2.10	2.24	2.28	2.40	2.44	2.11	2.20	0.00	
3QT0	2.21	1.54	2.50	1.57	1.94	2.16	2.00	2.05	1.50	2.34	2.61	1.81	2.39	2.40	3.02	0.00

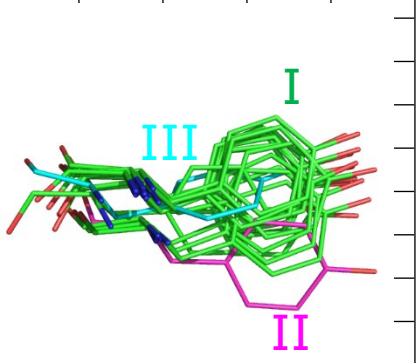


Figure S3. RMSD values (in Å) of the residues Phe282 (top left), Arg288 (top right), Phe363 (bottom left), Tyr473 (bottom right), for different PPARγ crystallographic structures involved in the best correlation models for set 1 (A), set 2 (B), and set 3 (C). RMSD < 2.0 Å are represented in blue, RMSD ≥ 2.0 Å and < 3.0 Å are represented in yellow, and RMSD ≥ 3.0 Å are represented in red.

A) (Set 1, Glide XP)

Phe282

	2F4B	2I4J	2PRG	2Q59	2VV0	2VV1	2XKW	3HOD	3NOA
2F4B	0.00								
2I4J	0.74	0.00							
2PRG	0.31	0.75	0.00						
2Q59	0.72	0.35	0.78	0.00					
2VV0	0.82	0.35	0.88	0.51	0.00				
2VV1	0.68	0.77	0.76	0.83	0.77	0.00			
2XKW	0.95	0.81	0.92	0.95	0.71	1.24	0.00		
3HOD	4.36	3.99	4.29	4.07	3.84	4.34	3.66	0.00	
3NOA	0.66	0.99	0.67	1.08	0.98	1.16	0.62	4.15	0.00

Arg288

	2F4B	2I4J	2PRG	2Q59	2VV0	2VV1	2XKW	3HOD	3NOA
2F4B	0.00								
2I4J	0.98	0.00							
2PRG	0.55	0.96	0.00						
2Q59	1.69	1.41	1.43	0.00					
2VV0	1.60	1.48	1.64	1.74	0.00				
2VV1	0.28	0.85	0.65	1.65	1.56	0.00			
2XKW	0.83	0.63	0.99	1.63	1.57	0.82	0.00		
3HOD	4.34	4.36	4.28	3.77	3.93	4.28	4.44	0.00	
3NOA	0.55	0.60	0.79	1.59	1.52	0.36	0.62	4.29	0.00

Phe363

	2F4B	2I4J	2PRG	2Q59	2VV0	2VV1	2XKW	3HOD	3NOA
2F4B	0.00								
2I4J	0.39	0.00							
2PRG	4.57	4.53	0.00						
2Q59	0.88	0.81	4.30	0.00					
2VV0	0.90	0.74	4.37	0.45	0.00				
2VV1	0.78	0.76	4.24	0.29	0.59	0.00			
2XKW	4.54	4.48	0.45	4.25	4.29	4.21	0.00		
3HOD	0.27	0.36	4.56	0.86	0.81	0.79	4.53	0.00	
3NOA	2.49	2.57	3.72	2.27	2.49	2.14	3.73	2.51	0.00

Tyr473

	2F4B	2I4J	2PRG	2Q59	2VV0	2VV1	2XKW	3HOD	3NOA
2F4B	0.00								
2I4J	1.22	0.00							
2PRG	0.54	1.53	0.00						
2Q59	0.94	0.96	1.26	0.00					
2VV0	1.31	0.87	1.69	0.73	0.00				
2VV1	1.27	1.00	1.64	0.70	0.20	0.00			
2XKW	1.33	1.51	1.63	1.01	0.89	0.71	0.00		
3HOD	1.33	0.44	1.66	0.84	0.62	0.72	1.24	0.00	
3NOA	2.57	2.25	2.76	2.49	2.12	2.14	2.24	2.20	0.00

B) (Set2, MMGBSA XP)

Phe282

	2I4J	2PRG	2Q59	2VV1	2VV2	3HOD	3NOA
2I4J	0.00						
2PRG	0.75	0.00					
2Q59	0.35	0.78	0.00				
2VV1	0.77	0.76	0.83	0.00			
2VV2	0.78	1.04	0.81	1.20	0.00		
3HOD	3.99	4.29	4.07	4.34	3.51	0.00	
3NOA	0.99	0.67	1.08	1.16	0.93	4.15	0.00

