Graduate Student Theses, Dissertations, & Professional Papers

Graduate School

2018

EXPLORING THE LIGAND-DEPENDENT CONFORMATIONAL ENSEMBLE OF PPARγ: MORE THAN A MOUSETRAP

Ian Michael Chrisman

Let us know how access to this document benefits you.

Follow this and additional works at: https://scholarworks.umt.edu/etd

EXPLORING THE LIGAND-DEPENDENT CONFORMATIONAL

ENSEMBLE OF PPARγ: MORE THAN A MOUSETRAP

By

Ian Michael Chrisman

Bachelor of Science in Biochemistry, University of Montana, 2012

Dissertation presented in partial fulfillment of the requirements

for Doctor of Philosophy in Biochemistry

The University of Montana

Missoula, MT

December 2018

Approved by

Scott Whittenburg, Dean of the Graduate School
Graduate School

Travis Hughes, Advisor
Department of Biomedical and Pharmaceutical Sciences

Bruce Bowler, Committee Chair
Department of Chemistry and Biochemistry

Stephen Sprang
Division of Biological Science

Kent Sugden
Department of Chemistry and Biochemistry

Xi Chu
Department of Chemistry and Biochemistry

i

Chrisman, Ian, M, PhD, Fall 2018
Biochemistry

Travis Hughes, Advisor

Abstract

The nuclear receptor PPARγ is a lipid-dependent transcription factor which regulates many pathways in lipid metabolism. Due to functions in adipogenesis, glucose uptake and insulin sensitivity it has been a target of type II diabetes mellitus therapeutics since the 1990s. Due to a poorly understood anti-inflammatory pathway it is also an emerging target in neurodegenerative disease. Despite extensive research the majority of the structural mechanism underlying the regulation of this proteins' function is poorly understood. It is often thought that nuclear receptors function in a simple two-state model with defined on and off states. This model only considers one region of the protein, the coregulator interaction surface. In this mechanism ligands function only to modulate the equilibrium population of the active and inactive states. In this work we examine the conformational ensemble of PPARγ in response to the binding of a wide variety of structurally and functionally diverse ligands and find clear evidence that this model is only correct for the most efficacious activating ligand. In all others the ligand induced conformational ensemble is highly diverse and contains multiple populations. Additionally, the mechanisms of inverse agonism and agonism are examined. Through this a novel mechanism of inverse agonism is determined. As well we demonstrate that the mechanistic underpinnings of agonism are not strictly through the coregulator interaction surface. The most efficacious ligands also make important interactions in other regions of the protein which are often disregarded. Among the most interesting findings is the identification of a highly conserved motif on helix 4 of the ligand binding domain of nuclear receptors. We demonstrate that this motif is responsible for a large degree of the specificity observed in coactivator interactions. This provides a substantial amount of insight into the mechanism by which these proteins function as it had previously been observed that coactivator recruitment was exceptionally specific, but no mechanistic reason was known. Through this work we have gleaned a deeper understanding of the mechanism of specific ligand-dependent conformations in this important protein. This fundamental understanding allows for more efficient and intelligent design of drugs which target PPARγ and other nuclear receptors.

Acknowledgments

I would like to thank everyone who helped me through this long and somewhat convoluted path. Particularly, my advisor, Travis Hughes, who gave me a chance after my previous advisor got dismissed and guided me towards productivity in an area in which I had limited experience. I would also like to thank my committee for their feedback and guidance through this experience. I would especially like to acknowledge my lab mates. Without their important contributions none of this work would have been possible. Their assistance in experimentation and trouble shooting was invaluable throughout my time here.

Most importantly I would like to acknowledge my family and my fiancée. Without their constant support it would have been difficult to navigate this long and arduous path.

Contents

Chapter 1 Introduction

1.1 Nuclear Receptors

The nuclear receptors (NR) are a diverse family of transcription factors that regulate a wide variety of signaling pathways. In the human genome there are 48 recognized NR proteins in this superfamily which make up the largest group of known transcription factors[1,2]. These proteins are highly sequence divergent but structurally conserved[3]. All members of this superfamily share a very modular fold wherein they contain an N-terminal intrinsically disordered region known as either the A/B domain or the activation factor 1 region (AF-1). This domain has very low homology between nuclear receptors and like many intrinsically disordered regions appears to function primarily in the recruitment of regulatory proteins. Though there is an emerging role for this domain in the modulation of drug response. Thus far this has only been reported with estrogen receptor alpha (ERα)[4,5] but this may also be true for other members of this superfamily. The next domain is the C domain or DNA-binding domain (DBD). This approximately 80 residue domain is highly conserved between nuclear receptors, ~80%, and contains two tetra-cysteine zinc fingers. This domain mediates the interaction with the respective response element for each of the NR. Receptors will bind to their response element as either heterodimers, most often dimerized with retinoid x receptors (RxR), homodimers or as half-site binders. This leads to highly divergent sequences for the response elements, but they are most often direct repeats for heterodimeric binding proteins or an inverted repeat if the receptor binds as a homodimer. The next domain, D, is the flexible hinge region of approximately 20 residues which connects the DBD to the following domain and allows conformational freedom between the domains of the protein. For many nuclear receptors the E domain or ligand-binding domain (LBD) is the C-terminus of the protein. This domain has a very conserved fold made up of approximately 12 α-helices and a single region which typically contains three β-sheets. Not all nuclear receptor LBDs have exactly 12 helices, but they are numbered as if they do. The LBD contains the ligand binding pocket as well as the major coregulator

interaction surface known as the activation factor 2 (AF-2). In some receptors, primarily the hormone receptors, there is an additional F domain which is moderately structured and makes up the C-terminus of the protein. The role of the F domain is not known but has been shown in some receptors such as ERα or glucocorticoid receptor (GR) to be essential for protein stability. Recent data in ER α and β indicates this domain may also be required for agonist behavior of certain ligands[6,7].
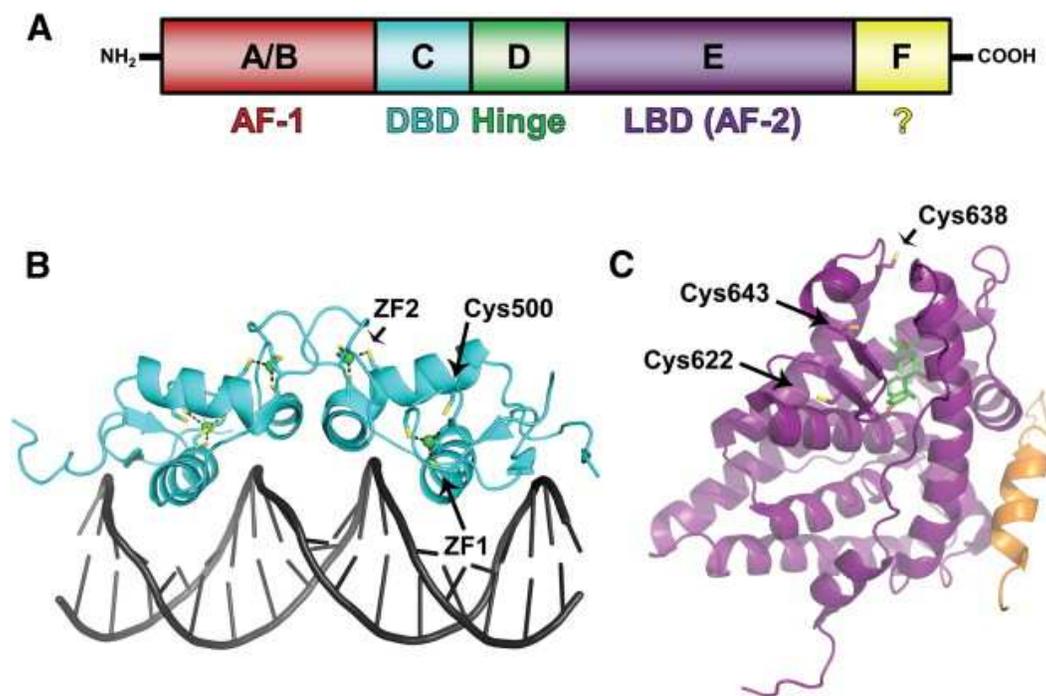
Figure 1. Conserved structure of nuclear receptors. A: Primary structure of nuclear receptors showing the four conserved domains seen in nuclear receptors. Domain F is unique to hormone receptors and is not present in the majority of nuclear receptors. B: Crystal structure of the DBD of GR bound to a response element of DNA (PDB 1GLU). C: Crystal structure of the LBD of GR bound to a peptide of the coactivator TIF2 (PDB 1M2Z). Both the DBD and LBD have a conserved fold in nuclear receptors, so the representative structures shown will be very similar to those of other nuclear receptors. Figure was adapted from Carter et al[8].

These receptors are primarily divided into three families (NR1, NR2 and NR3) which respond to different types of ligands. The NR1 family is primarily lipid-dependent and binds either fatty acids or cholesterol metabolites. The NR2

family is by majority vitamin-dependent and binds a variety of essential vitamins such as retinoic acid (vitamin A). There are exceptions to this however, as the vitamin D receptor (VDR) is a member of the NR1 family while the orphan receptor testicular receptor 4 (TR4) which has no known ligands specific to this protein but can bind peroxisome proliferator-activated receptor gamma (PPARγ) synthetic ligands is a member of NR2. The NR3 family is composed primarily of hormone receptors and include heavily studied receptors such as GR or ERα. There are three additional small families which are the NR0, NR4 and NR5 families. These are all orphan receptors, with no known ligands, and several, including all members of the NR4 family, are considered to be constitutively active and ligand-independent[9,10]. This is based on crystallographic observations that these proteins do not contain a true ligand binding pocket due to impingement by helix 3 and a large number of bulky hydrophobic side chains in the ligand binding pocket[9]. However, a recent paper examining the LBD of Nurr1, NR4A2, showed that the residues around what would be the ligand binding pocket are highly plastic[11]. This may indicate that what is observed in crystal structures is only one of several accessible conformations in this region and that in some accessible conformations there is a properly formed ligand-binding pocket. The NR0 family is also interesting, these proteins have lost the AF-1 and DBD through the evolutionary process and now function more in the manner of a coregulator. They still form heterodimers with other nuclear receptors but due to the lack of their own DBD they are almost exclusively transrepressive[12,13].

1.2 Mechanism of transcriptional regulation

Nuclear receptor function is regulated through the recruitment of coregulator complexes. These occur through direct interactions between the nuclear receptor and either a coactivator or a corepressor protein in a ligand-dependent manner. The binding of a coregulator to a NR will then nucleate the formation of a large complex which regulates the transcription of downstream genes in a variety of ways. It is generally thought that in an apo state of the nuclear receptor it is almost exclusively bound with corepressors and when bound

to an endogenous ligand or a synthetic agonist, corepressors are displaced and coactivators are recruited. These interactions are primarily mediated by small motifs on the coregulator protein. In a coactivator this interaction motif is comprised of an LxxLL sequence known as an NR box[14–16] (In this work referred to as an LxxLL box to avoid confusion with the abbreviation for nuclear receptor). For a corepressor this motif is a somewhat longer and there is less clarity in what makes up the sequence, but it is postulated that this motif contains at least I/LxxI/VIxxxF/Y/L, LxxxL/IxxxL/I or L/VxxI/VI and is referred to as a CoRNR box[17–20]. The LxxLL or CoRNR box interacts most often on the AF-2 surface of the nuclear receptor in a conserved manner where the interaction is mediated primarily by a single α-helix of the coregulator protein. These interactions are exceptionally specific with each NR binding only certain coregulators and often only specific LxxLL or CoRNR boxes within a coregulator. Currently, little is known of how these interactions are specific but in chapter three of this work we present evidence for a helix 4 motif on the AF-2 surface of nuclear receptors which appears to be responsible for the majority of the specificity observed in coactivator binding.

What is known about coregulator recruitment is almost completely in the context of coactivators. It has been shown that the AF-2 surface of a nuclear receptor has a hydrophobic cleft which allows a docking event between the leucine residues in the LxxLL box and the nuclear receptor[21]. Following this docking the most common hypothesis for interaction is from the "mousetrap model" of the charge clamp. The charge clamp is made up of a glutamate on helix 12 and a lysine on the C-terminus of helix 3 which are conserved in almost all nuclear receptors[22]. In this model helix 12 relocates into a conformation where it is more proximal to the coactivator LxxLL box. This moves the coactivator into a conformation where it is located near both charge clamp residues and allows for the formation of extensive hydrogen bonds between the charge clamp residues and the peptide backbone of the LxxLL box. These interactions have been shown experimentally to convey the bulk of the binding energy for the coregulator but as

stated earlier confer no selectivity since they are to a conserved region of the peptide backbone of the coactivator.

Additionally, there are several nuclear receptors which have been shown to interact with coregulators through the intrinsically disordered AF-1 region[23,24]. The nature of this interaction is not well understood. The hypothesis for how this binding leads to a functional output is by driving a self-association between the AF-1 and the LBD which is required for activity, but why this is required for activation is not currently clear. This often leads to coregulators binding to both the AF-1 and AF-2 surface simultaneously[25,26]. It has been demonstrated that these spanning interactions are important for function but the structural reason why remains unclear. In the case of a small number of nuclear receptors, the NR4 family and the testicular receptors (TR2 and 4)[27,28], the AF-2 surface is degenerate and unsuitable for coactivator binding, so all coactivator interactions are mediated purely through the AF-1. For the NR4 proteins this is primarily thought to be due to a reversal of the charge clamp residues. In all three members of this family the helix 12 charge clamp residue is instead a lysine and the helix 3 charge clamp residue is a glutamate. This is in contrast to all other nuclear receptors which have the conserved lysine on helix 3 and the glutamate on helix 12. It is not totally clear why this change prevents coactivator recruitment on the AF-2 as these residues are only involved in non-specific hydrogen bonds to the backbone of coactivators but it has been shown experimentally that only the AF-1 contributes to coactivator binding in Nur77 and that the presence of the LBD actually lowers coactivator affinity[29]. For the other nuclear receptors which do not bind coactivators through the AF-2, TR2 and TR4, the reason appears to be due to a change in the polarity of the normally hydrophobic cleft where the leucines in the LxxLL box interact. In these proteins this cleft is much shallower and more polar which creates an unfavorable environment for these hydrophobic interactions[27,28]. Interestingly, in at least Nurr1, NR4A2, it has been found that not only do coactivators not bind in the AF-2 region but corepressors may also bind in a novel position which is located in a cleft between helices 11 and 12[30].

1.2.1 Transcriptional activation through the recruitment of coactivator complexes

Coactivator proteins recruit large complexes which depending on the coactivator have distinct functionality and regulate different pathways[31]. The largest family of coactivator proteins is the p160 family[32]. This includes the most commonly studied coactivator proteins nuclear receptor coactivator 1 and 2 (NCoA 1 and 2), also commonly known as SRC 1 and 2, as well as several other proteins. These proteins function primarily through the formation of histone acetyltransferase (HAT) complexes which covalently modify histone proteins and lead to the release of DNA thereby making it more accessible for transcription[33]. Other common coactivators such as p300 and creb-binding protein (CBP) nucleate the formation of similar HAT complexes[34]. There is significant crosstalk between the NCoA proteins and CBP or p300 as it has been shown for each they are often in the complexes assembled by the other[35] leading to a synergistic activation. In addition, the yeast HAT protein GCN5[36], closely related to the human analog KAT2A, has been shown to directly bind some nuclear receptors in a drug dependent manner as a coactivator. This implies that the formation of large complexes may not always be necessary for coactivator function, but it should be stated that it is not clear that if these HAT proteins were directly recruited they would not nucleate larger complexes as well.

Not all coactivators are dependent on histone modification for function. Two examples of this are the mediator of RNA polymerase II transcription subunit 1 (MED1, also known as TRAP220 and DRIP205) and PPARγ coactivator 1α (PGC1α)[37–39]. With these coactivators the mechanism operates through a direct interaction with the promoter rather than through the recruitment of HAT proteins. Both of these proteins function in the same manner where upon binding of the nuclear receptor they assemble the 33-protein mediator complex. This complex will then interact directly with the promoter of the gene to be transcribed and facilitate the assembly of the RNA polymerase II holoenzyme.