Arg288

	2I4J	2PRG	2Q59	2VV1	2VV2	3HOD	3NOA
2I4J	0.00						
2PRG	0.96	0.00					
2Q59	1.41	1.43	0.00				
2VV1	0.85	0.65	1.65	0.00			
2VV2	0.65	0.78	1.66	0.39	0.00		
3HOD	4.36	4.28	3.77	4.28	4.38	0.00	
3NOA	0.60	0.79	1.59	0.36	0.22	4.29	0.00

Phe363

	2I4J	2PRG	2Q59	2VV1	2VV2	3HOD	3NOA
2I4J	0.00						
2PRG	4.53	0.00					
2Q59	0.81	4.30	0.00				
2VV1	0.76	4.24	0.29	0.00			
2VV2	4.16	1.74	4.11	4.00	0.00		
3HOD	0.36	4.56	0.86	0.79	4.23	0.00	
3NOA	2.57	3.72	2.27	2.14	3.41	2.51	0.00

Tyr473

	2I4J	2PRG	2Q59	2VV1	2VV2	3HOD	3NOA
2I4J	0.00						
2PRG	1.53	0.00					
2Q59	0.96	1.26	0.00				
2VV1	1.00	1.64	0.70	0.00			
2VV2	1.02	1.67	0.83	0.29	0.00		
3HOD	0.44	1.66	0.84	0.72	0.73	0.00	
3NOA	2.25	2.76	2.49	2.14	2.10	2.20	0.00

C) (Set3, Glide SP)

Phe282

	2ATH	2PRG	2Q59	2VV0	2VV1	3B3K	3CDS	3QT0
2ATH	0.00							
2PRG	0.90	0.00						
2Q59	1.10	0.78	0.00					
2VV0	0.88	0.88	0.51	0.00				
2VV1	1.28	0.76	0.83	0.77	0.00			
3B3K	4.11	4.54	4.35	4.10	4.55	0.00		
3CDS	3.92	4.40	4.16	3.93	4.45	0.67	0.00	
3QT0	0.76	0.72	1.18	1.05	0.98	4.56	4.42	0.00

Arg288

	2ATH	2PRG	2Q59	2VV0	2VV1	3B3K	3CDS	3QT0
2ATH	0.00							
2PRG	0.74	0.00						
2Q59	1.69	1.43	0.00					
2VV0	1.57	1.64	1.74	0.00				
2VV1	0.20	0.65	1.65	1.56	0.00			
3B3K	3.66	3.61	3.15	3.25	3.62	0.00		
3CDS	4.03	3.94	3.41	3.61	3.99	0.61	0.00	
3QT0	1.01	1.02	1.45	1.66	0.97	3.72	4.07	0.00

Phe363

	2ATH	2PRG	2Q59	2VV0	2VV1	3B3K	3CDS	3QT0
2ATH	0.00							
2PRG	4.58	0.00						
2Q59	1.89	4.30	0.00					
2VV0	1.95	4.37	0.45	0.00				
2VV1	1.72	4.24	0.29	0.59	0.00			
3B3K	1.32	4.39	0.95	0.91	0.80	0.00		
3CDS	1.30	4.59	0.93	0.96	0.80	0.40	0.00	
3QT0	4.59	0.87	4.34	4.39	4.26	4.41	4.61	0.00