There may also be other activity with two specific coactivator proteins, NCoA1[40] and MED1[41]. These have been shown in cell-based assays to increase gene expression when overexpressed. The mechanism of this is not fully clear as it occurs in the absence of stimulation of any ligand for receptors which interact with these proteins. The most likely hypothesis for this observation is that the higher concentration of overexpressed MED1 or NCoA1 led to more instances where the NR was bound to the coregulator even in the absence of ligand. This is consistent with the observation that NR have an innate low affinity for coactivators when the NR is in the apo state. However, this hypothesis has not been experimentally validated.
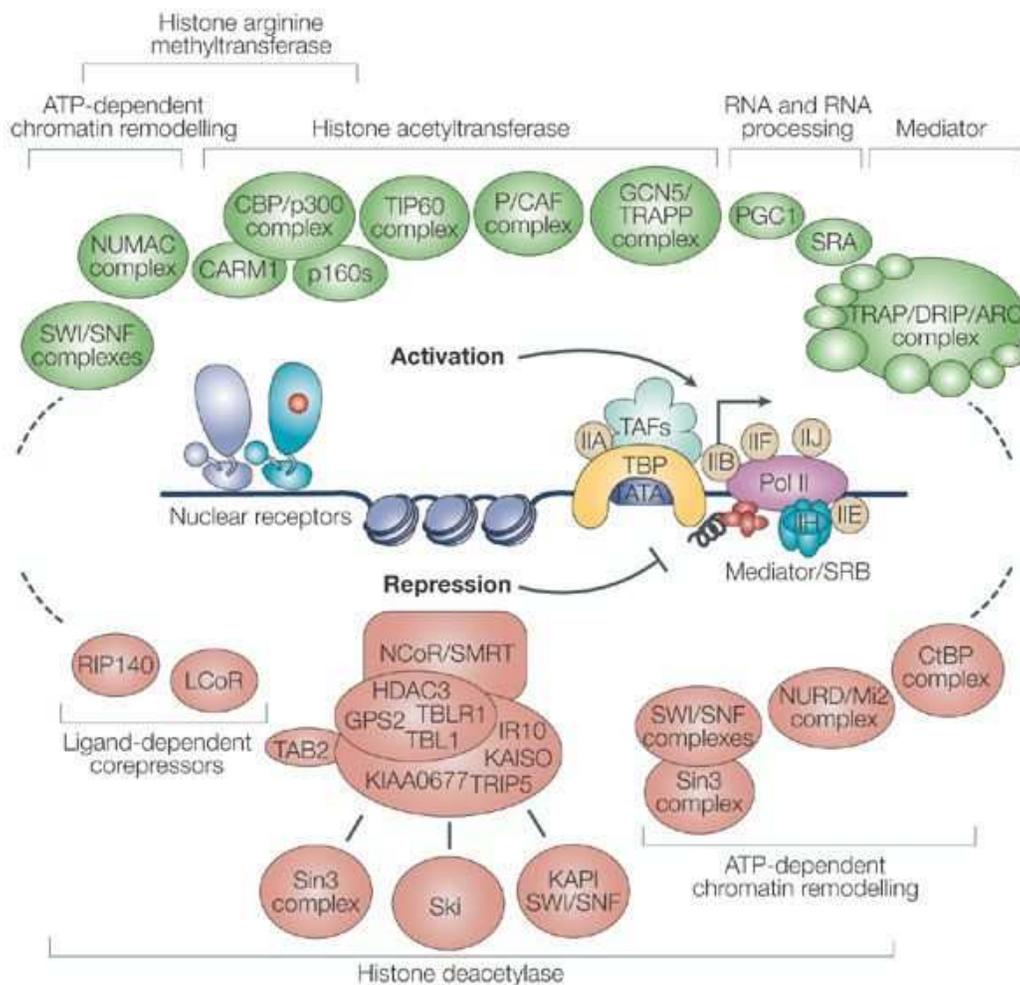
Figure 2. Nuclear receptor activity is mediated by coregulator recruitment. Nuclear receptor regulated transcription of downstream genes is mediated by the recruitment of coregulators. For most nuclear receptors depending on the ligand bound state either coactivator complexes (top) will be recruited to drive transcription. Alternatively, when bound to an inverse agonist or in the apo state there is a preference for corepressor binding which will downregulate the transcription of most downstream genes (bottom). Figure was adapted from Parissi et al[42].

1.2.2 Transcriptional repression and corepressor complexes

Nuclear receptor corepressors are a much less diverse group. In humans there are two primary corepressors, nuclear receptor corepressor 1 and 2 (NCoR 1 and 2, NCoR2 is often referred to as SMRT)[43,44]. These are both structurally similar large proteins, approximately 2500 residues, which are primarily intrinsically disordered. They share a mechanism of action whereby they nucleate the formation of complexes which contain one or more histone deacetylase (HDAC) proteins. The activity of HDAC will induce compression of genomic DNA and thereby reduce transcription of nearby genes. It is generally thought that the resting state of nuclear receptors in the absence of ligand is bound to one of these corepressors. Due to somewhat high structural similarity and a similar mechanisms of action it is often thought that these proteins are redundant and perform the same role. However, this seems unlikely mostly due to the high cost of producing such large and complex proteins. If these two NCoR proteins did not have discrete roles it is likely that one would have since been selected against and become a pseudogene. As well, there is some literature that suggests that one cannot compensate for the absence of the other. For instance, it was shown in mouse adipocytes that the selective ablation of NCoR1 created a phenotype very similar to that seen with treatment by a PPARγ activating drug[45]. If it were the case that NCoR1 and NCoR2 were redundant than it would be expected that the ablation would have had minimal effect due to compensation through NCoR2.

There is a third less well understood corepressor, nuclear receptor interacting protein 1 (NRIP1). This protein is particularly interesting as it does not contain a CoRNR box like the other corepressors but rather contains the LxxLL

box seen in coactivators[46,47]. It has been shown to bind in a ligand-dependent manner with a binding preference for agonist bound receptors. It has similar corepressor activity as is seen with NCoR1 and SMRT arising primarily through the assembly of corepressor complexes containing HDAC proteins. However, it also has been shown to have some coactivator characteristics. This mechanism is not well understood but it is thought that the regulation of NRIP1 activity as either a coactivator or corepressor is regulated through post-translational modification of the NRIP1 protein or by the sequence of the DNA the nuclear receptor is bound to. The presence of an LxxLL box instead of a CoRNR box would also suggest this protein would have a distinct binding mode from the NCoR proteins. This is due to the fact that the CoRNR box has a much larger footprint on the nuclear receptor and binding is thought to require either displacement of helix 12 or as has more recently been shown a relocation of helix 12[48]. This implies that despite primarily having corepressor activity this protein likely binds in a manner similar to coactivators.

1.3 Peroxisome proliferator-activated receptor gamma (PPARγ)

The work in this project focuses on the nuclear receptor PPARγ. This protein along with PPAR α and δ (PPARδ is also referred to as PPARβ) make up the NR1C family. All the PPAR proteins function as lipid-dependent transcription factors and regulate a variety of pathways related to fatty acid metabolism[49]. The PPAR proteins like other fatty acid receptors transactivate as obligate heterodimers with the retinoid x receptors (RxR α/β/γ) through the binding of direct repeat sequences of five nucleotides with a single nucleotide spacer, known as PPAR response elements (PPRE)[50]. It is generally thought the PPAR proteins have a preference to interact with RxRα but recent in cell data showed that PPARγ was actually most often interacting with RxRβ[51]. The PPAR proteins regulate a number of pathways related to lipid metabolism and differ primarily in tissue distribution. Particularly PPARα and PPARδ appear to have very similar functions but differ in tissue distribution with PPARα being most heavily expressed in the liver[52] and PPARδ having highest expression in skeletal

muscle[53]. Both these proteins are involved in signaling which leads to the metabolism of lipids as an energy source. Due to this there has been significant research on PPARα as a therapeutic target for dyslipidemia leading to the fibrate family of FDA approved drugs[54]. There has been significantly less research on PPARδ as a therapeutic target. Some experimental ligands have been created but none were suitable for human use[55,56].

PPARγ is expressed at low levels ubiquitously in almost all tissue types. There are two common isoforms which exist, PPARγ1 and PPARγ2, these isoforms differ only by the presence of an additional 28 residues on the N-terminus of the protein in PPARγ2, in mice PPARγ2 contains an additional 30 residues. For the purpose of clarity in the introduction all PPARγ residue numbers will be referred to in the form PPARγ1 #/ PPARγ2 # since both numbering systems are used extensively in literature. There are no currently known differences in the activity of the two isoforms of this protein. They do however differ significantly in tissue distribution. PPARγ1 is expressed at low levels in the majority of tissues while PPARγ2 is expressed almost completely in adipose tissue[57]. It has been shown that a high fat diet (HFD) in mice will lead to the expression of PPARγ2 in other tissues[58,59]. Depending on the tissue there appears to be distinct pathways for this protein which will be explained in greater detail below.
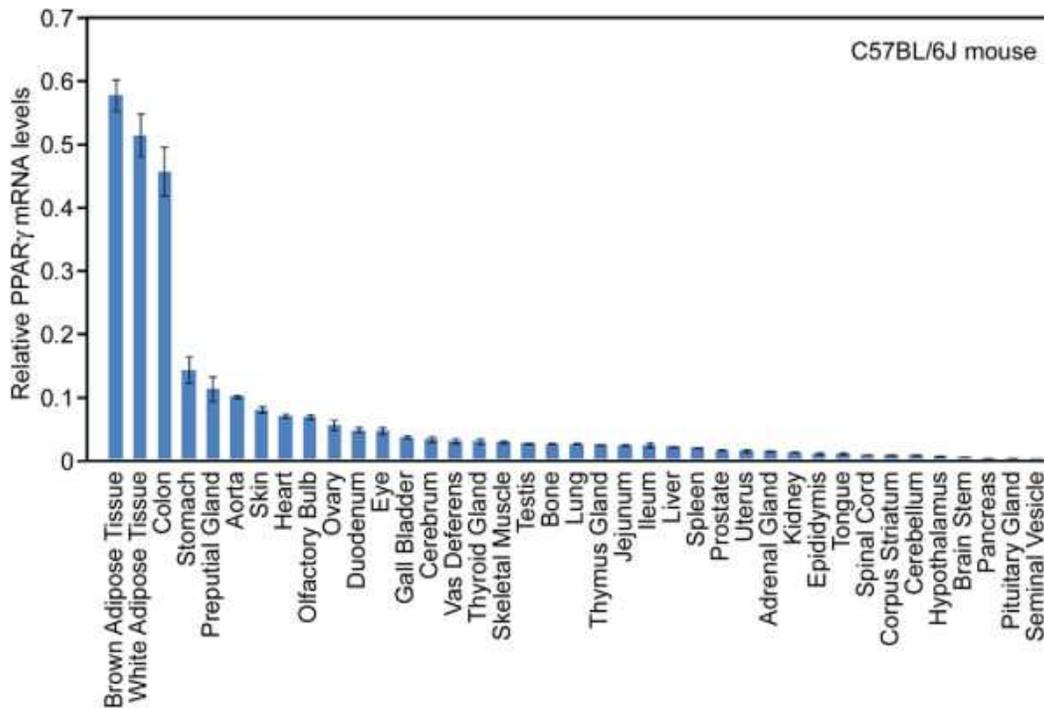
Figure 3. Tissue specific expression of PPARγ in mice. The relative abundance of the PPARγ mRNA in a wide variety of tissue types shows highest expression in adipose tissue and the colon but some expression in almost all tissue types. Figure was adapted from Lee et al[60].

1.3.1 Function of PPARγ in adipose tissue

The activity of PPARγ was first determined through its induction during adipogenesis[61,62]. The function of this protein is best understood through the roles it plays in adipogenesis and lipogenic pathways. As stated above the majority of PPARγ is expressed in adipose tissue. Adipose PPARγ has been shown to act as the master regulator of adipogenesis. This was first observed from PPARγ null mice which were completely unable to generate any adipose tissue[63]. Later experiments demonstrated that PPARγ knockout mouse embryonic fibroblasts were unable to be differentiated into adipocytes consistent with this protein playing an essential role in adipocyte differentiation[64]. Interestingly, it was also found through the selective ablation of PPARγ in mouse adipocytes that this protein is also required for maintenance as ablated cells died within a few days[65]. PPARγ was also identified as playing a role in the self-renewal of adipose tissue[66]. This was determined through the identification of a niche of adipocyte progenitor cells in the adipose vasculature which heavily expressed PPARγ and

could be differentiated into adipocytes by treatment with a PPARγ ligand unlike similar cells obtained from other tissues. Interestingly, adipose tissue is also the only location where PPARγ2 is expressed to significant levels in healthy animals. The specific role this isoform plays is not well understood as this tissue also heavily expresses PPARγ1.

A function unique to PPARγ in adipose tissue is the "beiging" of white adipose tissue (WAT)[67]. This leads to a phenotypic shift which makes this tissue behave more similarly to brown adipose tissue (BAT). It had been thought that humans lack this tissue type, but recent research determined that there is a somewhat small reservoir of BAT in humans and like other adipose cells PPARγ was essential for its differentiation and maintenance[68]. This tissue unlike WAT is highly metabolically active and consumes lipids through uncoupled mitochondrial respiration to generate heat[69]. There has been significant research into this "beiging" phenomenon as due to the uncoupled oxidation of fatty acids this could be exploited for weight loss or the treatment of certain metabolic conditions. Thus far this "beiging" has only been observed following treatment with strong PPARγ activators and the exact mechanism is not fully clear[70]. The current hypothesized pathway is through stabilization of PRDM16, a transcription factor linked to BAT development. It has been found that the activation of PPARγ leads to an increased half-life for PRDM16[71]. This is supported by findings that deacetylation of PPARγ on two lysine residues, 240/268 and 265/293, leads to an association between PPARγ and PRDM16[72]. These findings have led to research to determine if this pathway is PPARγ transactivation independent and whether a ligand can be found which will preferentially increase this "beiging" without leading to canonical PPARγ signaling.

1.3.2 Whole body signaling of PPARγ expressed in adipose tissue

As well as regulating adipogenesis PPARγ also has important roles in adipose tissue which regulate glucose homeostasis. PPARγ activity increases the expression of the glucose transporter Glut4[73,74] and the c-cbl associated protein (CAP)[75]. Both of these proteins play important roles in insulin signaling. There is

also thought to be function in increasing insulin sensitization through the modulation of the release of cytokines and adipokines. It has been shown that PPARγ activation increases the release of adiponectin[76]. This adipokine is important in insulin signaling through signaling an increase in glucose uptake and a decrease in gluconeogenesis. Another important adipokine that is downregulated is leptin which is important in the regulation of the hunger response[77,78]. A significant amount of the PPARγ induced insulin sensitivity also comes from the reduction of pro-inflammatory molecules. For instance, high circulating glucose levels are highly inflammatory and have been linked to insulin resistance. As well PPARγ activation has been shown to reduce the expression of the pro-inflammatory cytokine TNF-α[79]. This cytokine is also implicated in insulin resistance when at high levels and is frequently upregulated with type II diabetes mellitus (T2DM)[80].

Due to these strong effects in insulin sensitization PPARγ has been studied extensively as a medicinal target for T2DM. This has led to the development of multiple drugs which have been approved for human use by the FDA. Unfortunately, all of these drugs have been tied to severe negative side effects leading to the restriction or recall of some. Medicinally relevant ligands of PPARγ will be covered in more detail below in the section on synthetic ligands with the FDA approved ligands covered in the section on full agonists.

1.3.3 Negative effects of PPARγ activation

As stated earlier, PPARγ is expressed to low levels in the majority of tissue types. In many tissues this leads to beneficial effects such as are seen in adipose tissue. PPARγ activation in skeletal muscle[81], the pancreas[82,83] or the liver all leads to insulin sensitizing effects similar to those seen in adipocytes. Though, in the liver it is hypothesized that PPARγ activity leads to steatosis, a damaging accumulation of adipocytes[84]. However, literature also exists that suggests that PPARγ reduces liver steatosis[85,86]. So, the overall effect on liver health is unclear at this point. There is also the observed effect of a pronounced decrease in inflammation in the brain. This has been of great interest in recent research

towards the treatment of neurodegenerative diseases. This will be covered in greater detail later in the section on anti-inflammatory effects. However, in some tissue types there are severe effects related to the activation of PPARγ which give rise to the negative side effects that have caused the recall or restriction of drugs targeting this protein.