Tyr473

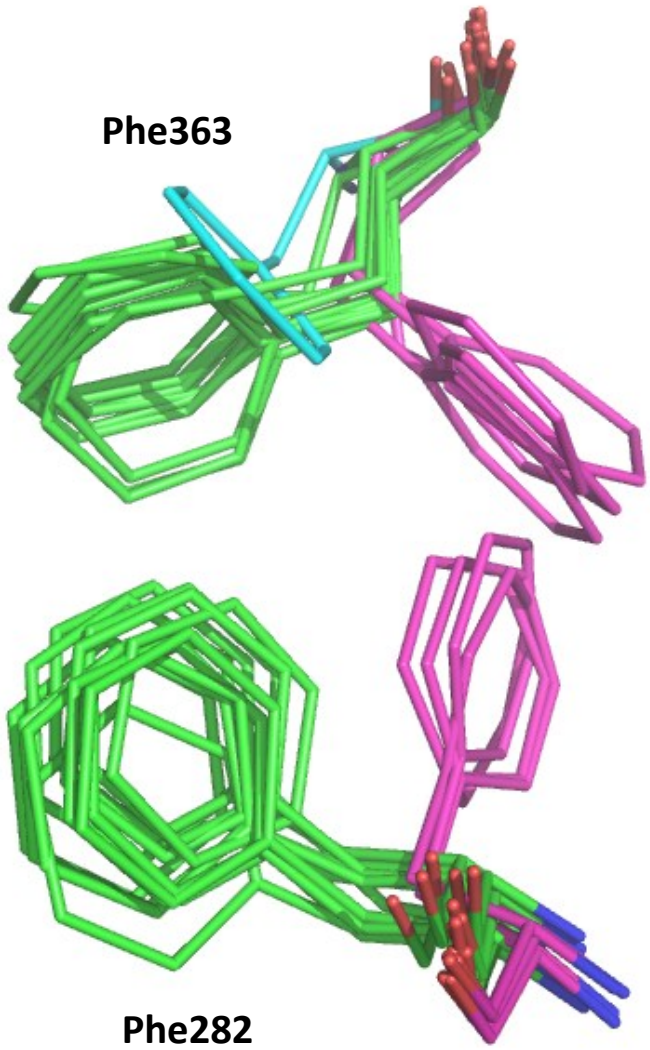
	2ATH	2PRG	2Q59	2VV0	2VV1	3B3K	3CDS	3QT0
2ATH	0.00							
2PRG	2.07	0.00						
2Q59	1.96	1.26	0.00					
2VV0	1.54	1.69	0.73	0.00				
2VV1	1.52	1.64	0.70	0.20	0.00			
3B3K	1.72	1.51	0.67	0.72	0.80	0.00		
3CDS	1.93	1.68	0.86	0.91	1.02	0.35	0.00	
3QT0	2.21	1.57	1.94	2.16	2.00	2.34	2.61	0.00

Figure S4. (A) RMSD values (in Å) for Phe282- Phe363 pair, for different PPARγ crystallographic structures. RMSD ≤ 1.61 Å are represented in light blue, RMSD ≥ 1.64 Å and ≤ 1.95 Å are represented in dark blue, RMSD ≥ 2.5 Å and ≤ 2.73 Å are represented in yellow, RMSD ≥ 2.77 Å and ≤ 3.35 Å are represented in red, RMSD ≥ 3.43 Å and ≤ 3.58 Å are represented in pink, and RMSD ≥ 3.89 Å are represented in purple. (B) Examples of conformational pairs compared in A observed for each defined color.

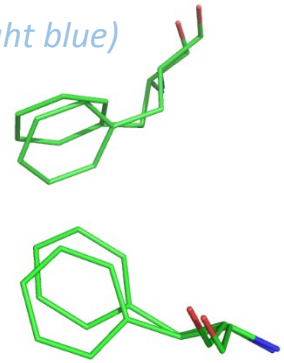
A)

	2ATH	2F4B	2I4J	2PRG	2Q59	2VV0	2VV1	2VV2	2XKW	3B3K	3CDS	3GBK	3HO0	3HOD	3NOA	3QT0
2ATH	0.00															
2F4B	1.17	0.00														
2I4J	1.23	0.59	0.00													
2PRG	3.30	3.24	3.25	0.00												
2Q59	1.54	0.80	0.62	3.09	0.00											
2VV0	1.51	0.86	0.58	3.15	0.48	0.00										
2VV1	1.52	0.73	0.77	3.05	0.62	0.69	0.00									
2VV2	2.83	3.08	2.99	1.43	2.96	2.99	2.95	0.00								
2XKW	3.24	3.28	3.22	0.72	3.08	3.08	3.10	1.31	0.00							
3B3K	3.05	3.26	3.02	4.46	3.14	2.97	3.27	3.89	4.16	0.00						
3CDS	2.92	3.15	2.90	4.50	3.01	2.86	3.20	3.92	4.18	0.56	0.00					
3GBK	0.95	0.57	0.68	3.32	0.88	0.91	1.00	3.05	3.29	2.96	2.79	0.00				
3HO0	3.02	3.17	2.90	4.47	3.00	2.84	3.18	3.95	4.16	0.61	0.42	2.82	0.00			
3HOD	2.90	3.09	2.83	4.43	2.94	2.77	3.12	3.89	4.12	0.63	0.42	2.73	0.20	0.00		
3NOA	1.64	1.82	1.95	2.68	1.78	1.89	1.72	2.50	2.67	3.58	3.47	1.74	3.52	3.43	0.00	
3QT0	3.29	3.28	3.33	0.80	3.18	3.19	3.09	1.61	1.01	4.48	4.52	3.35	4.51	4.45	2.69	0.00