The best understood of these is a decrease in bone strength. This is caused by a decrease in osteoblastogenesis, the formation of bone tissue, and an increase in osteoclastogenesis, the resorption of bone cells. It has been found that the activation of PPARγ inhibits bone formation through a reduction in differentiation of osteoblasts[87,88]. To support this, it has also been shown that PPARγ-null embryonic stem cells will spontaneously differentiate to osteoblasts unlike normal embryonic stem cells[89]. As well, PPARγ-null hematopoietic mouse lineages are deficient in osteoclastogenesis[90]. This data indicates that the loss in bone integrity with PPARγ ligands occurs through the downregulation of osteoblastogenesis and upregulation in osteoclastogenesis. There is also evidence that PPARγ ligands can cause spontaneous differentiation of osteoblast precursors to adipocytes and lead to fat deposits forming in bone further decreasing bone integrity[91]. This was observed to lead to a large increase in hip fracture particularly in older women.

Another common side effect seen with drugs targeting PPARγ is weight gain. This was originally thought to be due to the increase in adipogenesis caused by the activation of PPARγ. However, later work in two independent studies performed in mice implicated that this was actually due to the activation of PPARγ in the brain. This was determined in mice using both pharmacological and genetic methods[92,93]. Such findings were somewhat surprising as it was not thought that PPARγ ligands were capable of crossing the blood-brain barrier. These findings suggest that to understand the overall physiological ramifications of activating this protein it is not prudent to consider only the adipose tissue where these ligands have their most pronounced effect.

There are other severe side effects with major ramifications regarding cardiac health. The first of these arises from PPARγ activity in the kidneys. It has

been shown that treatment with PPARγ agonists helps improve the pathogenesis of diabetic nephropathy (kidney damage)[94], but this treatment also leads to edema. This arises due to an altered sodium and fluid reabsorption in the kidney. But, the exact mechanism is not totally clear as multiple studies on this topic have reported conflicting results[95–97]. Additionally, there are cardiac issues which occur from the activation of PPARγ in the heart[98]. In a healthy individual PPARγ is expressed to very low levels in the heart but with individuals who have metabolic disorders such as T2DM, heart PPARγ levels are greatly increased. Treatment with PPARγ ligands has been shown to increase the mass and fat content of the heart in mice[99]. As well, mice treated with PPARγ ligands have mitochondrial damage in the heart and develop cardiac myopathy[100].

1.3.4 Medical Ramifications

Given the findings outlined above it is clear that PPARγ is a medicinal target of high value. This is particularly true for metabolic disorders such as T2DM and atherosclerosis. However, the current system of treatment does not appear to be optimal due to the negative side effects caused by increased activity of this protein. This has led research to shift to more selective modulators which activate the protein only fractionally but may give beneficial results through other signaling pathways. Unfortunately, the mechanisms which lead to such selective activation are incompletely or in some cases very poorly understood. This indicates that significantly more research is needed into specific mechanisms of PPARγ ligands which would allow for more intelligent drug design which could avoid or reduce negative side effects while retaining a majority of the desired therapeutic outcome.
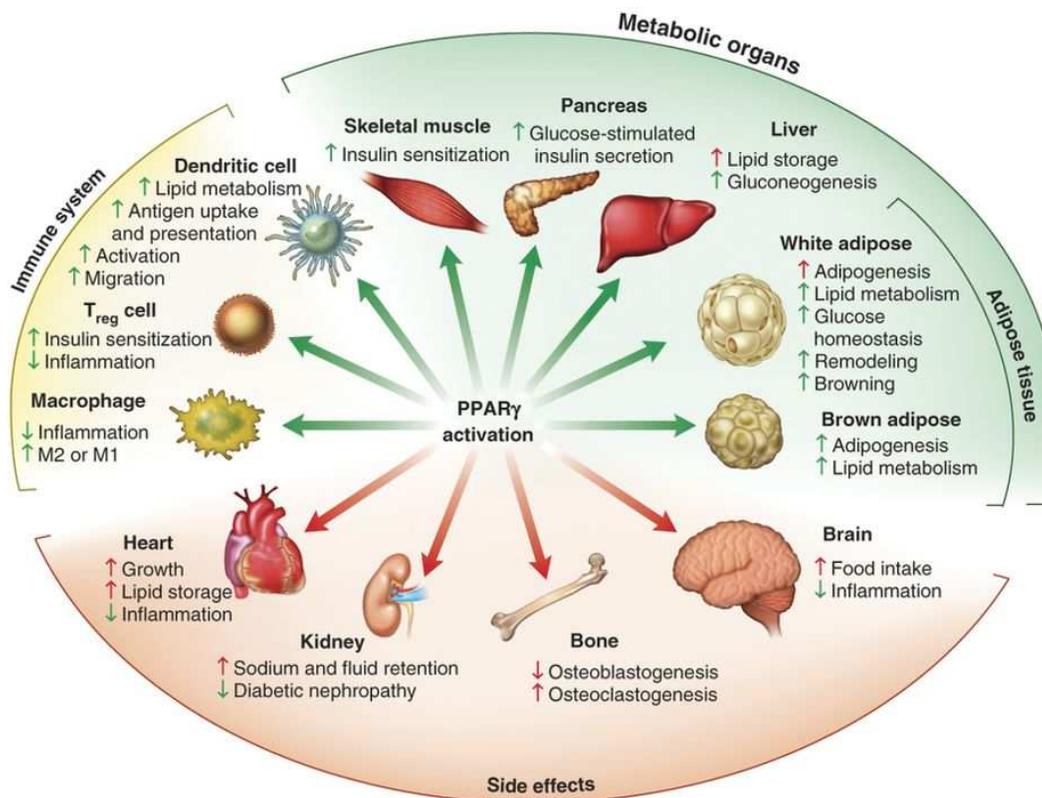
Figure 4: Effect of PPARγ activation in different tissues. The activation of PPARγ leads to beneficial effects (Green arrows) in certain tissues but also very negative effects in other tissue types. Figure was adapted from Ahmadian et al[101].

1.4 Endogenous ligands of PPARγ

There is some debate as to the natural ligands of PPARγ. It is technically classified as an orphan receptor but several endogenous lipids which can bind to PPARγ have been identified. The concern with these natural ligands is that the identified fatty acids which bind PPARγ have very poor affinity, usually >10µM, this raises the concern that under physiological conditions these are never at a high enough concentration to bind. Additionally, they only function as partial agonists and it is thought that a more efficacious endogenous ligand should exist. The majority of these fatty acids which bind to PPARγ do so in a covalent manner to C285/313 on helix 3. Such a ligand would circumvent the low affinity. Only at the initial non-covalent binding step would concentration of the lipid have to be very high and then once the covalent bond is established there would be no risk of

ligand release due to low affinity. That has raised some concern as this would reduce the temporal control of signaling from this protein. This can be rationalized by the fact that many synthetic ligands likely have very low off rates due to their high affinity and would be bound for prolonged periods. Additionally, PPARγ has a relatively short half-life so the covalent binding of a ligand would not have too prolonged an effect[102,103].

All of the postulated endogenous ligands that have been identified thus far are polyunsaturated fatty acids (PUFA). The first ligand to be identified in 1995 is prostaglandin D_2 with metabolites in the Prostaglandin J_2 pathway having stronger effects. 15-deoxy-Δ12,14-Prostaglandin J_2 (15d-PGJ$_2$) showed the greatest effect by a fairly large margin, approximately 65% of the efficacy of the synthetic full agonist rosiglitazone[104,105]. This fatty acid requires the presence of C285/313 for function, though it will still bind at low affinity when this residue is mutated but ligand activity is abolished. Crystallography revealed this dependence on C285/313 to be due to the formation of a covalent bond between one of the sites of unsaturation and this cysteine residue in PPARγ[106]. The affinity is very poor with a transcriptional reporter assay failing to plateau at a concentration of 10μM. This may indicate that this lipid could activate the protein to a greater degree and the level seen is a function of only partial binding. However, when dealing with an endogenous ligand the likelihood of micromolar concentrations being present is very low. Particularly for a lipid such as this which can readily bind to a variety of proteins and will have very poor water solubility.

Later work examined the effect of the binding of several other PUFA to PPARγ and was able to identify three new eicosanoid lipids which were able to activate PPARγ to some degree[107]. These were examined in both a hydroxy and an oxo state to determine if covalent attachment was necessary for function, since the hydroxy state is unable to form a covalent bond to cysteine. From this they observed that the covalent interaction was not actually necessary for the PUFA to act as an agonist but in the absence of a covalent bond significantly higher concentrations were required to observe the same response. Even with the

covalent PUFA micromolar quantities were still required to reach the EC_{50} in a transcriptional reporter assay, between 6-10μM. These concentrations seem physiologically unlikely indicating that these many not be relevant in a biological system. Another criticism is that the transcriptional reporter assay contained no reference ligand, so it is very difficult to say to what level these lipids actually increased transcription. Due to variability in the experimental system a reference synthetic ligand of well-characterized behavior is typically included to more accurately determine the effect of ligand binding. The most interesting finding in this is that with one of the PUFA, 13-S-HODE, the authors were able to obtain a crystal structure of PPARγ bound to two molecules of the lipid. This was one of the first reports of PPARγ binding two ligands simultaneously[108]. This phenomenon has recently been more thoroughly studied with synthetic ligands and it does appear that the co-binding of endogenous lipids with synthetic ligands may be a phenomenon which modulates the ligand behavior. The author in this work postulated that these PUFA may not be the endogenous ligand but that it may be a related lipid. They suggest that the more common metabolites of arachidonic acid, 5-oxo-EPA and 5-oxo-ETE could readily reach sufficient concentration in peripheral blood mononuclear cells for binding[109]. However, they do not make it clear if the concentrations they report, 10-150μM, are total concentration for these lipids which would suggest that a large percentage of this lipid would likely be bound in membranes and inaccessible for binding to PPARγ. It is worth noting that one of these postulated PUFA has been demonstrated to bind to PPARγ in cancerous cell lines and arrest the cell cycle thereby promoting apoptosis[110].

With these findings it is still unclear if PPARγ is an orphan receptor or if the endogenous ligand has been discovered. Given the low affinity seen in the postulated endogenous PUFA ligands it seems likely that none of these would be suitable ligands for signaling in a cellular environment. But to fully demonstrate this will require further research and it is possible that the exact endogenous ligand has not been identified but that the endogenous ligand may be closely structurally related to those which have been identified. It also should be noted

that the majority of research postulates that there is a single endogenous ligand. This is true for some nuclear receptors such as RxR and other vitamin receptors. However, PPARγ is lipid-dependent and expressed in a wide variety of tissues. Different tissues have markedly different lipid profiles, so it may well be that there is not a single endogenous ligand for this protein but several that are tissue dependent and may regulate pathways which are more relevant in that tissue. To provide insight as to their structure several lipids which have been identified as putative PPARγ ligands are shown in figure 7.

1.5 Synthetic ligands of PPARγ

Due to the medicinal importance of PPARγ a very large number of synthetic ligands have been made for this protein. Interestingly they are not very structurally conserved. There are some moieties which are relatively common in structures, such as a central indole, but the only thing which is completely in common is an amphipathic nature. These ligands must be somewhat amphipathic as the ligand binding pocket of PPARγ is intended to bind fatty acids and is therefore relatively hydrophobic. Due to the large size of the ligand binding pocket[22], ~1500Å^3, it can be difficult to exploit all potential interactions, but there are important residues which can be utilized to drive hydrogen bonding and therefore increase affinity and specificity[111]. A very common example of this is S342/370 in the β-sheet region. This residue forms hydrogen bonds either through the side chain or more commonly the backbone to a very large number of ligands which are structurally and functionally distinct. Due to this bonding event not being correlated with ligand function it is likely only significant as a potential target to increase selectivity and affinity. There are other residues which are frequently involved in hydrogen bonding that are distinct to the ligand class which will be discussed in more detail later.

Another interesting characteristic of the binding of synthetic ligands to PPARγ is the so-called alternative binding site. This was discovered in 2015 through the binding of two separate ligand molecules to a single protein[112]. This binding does not occupy the standard binding pocket but instead the ligand binds

in a region between the N-terminal region of helix 3 and the Ω-loop. It was originally thought that this binding only occurred as a result of a second ligand binding event and could explain some of the odd behavior seen with some partial agonists at high concentration. The second binding event is typically only observed with partial agonists and typically occurs at very low affinity, EC_{50} values of approximately 10μM. However, since then there have been cases where ligands appear to preferentially occupy the alternative site. The partial agonists S35 and SR1664 (PDB 5DWL) have both been crystallized bound only to the alternative site in the absence of other ligand molecules. It was suggested in the original paper on S35 that in solution this ligand had a preference to bind the alternative site almost exclusively[113].

A more in-depth discussion of synthetic ligands by their effect on the functional behavior of PPARγ will follow below.

1.5.1 Ligand nomenclature used in nuclear receptors

In the nuclear receptor field, ligand function is named based on their effect on transcriptional reporter assays. The most commonly performed of these is the GAL4-fusion transactivation assay. In this assay the LBD of the nuclear receptor of interest is fused to the DBD of the yeast protein GAL4. A vector with a GAL4-UAS reporter, generally some form of luciferase, is also transfected to provide a readout. Based on the ligand binding state of the nuclear receptor LBD the endogenous coregulator proteins will be recruited to the reporter plasmid and regulate transcription of the reporter gene. This system has the advantage that the GAL4-DBD binds to DNA as a monomer so this circumvents the necessity of a dimer partner[114]. However, this has the limitation that only the LBD of the nuclear receptor is being considered. Generally, ligand function is regulated primarily by the LBD but there are rare instances whereby the DBD or the unstructured AF-1 region will affect the activity of a ligand[115]. It appears likely that the relative scarcity of such reports is due to the almost exclusive use of the LBD of nuclear receptors for both *in vitro* and cell-based assays. The primary reason for this appears to be the simplicity of the system. This is particularly

troublesome for *in vitro* assays since full-length nuclear receptors are not very stable and will readily undergo proteolysis. This eventually leads to all protein degrading to the LBD state which adds uncertainty to the assay due to the presence of partially degraded protein which is functionally different than full-length.
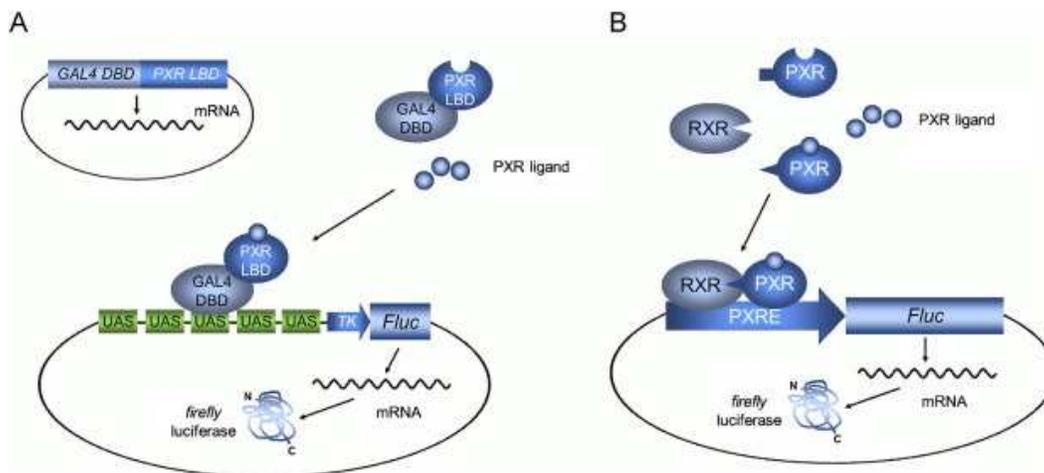
Figure 5. Schematic diagram of a luciferase transcriptional reporter assays. Diagrams showing the mechanism of either a GAL4 transactivation assay (left) or a PXRE transactivation assay (right). The PXRE reporter is analogous to the PPRE reporters used with PPAR proteins. Figure was adapted from Luckert et al[116].

A variant of the above technique is the yeast two-hybrid assay[117]. This is a similar technique but instead of a fusion of the GAL4-DBD to the nuclear receptor LBD two expression plasmids are used. On the first is a fusion of the GAL4-DBD with an interacting protein of interest, usually the NR. The second contains a fusion of the coregulator of interest with the active domain of GAL4. Thereby when the two proteins bind transcription of a reporter gene is activated and through titration of a small molecule a ligand-dependent affinity can be determined. This has since been expanded into mammalian cell lines and is more commonly performed in this manner and referred to as a mammalian two-hybrid[118]. The only major difference in the mammalian system is that the GAL4 active domain is unsuitable in a mammalian cell line so the active domain of a viral protein, VP16, which is constitutively in an active state with high coactivator affinity is used instead.

As the DNA recognition of nuclear receptors began to become better understood a similar assay grew in popularity. This is commonly referred to as the PPRE transcriptional reporter. In this assay, instead of relying on a fusion of the LBD, full-length PPARγ is overexpressed and the reporter plasmid has PPRE enhancer sequences upstream of the promoter[119].

Both of these systems have the same drawbacks which are not straightforward to overcome. The first is that in these systems you do not know what coactivators or corepressors are being recruited. As stated before different coactivators control very different pathways in the cell and the ability to ascertain if a ligand could preferentially recruit specific coactivators, similar to the biased agonists seen in GPCR proteins[121,122], would be of great value. Additionally, since the transcriptional machinery endogenous to the cell is not being overexpressed this sets a cap to the response that can be observed. Due to this limitation it is not currently clear if different full-agonists show different efficacy since any full agonist will eventually saturate the response.

Based on transactivation assays described above, ligand function has been broken up into three primary groups. The first is the full agonists, which strongly drive transcription and have historically been the most studied ligands for medicinal purposes. The next are the non/partial agonists. These either do not drive transcription at all or show only a fraction of the response of the full agonists. With recent discoveries non/partial agonists have become one of the leading areas in PPARγ medical research but come with the caveat that they are structurally diverse, and their mechanism of action is poorly understood. The last group is the antagonists/inverse agonists. These two types of ligands are indistinguishable in transactivation assays but induce very different outcomes. The antagonist ligands function by having a portion of the ligand, often a bulky hydrophobic group such as a tert-butyl, that sticks directly out of the AF-2 surface and thereby disrupts the very specifically organized hydrophobic cleft required for any coregulator interaction. On the other hand, the inverse agonists are thought to specifically drive the recruitment of corepressor proteins while reducing affinity

for coactivator proteins. There is some thought that the inverse agonist behavior observed is not in fact due to a direct function of the ligand and instead occurs through a displacement of endogenous lipids bound to the protein. PPARγ is capable of binding a wide variety of fatty acids in the cell and even some as simple as nonanoic acid have been shown to have partial agonist character. Because of this, there is some argument whether or not inverse agonism is actually occurring in PPARγ, but recent data suggests that it is[123].

1.5.2 Alternate functional assays to determine ligand class

There has been increasing use of *in vitro* assays to characterize ligands since the mid-2000s. Typically these assays utilize fluorescence-based techniques to look at coregulator binding to a purified NR protein as a function of the ligand bound. One flaw in these assays is the large size and unstable nature of most coregulators which are at least partially intrinsically disordered. As such, recombinant expression of the entire coregulator protein is usually not possible. Because of this most assays rely on a peptide containing only a single known LxxLL or CoRNR box. Some work has been done using the receptor interaction domain (RID) of coregulators such as SRC-2 but these are not as popular as these domains are still not trivial to purify to any significant yield[124] and show only minor differences with assays utilizing the small peptides containing a single interaction motif. Another restriction is that at least one coregulator which binds with decent affinity must be known to perform these assays. Generally, for most nuclear receptors at least a portion of the coregulator interactome is known but specific affinity data is rare. Additionally, many coregulators contain multiple LxxLL[125] or CoRNR[126] boxes that bind different proteins with significantly different affinities which further complicates these studies since most of the interaction data was obtained from cells and the exact interacting motif is often ambiguous. There are two main forms of this assay which will be outlined below.

The first form of this assay performed to any great extent was based on Förster resonance energy transfer (FRET) often performed in a time-resolved (TR-FRET) fashion to reduce noise[127–129]. For simplicity, all FRET assays will be

referred to in this form, TR-FRET, for this work. In such an assay a fluorescently modified NR protein, most often through labeling with a fluorescently modified antibody, is pre-mixed with a fluorescently labeled coregulator peptide, often covalently modified with a fluorophore. To this mixture can then be titrated a solution of the ligand of interest and the change in FRET ratio can be fit to estimate binding of the coregulator to the NR. There are multiple downsides to this technique. The first is the very practical matter of cost, having to utilize multiple fluorophores and antibodies makes this a relatively expensive assay. The second more significant downside is that the data are somewhat challenging to interpret and relate to other data. Since the ligand is being titrated the EC_{50} value obtained from the fit is really only representative of the ligand's affinity for the NR protein and not the affinity of the NR for the coregulator. To look at that relative affinity you instead have to examine the FRET window between unliganded and ligand bound and relate the fold-change in FRET ratio to estimate the affinity of the NR for the peptide. As well these assays are not exceedingly reproducible, relative effects seen in ligand behavior remain constant, but other properties such as FRET ratio or window are somewhat different between runs making in depth interpretation difficult.
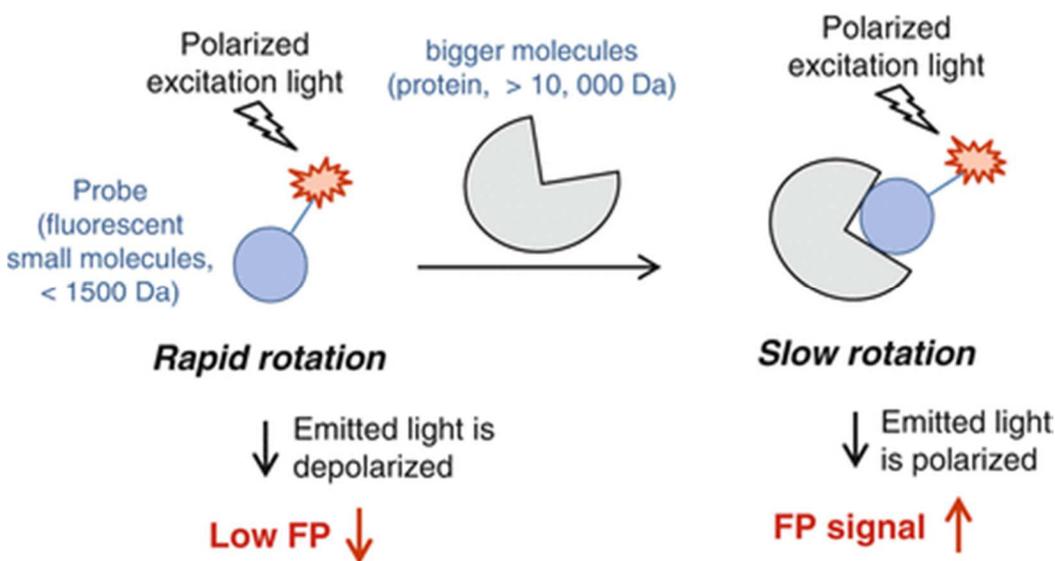
Figure 6. Fluorescence polarization binding assays. A schematic diagram of the fluorescence polarization assay. This technique is frequently used in NR through titrations to determine K_d

values of coregulator peptides. Figure was adapted from Fluorescence Polarization Assay to Quantify Protein-Protein Interactions in an HTS Format[130].

Once the technique became suitably adapted to a high-throughput assay the use of fluorescence polarization or anisotropy to look at binding events of NR with coregulators became popular, e.g. VDR[131], THRb[132], SRC1[133]. This system is much simpler as it requires only modifying either the protein or peptide with a fluorophore. This is typically done by the covalent addition of a fluorophore to the peptide. With the fluorescently modified component held at a fixed concentration the other protein can then be titrated in either pre-loaded with the ligand of interest or with a fixed ligand concentration in each well. The second method with a fixed ligand concentration is somewhat less common due to the low water solubility of nuclear receptor ligands. Given this poor solubility at low protein concentrations where the majority of the ligand is unbound there are concerns that the ligand will fall out of solution or aggregate and thereby have unforeseen effects. As the complex binds and the molecular mass increases the tumbling rate will lower increasing the polarization or anisotropy of the fluorophore. This increase can then be fit to determine a dissociation constant, K_d, for the coregulator and NR in each ligand bound state. This technique has become increasingly popular due to the low cost and relatively high throughput. It also allows for a more sensitive way to interpret the behavior of full agonists and inverse agonists than the cell-based assays.

In 2009 a high throughput modification of the TR-FRET protocol was developed by Pamgene. In this protocol 53 peptides are immobilized on a chip and GST-fusion NR proteins can be titrated in with ligands bound[36]. A fluorescently modified anti-GST antibody is then added and the binding affinity can be determined from a non-linear regression. This technique is quite interesting but has not generated a lot of use in academic research most likely due to the cost. At this point only two papers can be located that utilize this first-generation technique. The first is the initial methods paper from the company and the other is looking at coregulator binding patters in estrogen receptor alpha (ERα) obtained from tumor cells[134]. The same group released a newer version of the technology in

a paper from February 2018 which reported an increase i the number of coregulator peptides to 154 and looked at changes in coregulator affinity with both the apo state and a reference full agonist ligand[135]. However, the results reported in this paper were not entirely consistent with results from their previous paper and with some of our own data, so there may be some issues in this newer technique. There are a couple of publications using this newer technology in the orphan receptor NR2E1 (TLX)[136] or with ERα[4] which showed robust results so the inconsistencies in this method may be more of an issue with PPARγ and not a problem for all NR. The previous generation of the Pamgene assay does seem robust and we have found good agreement with their results and our own coregulator recruitment data.

1.5.3 Full agonist ligands of PPARγ

Many full agonist ligands for PPARγ were developed as potential therapeutics of type II diabetes. A significant number of these ligands have gone into clinical or preclinical trials but relatively few have been found to be suitable for human use. This unsuitability is due to a wide variety of negative side effects. Some of these arise as an effect of canonical PPARγ activation which has been a concern in the development of therapeutic agents which target this protein. However, there are a number of other severe side effects that occur from full agonists of PPARγ where the mechanism is unclear. Despite this, full agonist ligands of PPARγ are still an active area of study, though more modern compounds have been designed to not activate the protein as significantly as prior drugs. Through this the impact of side effects has lessened but is still a very significant concern.

Full agonism is the best understood mechanism of action in PPARγ. Almost all full agonists share the same core mechanism which consists of a direct hydrogen bond from the ligand to Y474/501 on helix 12[137,138]. This interaction stabilizes helix 12 in a conformation which is appropriate for interaction with the LxxLL motif on coactivators. These ligands also make other conserved hydrogen bonds within the protein to S289/317 on helix 3, H323/351 on helix 5 and

H449/477 on helix 11[139]. These aid in stabilizing other portions of the AF-2 region in a suitable conformation to interact with coactivators. NMR studies have also shown that full agonists reduce the available conformational ensemble more so than the other ligand classes[140]. The significance of this is not understood but it is hypothesized that limiting the plasticity of the protein will lock the protein in a conformation suitable for coactivator recruitment and thereby help to lower the energetic penalty associated with coactivator binding, which will increase binding affinity. Our own data, which will be shown in chapter 3, support this hypothesis and also indicates that it is not only the overall stability of the protein which is important for full agonist function but that the stabilization of the β-sheets, the Ω-loop and helix 5 in particular are important for the highest efficacy in agonism. Interactions in these regions are important in an optimal orientation of the AF-2 residues for productive interactions with the coactivator peptide. Additionally, these data illustrate that the AF-2 surface is not defined solely by interactions on helix 3 and helix 12 but that there are important interactions, both selective and general, arising from residues in helices 3' and 4.

Two major classes of full agonist ligands for PPARγ exist. These are the glitazones which contain an thiazolidinedione (TZD)[141] head group which mediates the important interactions with the AF-2 region and the glitazars which are derivatized from L-tyrosine[142]. Interactions are more complex with the glitazar ligands than with the glitazone ligands. This is due to the fact they are typically significantly more structurally complex than the glitazones and while the AF-2 interactions do arise from the conserved carboxylate within the tyrosyl moiety this group is frequently located in the central region of the molecule. Such greater structural complexity allows the glitazar ligands the ability to interact more extensively in significant regions than the glitazones, particularly around the Ω-loop and β-sheets. This greater degree of interactions appears to lead to higher affinity than is seen for the glitazones and in some cases greater agonist efficacy. It should be stated however, that higher efficacy is not a hallmark of the glitazar ligands as some including saroglitazar and aleglitazar only act as partial agonists for PPARγ. Unfortunately, the glitazar ligands have been strongly linked with

very negative side effects and only a single member, saroglitazar which is a pan-PPARα/γ ligand, has been approved for human use and is only approved in India[143]. There has been greater success with ligands from the glitazone family. Of these three have been approved for human use by the FDA (troglitazone, rosiglitazone and pioglitazone) and an additional one is approved for human use in Canada and the European Union, edaglitazone.

The first of the FDA approved full agonists was troglitazone (Rezulin). This drug was approved for human use in 1997 but within one year of use a severe side effect was observed. This side effect had not been apparent in either preclinical animal or human testing. It was found that in a small percentage of people taking this drug severe liver damage was occurring in some cases leading to liver failure. Due to this negative side effect the drug lost approval for human use in 2000. It is still not entirely clear what the cause of this side effect was but there is some evidence that, in mice with dysfunctional oxidative stress responses in the mitochondria of liver hepatocytes, liver damage is induced by troglitazone. This suggests that troglitazone may be causing liver damage through an increase in oxidative stress in hepatocyte mitochondria[144,145]. However, this damage is relatively mild which is inconsistent with the side effect observed in humans.

Research continued into this ligand family and the next compound to be approved for human use was rosiglitazone (Avandia) in 1999. This ligand also induced severe side effects including edema, weight gain, increased bone fracture in women and cardiac issues. There were a number of reports of heart failure associated with this medication[146–148] which led in 2007 to the addition of a black box warning which severely limited the availability in the U.S. Around the same time this medication was completely banned for human use by the European Medical Association (EMA). Since this time the black box warning has been lifted in the U.S. following a large randomized study by the manufacturer which showed the initial studies indicating heart failure were not reproducible, but this medication remains banned in Europe. Following this a similar medication, pioglitazone (Actos), was approved which carried much milder cardiac risks and

in some research has actually been shown to improve cardiac health in type 2 diabetes patients[149]. This medication unfortunately has been linked to bladder cancer risk leading to its ban in portions of Europe[150]. Additionally, it carries the same side effects seen in rosiglitazone with the exception of the cardiac issues. These shared side effects (edema, bone fracture and weight gain) appear to occur as a consequence of PPARγ signaling and so may occur with any full agonist drug.