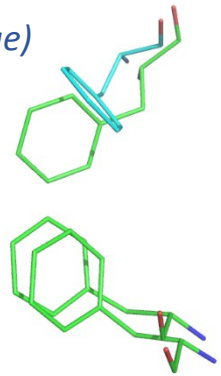
B)



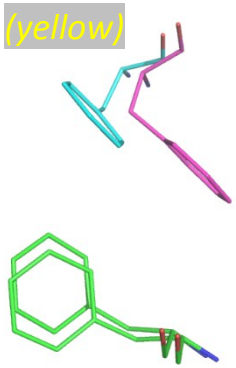
(light blue)



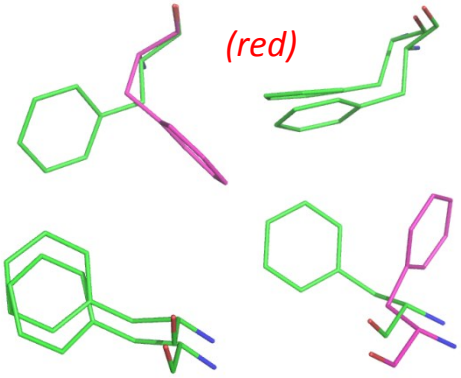
(dark blue)



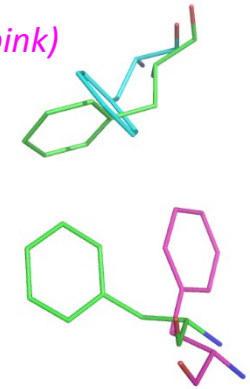
(yellow)



(red)



(pink)



(purple)

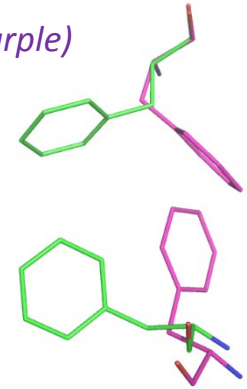


Figure S5. RMSD values (in Å) of the Phe282- Phe363 pair, for different PPARγ crystallographic structures involved in the best correlation models for set 1 (A), set 2 (B), and set 3 (C). RMSD ≤ 1.61 Å are represented in light blue, RMSD ≥ 1.64 Å and ≤ 1.95 Å are represented in dark blue, RMSD ≥ 2.5 Å and ≤ 2.73 Å are represented in yellow, RMSD ≥ 2.77 Å and ≤ 3.35 Å are represented in red, RMSD ≥ 3.43 Å and ≤ 3.58 Å are represented in pink, and RMSD ≥ 3.89 Å are represented in purple.

A) (Set 1, MMGBSA XP)

	2F4B	2I4J	2PRG	2Q59	2VV0	2VV1	2XKW	3HOD	3NOA
2F4B	0.00								
2I4J	0.59	0.00							
2PRG	3.24	3.25	0.00						
2Q59	0.80	0.62	3.09	0.00					
2VV0	0.86	0.58	3.15	0.48	0.00				
2VV1	0.73	0.77	3.05	0.62	0.69	0.00			
2XKW	3.28	3.22	0.72	3.08	3.08	3.10	0.00		
3HOD	3.09	2.83	4.43	2.94	2.77	3.12	4.12	0.00	
3NOA	1.82	1.95	2.68	1.78	1.89	1.72	2.67	3.43	0.00

B) (Set2, MMGBSA XP)

	2I4J	2PRG	2Q59	2VV1	2VV2	3HOD	3NOA
2I4J	0.00						
2PRG	3.25	0.00					
2Q59	0.62	3.09	0.00				
2VV1	0.77	3.05	0.62	0.00			
2VV2	2.99	1.43	2.96	2.95	0.00		
3HOD	2.83	4.43	2.94	3.12	3.89	0.00	
3NOA	1.95	2.68	1.78	1.72	2.50	3.43	0.00

C) (Set3, Glide SP)

	2ATH	2PRG	2Q59	2VV0	2VV1	3B3K	3CDS	3QT0
2ATH	0.00							
2PRG	3.30	0.00						
2Q59	1.54	3.09	0.00					
2VV0	1.51	3.15	0.48	0.00				
2VV1	1.52	3.05	0.62	0.69	0.00			
3B3K	3.05	4.46	3.14	2.97	3.27	0.00		
3CDS	2.92	4.50	3.01	2.86	3.20	0.56	0.00	
3QT0	3.29	0.80	3.18	3.19	3.09	4.48	4.52	0.00