Due to severe side effects, research has begun to shift away from these full agonists towards less efficacious partial agonists. Many of these compounds have gone into clinical or preclinical trials but as yet none have been approved for human use. What is known about partial agonists will be covered in greater detail below.
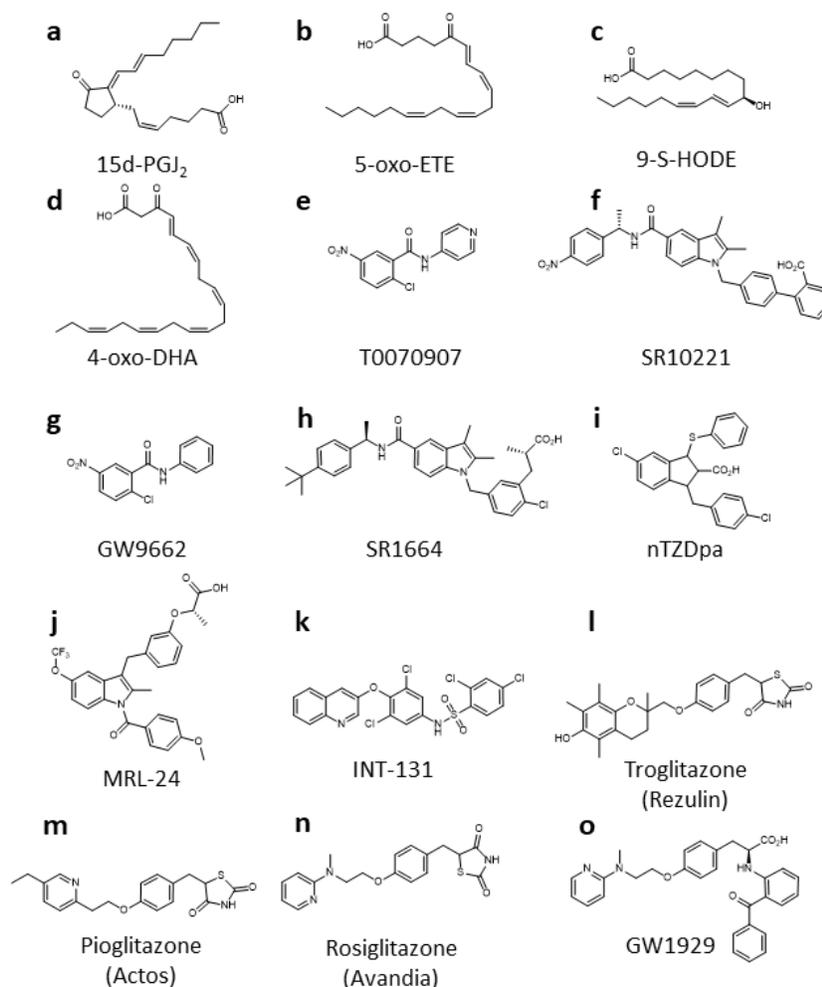
Figure 7. Chemical structures of representative ligands of PPARγ. A-D: The structure of four endogenous lipids which have been shown to bind PPARγ with an approximately 10μM affinity. All of these except 9-S-HODE bind covalently to C285/313. E-G: Structures of three inverse agonist ligands. T0070907 and GW9662 are both covalent and appear to act through a distinct mechanism to SR10221. H-K: The structure of 4 partial agonist ligands of PPARγ. These range in activity from no transactivation activity, SR1664, to fairly high activity, INT-131. I-O: Structures of four representative full agonist ligands. Troglitazone, Pioglitazone and Rosiglitazone are all TZD ligands which have been FDA approved for human use. GW1929 is a glitazar ligand which shows higher activity then the TZD ligands but is unsuitable for medical use.

1.5.4 Inverse agonists and antagonists of PPARγ

Comparatively few inverse agonists or antagonists have been explored in literature relative to the other ligand types. Typically, these are not thought to have a wide variety of medicinally useful effects, which limits research interest. There have been some reports of the covalent inverse agonists T0070907[151,152] and GW9662[153,154] having an anticancer effect by arresting the replication of human cancer cell lines. The issue in this research is that both of these ligands have rather strong off target binding, GW9662 can bind PPARα with a K_d of approximately 20nM[155] and T0070907 can do the same with a K_d of around 1μM[156], both of these concentrations are much lower than were used in these studies. The K_d values estimated for these two ligands are somewhat based on assumption since they are both covalent, they have no k_{off} value and so the K_d reflects only a contribution from the k_{on}. As well the mixed agonist/antagonist BADGE was able to promote apoptosis in cancer cells, but this was only at concentrations in excess of 50μM. This would suggest that the observed effects are almost certainly due to off target binding and not due to PPARγ signaling[157].

The mechanism of inverse agonism has been of some interest. A lot of this interest comes as a comparator to other NR where inverse agonists are of more value (e.g. ERα where the antagonist tamoxifen is the leading breast cancer medication). It is thought the mechanism of inverse agonism is shared among nuclear receptors so a greater understanding in one system should be of use in studies of a different NR. Typically, inverse agonists are thought to function through the displacement of helix 12 from the AF-2 surface[158]. This hypothesis is based on crystal structures of various other nuclear receptors bound to the corepressor SMRT in which helix 12 is never resolved and the helix 11- helix 12 loop points into space[159]. This is further supported by a pair of dominant negative NR, Rev-Erb α and β which completely lack a helix 12. The lack of this helix is thought to be what makes these proteins dominant negative due to its importance in coactivator interactions. That would be consistent with the observation that the

CoRNR box has a larger footprint when bound to the nuclear receptor then the LxxLL box so some conformational change must occur to prevent a direct steric clash with helix 12. Such a clash can be observed when a corepressor peptide from a crystal structure is docked on an apo PPARγ structure. Unfortunately, no crystal structure of a corepressor bound to PPARγ has been published as of yet. There are a small number of inverse agonist or antagonist ligands that have been crystallized bound to PPARγ. Recent structures of the antagonists SR10171 and SR11023 indicate that helix 12 is being pulled away from the AF-2 surface and towards the N-terminus of helix 3[48]. This conformational change would generate the additional space needed for the CoRNR box interaction but is surprising given that typically helix 12 appears to be pushed off the surface in other nuclear receptors not pulled to a different portion of the protein.

Our own data, which will be presented in chapter 4 of this work, indicates that there are two distinct mechanisms of inverse agonism in PPARγ. The first is the displacement of helix 12 which occurs with bulky ligands that appear to directly make sterically unfavorable interactions with helix 12 and lead to displacement of this helix from the AF-2 surface. The second which is unique to the comparatively small covalent inverse agonists appears to be driven by a relocalization of helix 12 to a position lower relative to helices 3 and 4 than it is in a partial agonist or agonist bound state but oriented in the same fashion as it would be in an agonist bound state. Such a mechanism leads to helix 12 actually making productive binding interactions with corepressors. It appears likely that these productive interactions with helix 12 are what make T0070907, one of the covalent inverse agonists mentioned above, the most efficacious known ligand of its class. Some example structures of representative inverse agonists can be seen in figure 7.

1.5.5 Partial agonist ligands of PPARγ

Recently the partial and non-agonists have become the most heavily studied group of PPARγ ligands[160]. The difference between these two groups is not well defined and many works will refer to the same ligand as either a partial

or non-agonist. For the purpose of brevity in this work all ligands from these groups will be referred to as partial agonists. Partial agonists are synthetic ligands that activate PPARγ transcription to only a small fraction of the full agonist response. These responses range from very weak: SR1664[161] or nTZDpa[162] have almost no transcriptional effect, to somewhat marked effects: Int131 activates up to approximately 40% the level of a full agonist[163]. This group is the most structurally diverse and therefore it has been difficult to determine what interactions are significant for their effects. Additionally they make relatively few direct interactions with the protein which further complicates a mechanistic understanding at an atomic level.

The mechanism of action for partial agonists is very poorly understood. It has been shown that they function in a manner independent of helix 12 conformation which is unsurprising as they do not form direct hydrogen bonds to Y473/501 as the full agonists do[164]. Rather they make extensive contacts on helix 3, some contacts to the Ω-loop and direct hydrogen bonds to S342/370 in the β-sheet region[165]. Particularly during the second or alternative binding event they make no contacts at all on the AF-2 surface. Additionally, they do not stabilize the protein as is seen with the full agonists. In complexes bound to partial agonists, the protein appears to remain highly plastic with a similar amount of rigidity as is seen in the apo state, when assayed by 2-dimensional NMR techniques. This is consistent with our own observations that the binding of partial agonist ligands does not restrict the conformational ensemble of either helix 3 or helix 12 of the AF-2 region, these data are presented shown in chapters 2 and 3. Despite this, a large number of these ligands still drive coactivator recruitment in *in vitro* assays and are seen to drive transcription through either transactivation assays or qPCR. How exactly they have this agonist behavior is not understood at this point.

While the mechanism remains elusive it has been seen that many partial agonists increase insulin sensitivity in cell-based assays or in model animals. They also carry far less side effects than seen with full agonists which often cause weight gain, edema or bone fracture. This reduction in side effect severity is

33

thought to be due to the lesser transactivation of PPARγ seen with this ligand family[160]. Some of these ligands, particularly SR1664[166] and MRL24[167], were observed to increase insulin sensitivity to almost the level seen with the full agonists. Due to this several of these ligands have gone into clinical or preclinical trials for type II diabetes treatment but as yet none have been approved for human use.

1.6 Other significant areas in drug design for PPARγ
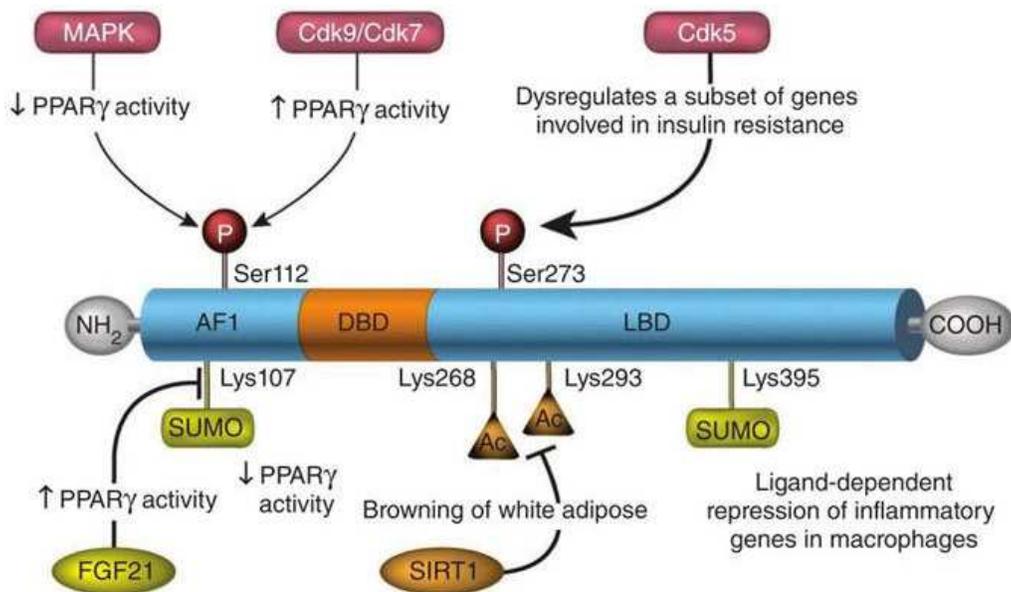
Figure 8. Post-translational modifications of PPARγ. A representation of PPARγ showing identified post-translational modification sites and their functional outcome. Figure was adapted from Ahmadian et al[101].

1.6.1 Post-translational modifications of PPARγ

Several post-translational modifications occur of PPARγ which have significant effects on the functional outcome. This includes a phosphorylation event on the AF-1 at S84/112 which has been well demonstrated to reduce the activity of PPARγ[168,169]. A sumoylation event on K79/107 has been shown to downregulate PPARγ activity. Interestingly an association with ubiquitin ligases, PIASxβ or Ubc9, at this site upregulates PPARγ activity in a manner independent of sumoylation[170]. As well, it was found that preventing this sumoylation event lead to insulin sensitizing effects in mice but no increase in adipogenesis[171]. This

34

may be a future avenue of drug discovery but this finding was only made this year so as yet no further research has been published. The deacetylation of PPARγ by SIRT1 has been determined to be important for the so-called "beiging" of adipose tissue which can be induced by PPARγ ligands. Acetylation of PPARγ was shown to directly inhibit association with PRDM16 which is essential for this observed behavior[72]. One of these modifications, sumoylation of K367/395, is suggested to be important for the anti-inflammatory effects seen with PPARγ[172]. However, there is debate as to whether this actually occurs with other literature directly stating that this residue is in fact not a target of sumoylation[173]. The most recently discovered post-translational modification is the phosphorylation of S245/273 which will be covered in significant detail below due to the fact it is the leading area of active drug development research.

1.6.2 Regulation of the phosphorylation of S245/273

The mechanism of increased insulin sensitivity seen with partial agonists remained elusive until 2010 when a post translational modification was discovered in the work of Choi et al[174]. This modification is a phosphorylation of S245/273 in a unstructured loop proximal to the β-sheets. The phosphorylation of this residue was confirmed by an alanine mutation which increased the expression of several proteins dysregulated in obesity. It was found that the phosphorylation of this residue almost completely abolished the insulin sensitizing effects seen with PPARγ ligands. This phosphorylation also seemed to lead to an overall downregulation of expression of PPARγ regulated genes. PPARγ contains a site which matches four of the five residues for a cyclin-dependent kinase 5 (CDK5) recognition motif[175] in a loop between the first of the β-strands and helix 2'. As such this appeared to be the mediating kinase[176]. The role of this kinase was demonstrated by the addition of a specific inhibitor of CDK5 or RNAi of CDK5 which strongly reduced the phosphorylation. Tracking levels of this kinase showed that mice fed a high fat diet (HFD), 60% fat, had greatly increased levels of CDK5 in adipose tissue. This suggests that the upregulation of this kinase and phosphorylation of PPARγ could be an important cause of dysregulation of many PPARγ controlled genes.

It was also found that the addition of PPARγ specific drugs significantly reduced this phosphorylation event. The addition of rosiglitazone to an *in vitro* kinase assay showed a 50% reduction in phosphorylation at 30nM, relatively close to the IC_{50} value for this ligand, though large doses of approximately 3μM were required to prevent phosphorylation. A much stronger effect was seen with the partial agonist MRL24 which was able to block most phosphorylation at a level of only 30nM. However, this ligand has a sub nanomolar IC_{50} value, so it is not clear if the difference is due to ligand behavior or merely an artifact of the fraction bound. Further examination of other partial agonist ligands which function well in insulin sensitization showed they also efficiently prevented this phosphorylation event. In fact the ligands most effective at preventing this phosphorylation were all partial agonists with the exception of a single full agonist, farglitazar. Other glitazars were not utilized in this study so it is not clear if other ligands of this class such as saroglitazar or GW1929 might also be very effective at blocking this post-translational modification. The authors postulate this reduction is due to a stabilization of this loop as evidenced by reduced hydrogen deuterium exchange of this loop and the neighboring β-sheet region. Later work by the same group suggests that thyroid receptor associated protein 3 (THRAP3) is recruited to PPARγ through this phosphorylated residue and that this is the root cause of dysfunction in gene expression caused by CDK5[177].

Based on these findings research into PPARγ therapeutics has shifted towards partial agonist ligands which can effectively block phosphorylation. The performance of ligands in an *in vitro* phosphorylation assay has become somewhat of a gold standard in papers covering novel ligands and their potential value in insulin sensitization. One concern with this assay however is that it seems to only perform well in a western blot[178]. Western blotting gives rise to complications as to quantification. Typically, these must be quantified through pixel counting which is not the most precise measurement and can have a smaller dynamic range than other techniques. The development of fluorescent indicators or automated systems which utilize chemiluminescent signals have greatly improved the potential accuracy in quantifying a blot but are not as commonly

used as the traditional horseradish peroxidase (HRP) colorimetric indicator and pixel counting[179]. In a few papers an ADP-Glo assay is used to measure phosphorylation through the consumption of ATP but this is not used commonly[180]. A single methods paper exists which outlined a protocol to perform this assay through enzyme-linked immunosorbent assay (ELISA)[181] which would improve the accuracy of quantification, but this method does not seem to have been adopted by the research community. Our own efforts to reproduce this assay have been unsuccessful.

1.6.3 Anti-inflammatory effects in nervous system

It has previously been observed that a number of PPARγ ligands have moderately strong anti-inflammatory effects. Interestingly these are observed to relatively similar levels in both full agonists such as rosiglitazone[182] or pioglitazone[183] as well as with partial agonists such as MRL24, UHC1[178] and INT131. The exact mechanism of this behavior is currently not well understood. It has been shown that this effect is independent of PPARγ DNA binding[184] and is therefore postulated to be due to a poorly understood transrepressive activity or possibly to occur through the S245/273 phosphorylation. The most common hypothesis for how these anti-inflammatory effects occur is through PPARγ binding to the NFκB subunit p65[185]. This interaction has been demonstrated to occur in cells via immunoprecipitation assays. The majority of anti-inflammatory effects observed with PPARγ ligands occur through the downregulation of NFκB signaling[186–188]. Due to this observation it is postulated that in a ligand dependent manner PPARγ associates with NFκB and prevents the transcriptional activity of this pro-inflammatory transcription factor. It is not clear if this is occurring in a manner similar to NFκBIB proteins which sequester NFκB and thereby downregulate activity or if possibly this effect may be due to PPARγ binding to NFκB which is bound to DNA in a repressive manner. Such transrepression has been reported with other nuclear receptors such as glucocorticoid receptor (GR)[189] but has not been seen to occur with this protein.
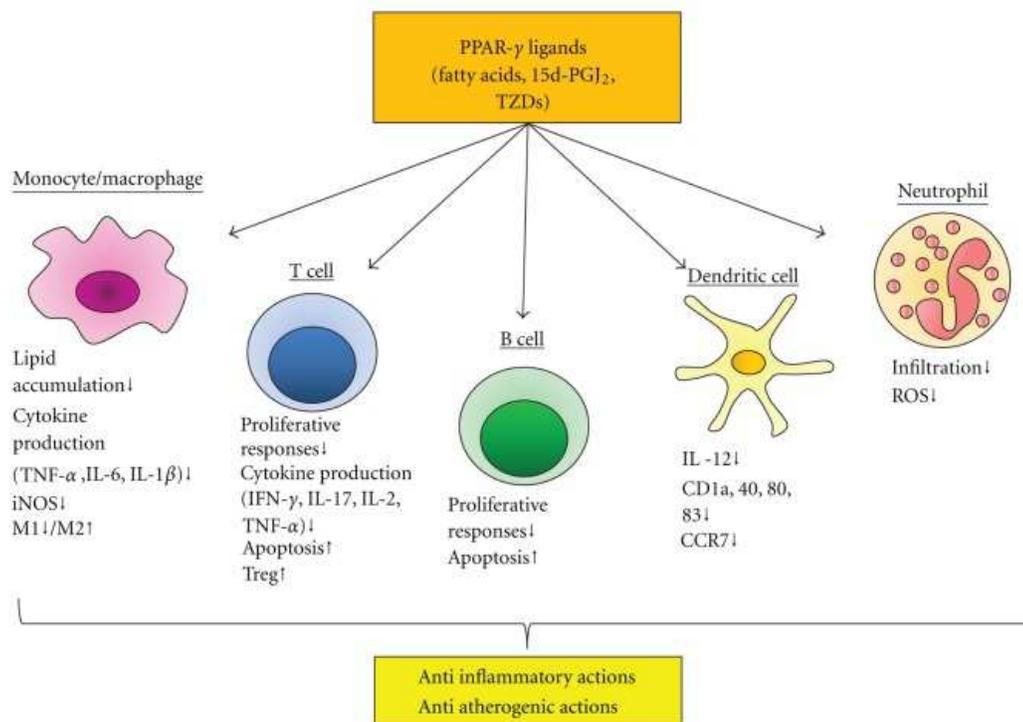
Figure 9. PPARγ activation leads to a variety of anti-inflammatory effects. Activation of PPARγ by agonist ligand introduction leads to a pronounced anti-inflammatory effect through a largely unknown pathway. These effects manifest in a variety of cell types and have been found to be exploitable to reduce chronic inflammation in the nervous system. Figure was adapted from Ohshima et al[190].

Due to their anti-inflammatory effects, several PPARγ synthetic ligands have been shown to have potential therapeutic value in treatment of neurodegenerative disease[183,191,192]. These conditions all cause a chronic inflammatory state and these ligands are thought to have beneficial effect through this transrepressive pathway. Examples of these effects include a reduction in insoluble Aβ42 load in rat brains with pioglitazone treatment[183]. Pioglitazone has also been shown to have beneficial effects on animal models of Parkinson's disease[192,193]. These effects are not explicitly due to full agonism. Multiple partial agonist ligands have been shown to be of benefit in therapy of neurodegenerative disease. In 2016 the partial agonist INT131, which has previously entered clinical trials as a type 2 diabetes therapeutic, began a new clinical trial as a potential therapeutic for multiple sclerosis. This has potentially opened a new pathway in

therapy for currently untreatable neurodegenerative diseases. However, as promising as some of the results observed may be there is still a large amount of work to be done before such a therapy would be suitable for human use.

1.7 Objectives of this work

The goal of this work is to gain a better understanding of the various mechanisms of PPARγ ligands at the molecular level. This information is incomplete for all PPARγ ligands and for certain ligand classes such as partial agonists almost nothing is known of their mechanism of action. Mechanistic information gained would be of great use in more intelligent drug design which could lead to the creation of therapeutics which retain beneficial effects while reducing negative side effects. The future goal would be to gain enough of an understanding of differential coregulator recruitment, coregulators are a highly diverse group which often lead to very different functional consequences, to develop so-called "biased ligands" which would preferentially recruit specific coregulator proteins. Biased ligands were first designed for the GPCR proteins and function through a preferential signaling of a single pathway instead of all outcomes of activation. Thus far they have shown great promise in providing similar therapeutic benefits as existing therapies but with significantly reduced side effects. To add to understanding of the conformational mechanisms of PPARγ ligands three projects were completed.

In the first project, which will be covered in chapter 2, the conformational effects of ligand binding on the C-terminal helix, helix 12, were examined. This helix is the most functionally significant portion of the AF-2 region in terms of overall coregulator affinity. So, a better understanding of how this region changes with the binding of a wide group of structurally and functionally distinct ligands is important to the understanding of how the protein is primed for coregulator binding. This study also allowed for better determination of the role of this helix in coregulator recruitment. For a long-time helix 12 has been the primary portion of the protein to be studied but it is clear that it is not the only significant region of the protein for activity.

The second project, described in chapter 3, had two goals. The first and most important was to identify a region of the protein that conveys the extreme selectivity seen in nuclear receptor and coactivator interactions. This is fundamental in the idea of creating a biased ligand since such bias would require the preferential recruitment of specific coactivators. As well the mechanism by which full agonists have differential efficacy was examined. The purpose of examining full agonism was not as much to understand full agonism, as this does not appear to be the optimal therapeutic pathway due to side effects. Rather the goal was to identify other regions of the protein where a particularly efficacious agonist might interact besides the direct helix 12 interaction. This could potentially allow the design of therapeutics that induce a more modest response in terms of protein activation and thereby have less deleterious side effects

The third project, described in chapter 4, examines one of the mechanisms of inverse agonism in PPARγ. To this point only a single mechanism is thought to occur leading to inverse agonism in nuclear receptors which involves the displacement of helix 12 from the AF-2 surface. This work identifies a novel mechanism which actually conveys greater inverse agonist efficacy, higher corepressor affinity, than the ligands which operate through the more standard pathway. Currently inverse agonists have not been studied much for their therapeutic value but there is some evidence they may be useful as cancer therapeutics through the arrest of the cell cycle. While these data may not be immediately useful for PPARγ research there may be value due to the high structural similarities among NRs. Which makes it quite possible that these mechanisms are conserved. This knowledge could be then utilized with other NR where inverse agonists are of great medical significance. If this is the case than it could lead to an entirely new route in drug design for these proteins.

Through the completion of this work and other work which can build on findings established herein it is hoped that a more fundamental understanding of this medicinally significant protein can be gained. This would allow for the design

of better therapeutic strategies for the treatment of the metabolic and neurodegenerative diseases in which PPARγ is directly or indirectly involved.

Chapter 2. Defining a conformational ensemble that directs activation of PPARγ

Ian M. Chrisman, Michelle D. Nemetchek, Ian Mitchelle S. de Vera, Jinsai Shang,

Zahra Heidari, Yanan Long, Hermes Reyes-Caballero, Rodrigo Galindo-Murillo,

Thomas E. Cheatham III, Anne-Laure Blayo, Youseung Shin, Jakob Fuhrmann,

Patrick R. Griffin, Theodore M. Kamenecka, Douglas J. Kojetin, and Travis S.

Hughes

2.1 Abstract

The nuclear receptor ligand-binding domain (LBD) is a highly dynamic

entity. Crystal structures have defined multiple low-energy LBD structural

conformations of the activation function-2 (AF-2) coregulator binding surface yet

it remains unclear how ligand binding influences the number and population of

conformations within the AF-2 structural ensemble. Here, we present a nuclear

receptor coregulator binding surface structural ensemble in solution, viewed

through the lens of fluorine-19 (^{19}F) nuclear magnetic resonance (NMR) and

molecular simulations, and the response of this ensemble to ligands, coregulator

peptides and heterodimerization. We correlate the composition of this ensemble

with function in peroxisome proliferator-activated receptor gamma (PPARγ)

utilizing ligands of diverse efficacy in coregulator recruitment. While the

coregulator surface of apo and partial agonist bound PPARγ is characterized by

multiple thermodynamically accessible conformations, the full and inverse agonist bound PPARγ coregulator surface is restricted to a few conformations which favor coactivator or corepressor binding, respectively.

2.2 Introduction

Nuclear receptors are ligand-regulated transcription factors that mediate the transcriptional actions of lipophilic endogenous ligands, including steroid hormones and lipids[194], and are the target of ~13% of US FDA approved drugs[195]. The binding of these natural ligands, as well as synthetic ligands and FDA-approved drugs, to the nuclear receptor ligand-binding domain (LBD) affects the recruitment of transcriptional coregulator proteins to target gene promoters, which influences chromatin remodeling and gene transcription[196]. Crystal structures of nuclear receptor LBDs have revealed in exquisite detail the molecular contacts created between the receptor and ligand, as well as low energy "active" and "inactive" conformations of helix 12[197–200]. Helix 12 is a critical regulatory structural element in the activation function-2 (AF-2) coregulator interaction surface of many nuclear receptors[201]. Over 100 crystal structures have been solved of the peroxisome proliferator-activated receptor gamma (PPARγ) LBD bound to ligands of various scaffolds and pharmacological activities[202]. Surprisingly, the backbone conformations of these structures, in particular the conformation of helix 12, are all very similar despite the fact that PPARγ is bound to ligands that produce a diverse range of functional outputs. Thus, it is difficult to understand the structural mechanism of action by which the binding of ligands with diverse activities affect helix 12 conformation from crystallography data alone. One

43

hypothesis is that helix 12 consists of a dynamic ensemble of conformations, and not one single or "static" conformation in the presence or absence of a bound ligand[138,203]. However experimental evidence describing this ensemble is lacking and it remains poorly understood how binding of pharmacologically distinct ligands affects the ensemble of coregulator binding surface and helix 12 conformations.

Solution structural methods indicate prevalent, ligand dependent helix 12 movement. Hydrogen deuterium exchange mass spectrometry (HDX-MS) demonstrates a relationship between helix 12 stability and agonist binding for nuclear receptors[161,204–206]. NMR studies implicate movement on the microsecond-millisecond (μs-ms) time scale between two or more conformations over a large portion of the apo and partial agonist bound PPARγ LBD. These movements result in very broad or unobserved NMR resonances that prohibit structural analyses. Full agonists robustly diminish these dynamics[140,207–209]. Furthermore crystal structures, HDX-MS, and protein NMR have provided complementary information revealing a relationship between structure and function for PPARγ (e.g. the presence or absence of critical hydrogen bonds between ligand and helix 12[22]) however a direct observation of the ligand dependent ensemble implied by these data is lacking. This raises the question: are there multiple long-lived conformations that correlate with functional efficacy (e.g. coregulator affinity) in nuclear receptors?

It remains challenging to quantify the number, relative population and kinetics of exchange between the conformations that compose this putative

ensemble and how the ensemble is influenced by binding small molecules and coregulators. ^{19}F (fluorine-19) NMR spectroscopy is exceptionally sensitive to structural and environmental changes, can reveal structural information from regions of a protein that are unobserved via 2D/3D NMR[210], and can be used to probe how ligands affect the conformational ensemble of proteins[211–215]. Here, using ^{19}F NMR combined with biochemical coregulator interaction analysis and molecular simulations, we define the ligand dependent conformational ensemble of the coregulator interaction surface, including helix 12, which controls the transcriptional activity of PPARγ. The data presented here indicate that helix 12 and the coregulator binding surface of apo and partial agonist bound PPARγ is found in a broad energy well with multiple local minima of similar potential energy separated by relatively small kinetic barriers, allowing exchange on the μs to ms time scale. In contrast, when PPARγ is bound to a full agonist or inverse agonist, helix 12 and the coregulator binding surface occupies narrow energy wells with fewer thermodynamically accessible conformations. In addition, simulations define some of the probable structures that compose these ensembles. These data better elucidate how ligands induce functional effects via nuclear receptors.

2.3 Results

2.3.1 Diverse activities of synthetic PPARγ ligands

We assembled a set of 16 pharmacologically distinct PPARγ ligands that we and others developed or previously characterized in cellular and structure-function studies (**Supplementary Fig. 16**). The set of ligands includes full and

partial agonists that robustly or mildly enhance transcriptional activation; antagonists/non-agonists that block activation or maintain constitutive basal cellular transcriptional activity; and inverse agonists that repress transcription compared to the basal receptor activity. In a time-resolved fluorescence resonance energy transfer (TR-FRET) coregulator interaction assay, full agonists such as GW1929 and rosiglitazone that induce robust PPARγ transcription[142,216], increase binding of a peptide derived from Mediator of RNA polymerase II transcription subunit 1 (MED1) coactivator (**Figure 1a**), and decrease binding of a peptide derived from the nuclear receptor corepressor 1 (NCoR1) referred to herein as NCoR (**Figure 1b**). In contrast, ligands that function as inverse agonists such as T0070907 and SR10221 lower PPARγ transcriptional responses relative to basal cellular activity[156,161,217], increase binding of NCoR (**Figure 1b**), and decrease binding of MED1 (**Figure 1a**). Crystal structures of ligand-bound PPARγ LBD typically show a non-crystallographic dimer configuration containing two chains, A and B (**Supplementary Fig. 2a**). The main difference between the two chain A and chain B protein molecules involves helix 12, which adopts distinct conformations commonly referred to as "active" and "inactive". The "active" conformation is assumed to be the conformation in solution when bound to a full agonist that induces increased transcription. Notably, these "active" and "inactive" helix 12 conformations are both influenced by crystal contacts (**Supplementary Fig. 2b**).
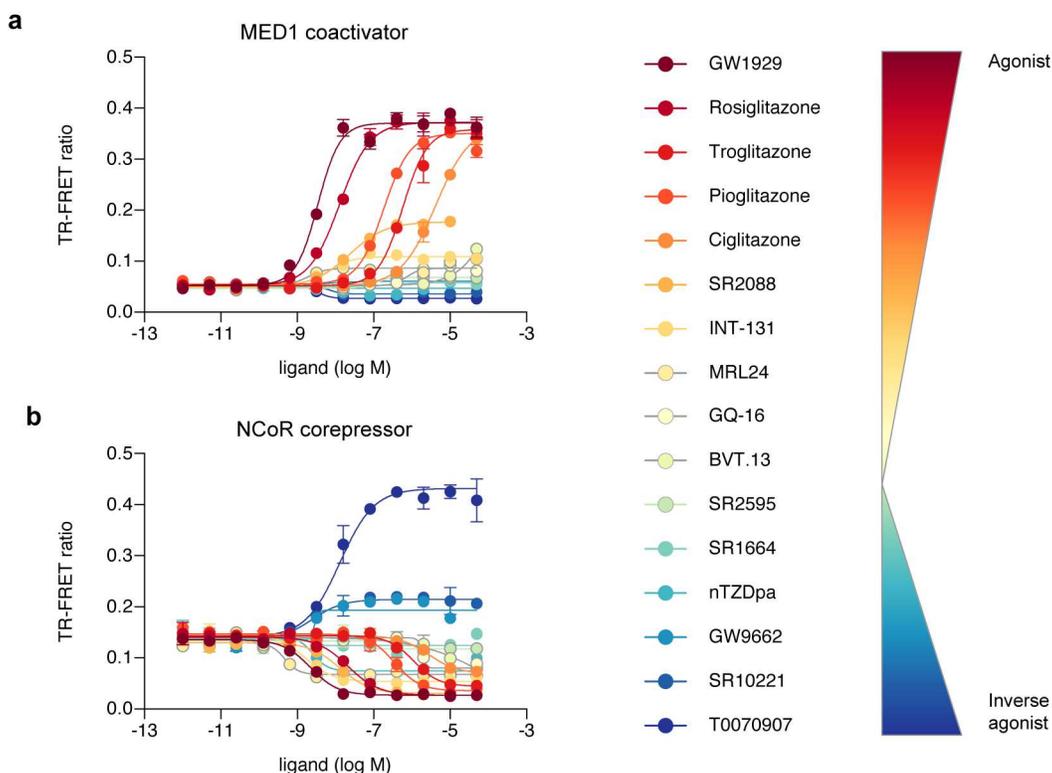
Figure 1. Differences in nuclear receptor coregulator interaction differentiates pharmacologically distinct synthetic PPARγ ligands. TR-FRET biochemical assay shows the effect of the compounds on the interaction between PPARγ LBD and peptides derived from the (**a**) MED1 coactivator and (**b**) NCoR corepressor, plotted as TR-FRET ratio (665 nm/620 nm) vs. ligand concentration (n=2, standard deviation). The data shown represents technical replicates from a single experiment and the experiment was repeated four times with similar results. The window of efficacy in these data is representative of ligand induced changes in coregulator affinity for PPARγ. An increase in TR-FRET ratio indicates a strengthening of affinity for MED1/NCoR compared to apo while a decrease indicates a ligand induced weakening of affinity for the coregulator. The effect of vehicle (DMSO) is negligible (**Supplementary Fig. 5**), furthermore the DMSO concentration is constant across the titration both in this figure and all other TR-FRET data presented.

2.3.2 NMR-detected coregulator binding surface structural ensemble

To facilitate ^{19}F NMR studies of the coregulator binding surface, which is composed of portions of helix 3, 4 and 12, we introduced a cysteine residue on the C-terminus of helix 12 at several locations (K502C, Y505C and Q498C) and on helix 3 (Q322C) of the PPARγ ligand binding domain (LBD) to allow covalent linkage of 3-Bromo-1,1,1-trifluoroacetone (BTFA), a small molecule containing a

trifluoromethyl ($-CF_3$) group. Q498C caused protein instability, whereas BTFA attached to Y505C did not show pronounced ligand induced changes to the ^{19}F NMR spectra of the PPARγ LBD, presumably due to its position at the unstructured C-terminus of helix 12 (**Supplementary Fig. 1**). However, Q322C and K502C yielded well-functioning protein with pronounced ligand inducible changes; we used these mutants to probe the conformational ensemble of the PPARγ coregulator binding surface (i.e. the AF-2 surface).

Molecular simulations indicate that K502 (wild-type residue) and K502C-BTFA (modified residue) are both solvent exposed in the "active" helix 12 conformation and are not likely to sterically hinder coregulator binding in this "active" conformation (**Figure 2, Supplementary Fig. 3**). Control experiments indicate that the introduced cysteine on helix 12 (C502) is preferentially labeled over the only native cysteine (C313), which could be because C313 points into the ligand binding pocket (**Supplementary Fig. 4**). We refer hereafter to BTFA labeled PPARγ K502C protein as PPARγ^{K502C}-BTFA, we also use PPARγ^{K502C}-BTFA with a C313A mutation (referred to as PPARγ^{C313A,K502C}-BTFA) for comparison, as this protein can only be labeled on helix 12. ^{19}F NMR signals in these labeled proteins could conceivably arise from BTFA labeled co-purified protein impurities, however, the spectrum of PPARγC313A, which lacks all cysteines, reveals no detectable signal from impurities (**Supplementary Fig. 5**).

Next, we determined the effects of mutations and labeling on the function of the PPARγ LBD. First, we measured ligand K_i values using a competitive ligand displacement assay to determine if the mutations and BTFA label affects

ligand affinity. Compared to wild-type, the PPARγ[K502C]-BTFA and PPARγ[C313A,K502C]-BTFA proteins exhibit a 5- and 11-fold median reduction in ligand affinity respectively (**Supplementary Table 1**) however, only four ligands (GQ-16, BVT.13, ciglitazone and troglitazone) are predicted to be at less than 93% occupancy in our samples under the conditions used for the NMR experiments reported here (**Supplementary Table 2**). Second, we measured coregulator recruitment efficacy and EC_{50} using TR-FRET. As expected from the calculated K_i values, PPARγ[K502C]-BTFA has a 2- and 5-fold median reduction in EC_{50} and PPARγ[C313A,K502C]-BTFA has a 1- and 9-fold median reduction in EC_{50} in recruiting MED1 and NCoR respectively compared to wild-type PPARγ (**Supplementary Table 3**). Importantly, ligand induced coregulator recruitment efficacy is highly correlated for PPARγ[K502C]-BTFA and WT indicating that PPARγ[K502C]-BTFA is functionally similar to wild type PPARγ LBD (R^2=0.8 for NCOR and R^2=0.98 for MED1; **Supplementary Fig. 6d and 6f**). In contrast, the relative NCOR recruitment efficacy differs between PPARγ[C313A,K502C]-BTFA and WT, while the relative MED1 recruitment efficacy is comparable to WT in this labeled mutant (R^2=0.11 for NCOR and R^2=0.8 for MED1; **Supplementary Fig. 6d and 6f**). We also found that the mutations and labeling had little effect on recruitment efficacy and EC_{50} of a coactivator peptide derived from CREB-binding protein (CBP) for both labeled forms of PPARγ (**Supplementary Fig. 6b and 6e** and **Supplementary Table 4**).

In addition, we compared mutant and WT PPARγ affinity for fluorescein isothiocyanate (FITC) labeled NCoR, MED1 and a peptide from the silencing

mediator of retinoic acid and thyroid hormone receptor (SMRT) corepressor utilizing fluorescence polarization (FP). Labeling did not significantly affect affinity of NCoR or MED1 for the four tested PPARγ-ligand complexes (**Supplementary Fig. 7**) and consistent with the TR-FRET data, PPARγ^{K502C}-BTFA affinity for NCOR and MED1 correlated most closely with WT PPARγ. Both labeled PPARγs affect SMRT peptide affinity, although PPARγ^{K502C}-BTFA does so to a lesser extent, indicating that the label may directly interfere with SMRT binding (**Supplementary Fig. 7**). In general, these data indicate that the relative coregulator affinities are consistent between WT and especially PPARγ^{K502C}-BTFA. Given these data we focused on PPARγ^{K502C}-BTFA to correlate structure with function and utilize PPARγ^{C313A,K502C}-BTFA to confirm specific labeling of PPARγ^{K502C}-BTFA and in cases where increased signal to noise is required. PPARγ^{K502C}-BTFA is incompletely labeled (**Supplementary Fig. 4d**) to avoid labeling C313 (**Supplementary Fig. 4a-c**), which decreases the NMR signal. In addition, PPARγK502C that is not labeled with BTFA is not TR-FRET active (**Supplementary Fig. 4e**) but has the same affinity for coregulators as PPARγ^{K502C}-BTFA (**Supplementary Fig. 7b**), therefore any unlabeled portion of PPARγ^{K502C}-BTFA should not affect FP or TR-FRET results.
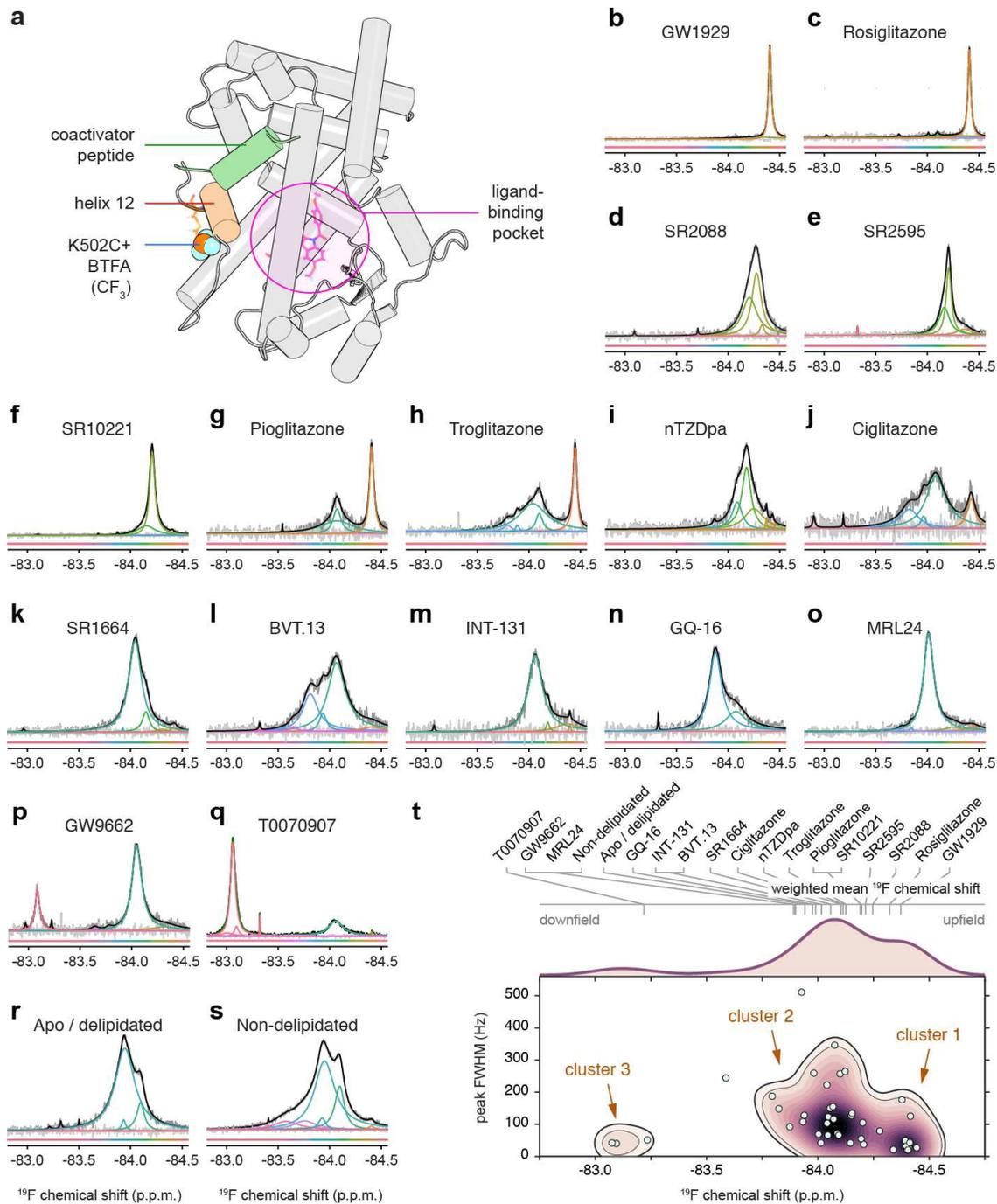
Figure 2 | ^{19}F NMR analysis of PPARγ^{K502C}-BTFA bound to 16 pharmacologically distinct ligands. (**a**) Location of covalently attached BTFA tag. (**b-s**) ^{19}F NMR spectra (medium grey lines) of PPARγ^{K502C}-BTFA (**b-q**) bound to ligand and ordered according to mean weighted ^{19}F chemical shift or (**r**) delipidated and (**s**) non-delipidated apo-protein. Fitted peaks are colored according to fitted ^{19}F chemical shifts; deconvoluted spectra and residuals are shown in black and light grey lines, respectively. Some of these spectra have been replicated (**Supplementary Fig. 10**) (**t**) Plot of the full width half max (FWHM) of the major fitted peaks (only peaks that comprise >5% of the spectrum area are included) versus fitted chemical shifts. Clusters were detected by

bivariate kernel density estimation (purple contours) with weighted mean ^{19}F chemical shift values annotated on top of the plot. Computational simulations were run of labeled PPARγ (PPARγ^{K502C}-BTFA) bound to one of these 16 ligands (GW1929) that demonstrate that the tag stays on or near the surface of the protein (**Supplementary Fig. 3**)

We collected ^{19}F NMR spectra of PPARγ^{K502C}-BTFA bound individually to each of the 16 synthetic ligands (**Figure 2b–q**), as well as spectra of apo-protein with and without delipidation (**Figure 2r,s**). Bacterially expressed PPARγ often contains fatty acids, which can have functional effects[218] and therefore we remove these lipids, additional data regarding the effects of delipidation on structure and function are presented below. ^{19}F NMR provides a time-averaged view of the conformational ensemble of helix 12, thus long-lived major conformations (>ms lifetime) show up as distinct peaks, while conformations with lifetimes on the μs-ms timescale show up as a broad single peak. A single conformation (i.e. fast exchange among minor conformational variants) produces a single narrow peak. We used an objective deconvolution[219] method to determine the number of peaks and corresponding peak line widths that compose the recorded spectra. The line widths of the peaks obtained from the objective deconvolution are consistent with measured NMR lifetimes of the ^{19}F (T$_2$) that we obtained for several liganded states (**Supplementary Table 5**). Bivariate kernel density estimation of all the deconvoluted spectra revealed two primary clusters of peaks, or conformations and a third lowly populated cluster (**Figure 2t**). In general, ligands known to be efficacious or strong agonists (e.g. GW1929, rosiglitazone, pioglitazone, and troglitazone) populate the most upfield (right) group of the narrowest peaks in cluster 1. Ligands of other pharmacological types, including partial agonists and inverse agonists are found in cluster 2, which is

52

composed of wider peaks, indicative of ensembles composed of multiple conformations. A third lowly populated cluster of downfield chemical shifts and relatively narrow peak widths, cluster 3, occurs for T0070907 and GW9662 ligands that covalently bind to C313 in the canonical ligand-binding pocket and function as inverse agonists in cell-based assays[156,217]. These covalent ligands do not attach to the introduced cysteine on helix 12 (C502; **Supplementary Fig. 5a-b**). PPARγ[C313A,K502C]-BTFA loaded with or without the same 16 ligands show very similar ^{19}F NMR spectra to PPARγ[K502C]-BTFA, except for ligands which covalently bind to C313, which as expected look similar to ligand-free/apo-protein because PPARγ[C313A,K502C]-BTFA lacks C313 (**Supplementary Fig. 8**). Overall these data indicate a diverse ligand dependent helix 12 conformational ensemble.

A contribution to the ligand-dependent differences in ^{19}F NMR chemical shift values could be the relative solvent exposure of the BTFA probe $-CF_3$ group, which can be determined by collecting ^{19}F NMR spectra as a function of D_2O concentration. When increasing amounts of D_2O are added to the sample, a large upfield shift (i.e., to the right) of the ^{19}F NMR peak indicates high solvent exposure of the BTFA probe while a smaller shift indicates low solvent exposure[215], for example a solvent exposed 5-fluorotryptophan shifted 0.217 ppm upfield in 90% D_2O buffer[220]. The chemical shift of free BTFA, which has high solvent exposure, has nearly the largest chemical shift change of any measured here (0.115 ppm upfield), whereas the solvent protected MRL24 ligand CF_3

group, which is buried deep in the ligand binding pocket, moves in the opposite

direction with increasing D_2O concentrations (**Figure 3 and Supplementary Fig.**

13). Given the solvent protected position of MRL24 CF_3 in the ligand binding

pocket[221], the movement in the opposite direction of the MRL24 ligand CF_3 group

likely indicates interaction between the ligand CF_3 group and the protein that is

perturbed upon addition of D_2O[32]. GW1929 bound PPARγ[C313A,K502C]-BTFA has

the largest upfield chemical shift (0.116 ppm; **Figure 3**), which is consistent with

molecular simulations that show that K502 and C502 both have considerable

solvent exposure in the "active" conformation (**Supplementary Fig. 3**). These

data indicate full agonists primarily populate the "active" crystalized

conformation in solution, while apo, partial agonists and inverse agonists are

found in other distinct conformations.
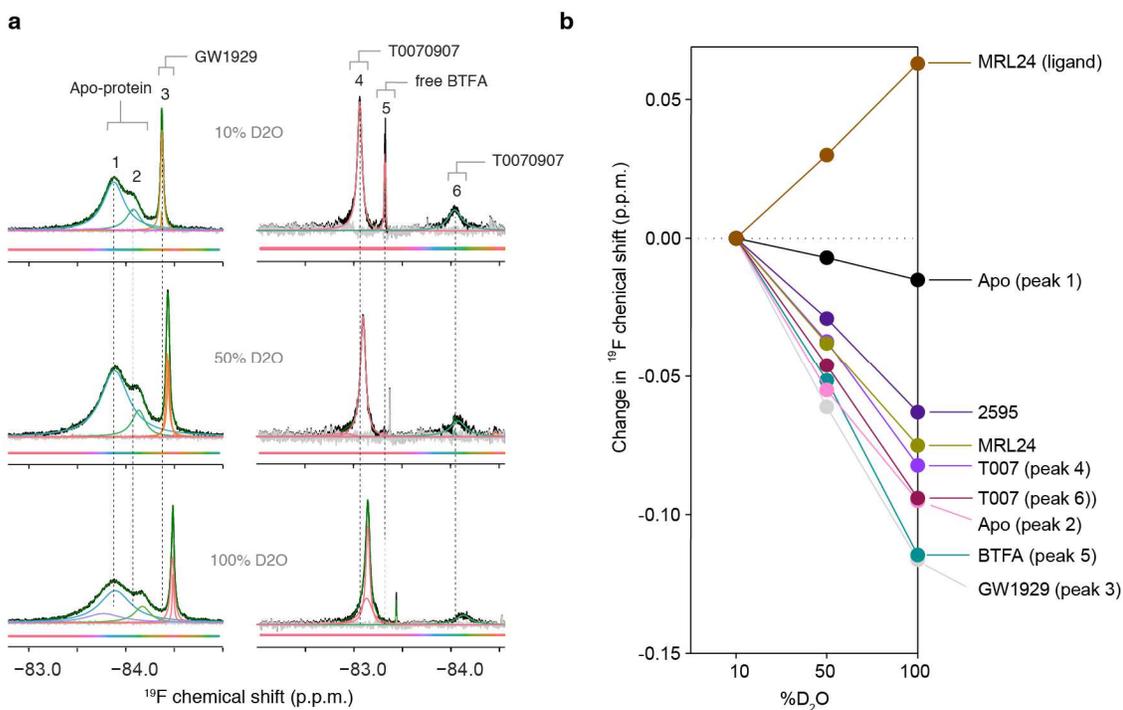
Figure 3 | Helix 12 solvent exposure is distinct for the "active" helix 12 conformation, apo-protein and inverse agonist-bound conformations. (a) ^{19}F NMR spectra of delipidated apo-PPARγ^{C313A,K502C}-BTFA bound to 0.2 molar equivalent of GW1929 (left); and T0070907-bound PPARγ^{K502C}-BTFA with excess free BTFA; at the indicated concentrations of D$_2$O. Due to the high binding affinity of GW1929 (4 nM), all three expected peaks (two apo-protein peaks and one GW1929 peak) are present in slow exchange on the NMR time scale, allowing analysis of the three conformations simultaneously. **(b)** D$_2$O solvent isotope-induced changes, plotted as % D$_2$O versus change ^{19}F NMR chemical shift values for various peaks; a dotted grey line is shown to highlight no D$_2$O-induced change in ^{19}F NMR chemical shift. These experiments were performed once.

PPARγ protein expressed in bacteria can pull down medium chain fatty acids, which function as weak partial agonists[218]. In agreement, non-delipidated (native) apo-protein that is partially bound to endogenous *E. coli* lipids (**Figure 2s**) afforded ^{19}F NMR spectra with two peaks of similar chemical shift to delipidated apo-protein (**Figure 2r**) along with several other lowly populated peaks. Comparison of 2D NMR spectra of delipidated and non-delipidated PPARγ LBD and PPARγ^{C313A,K502C}-BTFA in apo form or bound to ligands indicate that *E. coli* lipids have relatively minor effects on helix 12 and backbone structure (**Supplementary Fig. 9a-b**). In addition, TR-FRET data collected using non-delipidated or delipidated PPARγ-LBD shows essentially the same ligand dependent coregulator recruitment potency and efficacy (**Supplementary Fig. 9c and 9d**). In separate protein preparations we did observe some variation in T0070907 bound spectra (**Supplementary Fig. 10c**) which could be due to variable amounts of residual co-bound lipids remaining after delipidation as PPARγ has a very large ligand binding pocket that can accommodate more than one bound ligand[208]. Overall, these data indicate that some inverse agonists do more than simply displace activating lipids, but instead induce a distinct PPARγ LBD state with higher affinity for corepressors than delipidated apo protein.

We probed ligand induced changes in another region of the PPARγ coregulator binding surface using a Q322C mutation, which is located in helix 3, generating PPARγ^{C313A,Q322C}-BTFA. The fluorine probe in this variant is solvent exposed in the "active" conformation and maintains wild type like recruitment of coregulators (**Figure 4 and Supplementary Fig. 11**). The chemical shift difference induced by ligands is relatively small; however, just as with the helix 12 probe, ligands decrease the conformational diversity of this region with the strongest agonists (rosiglitazone and GW1929) yielding the narrowest peaks indicative of a single main conformational ensemble. Similar to apo PPARγ^{K502C}-BTFA (**Figure 2r**), apo PPARγ^{C313A,Q322C}-BTFA produces a spectrum with two main wide peaks, indicative of two sub ensembles composed of multiple conformations with slow exchange between the sub ensembles (**Figure 4**). These data indicate that the ligand free coregulator binding surface is composed of two main structurally diverse (i.e. wide NMR peak) ensembles in slow exchange. Alternatively, one of the peaks could be PPARγ bound to residual *E.coli* lipids, however this unlikely given that the spectra is stable over time (**Supplementary Fig. 11**) whereas *E.coli* lipid bound PPARγ^{K502C}-BTFA spectra change with time (**Supplementary Fig. 10a**). The exchange rate between these two apo NMR peaks is discussed below.
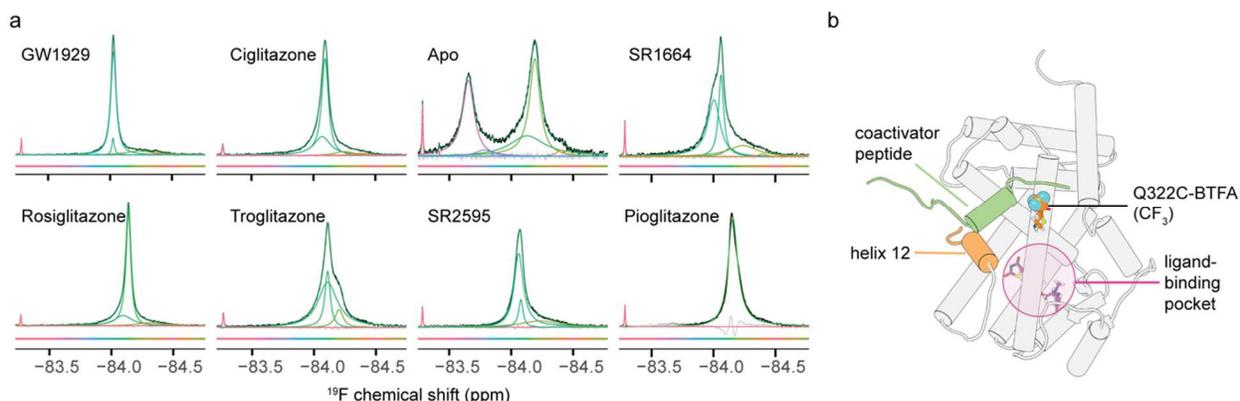
Figure 4 | A helix 3 probe on the coregulator binding surface confirms agonist induced reduction in conformational complexity of the coregulator binding surface. a) Fluorine NMR spectra of PPARγ^{C313A,Q322C}-BTFA bound to the indicated ligands. The small sharp left shifted peak in all the spectra is free BTFA. b) Trajectory frame from a simulation of PPARγ^{C313A,Q322C}-BTFA bound to a coactivator peptide (MED1; green) with 322C-BTFA shown in orange as spheres (fluorine atoms are turquoise). These experiments were performed once.

2.3.3 Slow exchange between ensemble structures

A single well populated ^{19}F NMR peak is observed for BTFA probes placed in two areas of the coregulator binding surface (helix 3 and 12) when PPARγ-BTFA is bound to a strong agonist such as GW1929 or rosiglitazone. However, for less efficacious ligands, multiple-well populated ^{19}F NMR peaks are observed in slow exchange on the NMR time scale indicating that the peaks represent distinct coregulator binding surface conformations with lifetimes on the order of milliseconds or longer.

To confirm the apparent slow exchange between different conformations, we performed ^{19}F chemical exchange saturation transfer (CEST) NMR experiments on PPARγ where multiple peaks, or conformations, are present (**Figure 5a and Supplementary Fig. 12**). Of the five ligands and apo protein studied, four ligands and apo protein show obvious slow chemical exchange

between a well-populated minor resonance and the most abundant resonance. In the case of T0070907, exchange was difficult to detect, likely because exchange rates less than ~0.4 s^{-1} are too slow to be effectively detected using this method under our experimental conditions (**Supplementary Fig. 12a and 12b**).

We next quantified the exchange rate between peaks. PPARγ^{K502C}-BTFA bound to GW9662 and troglitazone each showed very slow exchange between peaks, making precise measurement difficult (0.3 and 0.4 s^{-1}; Supplementary Fig. 12d). However, for pioglitazone and ciglitazone, the exchange between the two prominent peaks is 1.2 and 1.4 s^{-1} (95% CI 1.1 to 1.3 s^{-1} for pioglitazone and 1.3 to 1.5 s^{-1} for ciglitazone; Figure 5b). In addition, we found exchange of 1.0 s^{-1} (95% CI 0.9 to 1.1 s^{-1}) between the two resolved apo peaks of PPARγ^{C313A,Q322C}-BTFA (**Supplementary Fig. 12c**). These slow exchange rates are consistent with the notion that the detected conformations originate from larger scale movements involving many atoms[223], indicating helix 12 and the coregulator binding surface exchanges between two or more distinct conformations. Almost all of the signal in the ligand bound spectra do not overlap with the apo spectrum, thus implying that exchange between conformations is occurring while bound to ligand. These data raise the possibility that these ligands are found in at least two functionally distinct main conformations exchanging on the seconds time scale. This idea is consistent with the fact that the PPARγ non-agonist/antagonist SR1664 has been crystalized in two distinct conformations[161,224].
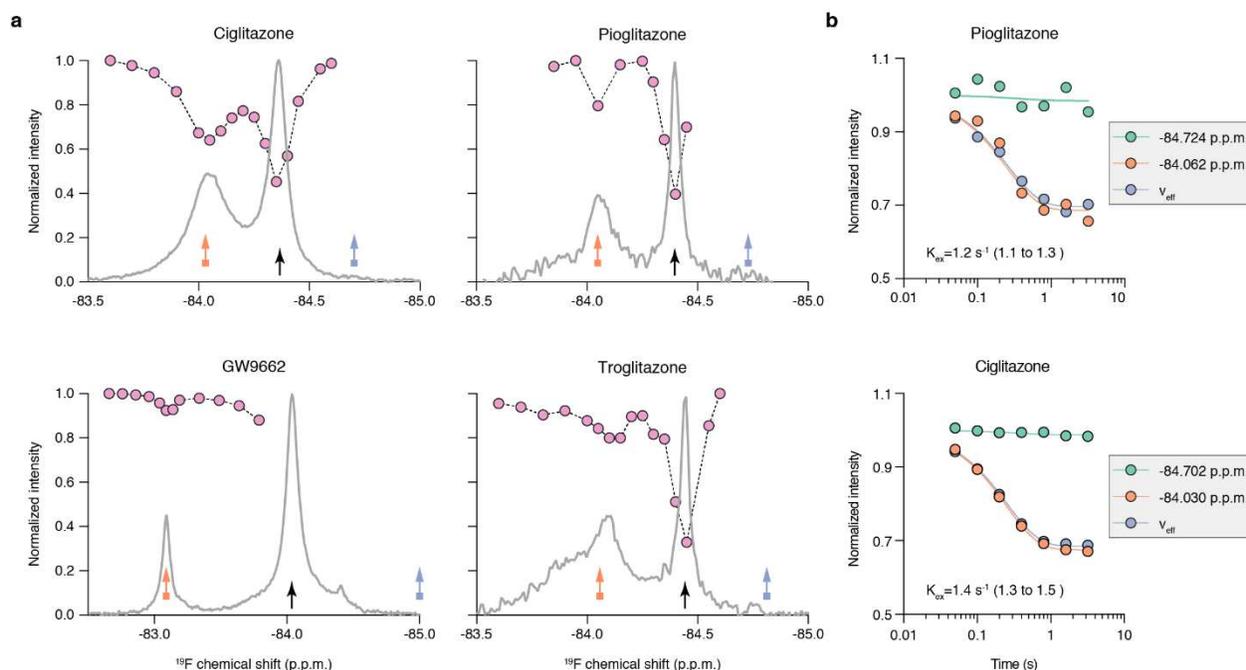
Figure 5 | Chemical exchange saturation transfer indicates slow exchange between liganded states with multiple helix 12 conformations. (**a**) A selective Gaussian pulse was used to saturate the ^{19}F spectra bound to the indicated ligands at locations indicated by pink circles. The height of the pink circles indicates the height of the peak indicated by the black arrow when the selective pulse was carried out at the chemical shift of the pink circle. If exchange is occurring between the two peaks the pink dots should mirror the spectrum. (**b**) A selective Gaussian pulse was used to saturate the spectrum at an on resonance (**a**; orange box arrow) and off resonance (**a**; blue box arrow) location, the peak height of the most abundant resonance (**a**; black arrow) was monitored as a function of the duration of the saturation pulse, and the resulting peak intensities were fit to extract the exchange rate. 95% confidence intervals for the fit of the calculated exchange rates are shown in parentheses. PPARγ^{K502C}-BTFA was used for these data except for ciglitazone which used PPARγ^{C313A,K502C}-BTFA for increased signal to noise. These experiments were performed once.

2.3.4 Connecting the helix 12 structural ensemble to function

To test whether the ^{19}F NMR detected helix 12 conformations correlate to function within our set of 16 ligands, we compared the weighted mean ^{19}F NMR chemical shift values to ligand efficacy for recruitment of MED1 (coactivator), NCoR (corepressor) and for a subset of these ligands, CBP (coactivator) in the TR-FRET coregulator interaction assay. For PPARγ^{K502C}-BTFA there is a correlation between mean ligand induced ^{19}F NMR chemical shift and ligand induced MED1, CBP and NCoR recruitment efficacy (**Figure 6a)**. There are

similar correlations for PPARγ^{C313A,K502C}-BTFA (**Supplementary Fig. 14**) but not

for PPARγ^{C313A,Q322C}-BTFA. In addition, we used fluorescence polarization (FP)

to measure the affinity of a subset of PPARγ-ligand complexes for coregulators

(MED1, NCoR, and SMRT) and compared these to labeled and wt PPARγ and

found that ligand specific affinity correlates with NMR chemical shift (**Figure 6b**

and **Supplementary Fig. 14**). The most efficacious agonist for MED1 and CBP

binding in our ligand set (GW1929) induces the most upfield mean chemical shift

of the PPARγ^{K502C}-BTFA probe. Similar trends are observed for other agonists,

including the right shifted peaks of thiazolidinedione (TZD/glitazones) ligands. In

contrast, downfield mean chemical shifts are prevalent in apo-protein and inverse-

agonist (T0070907) bound PPARγ^{K502C}-BTFA, which show the highest efficacy

and affinity for NCoR and SMRT binding (**Figure 6**). These data indicate that the

conformations detected by ^{19}F NMR are functionally distinct.
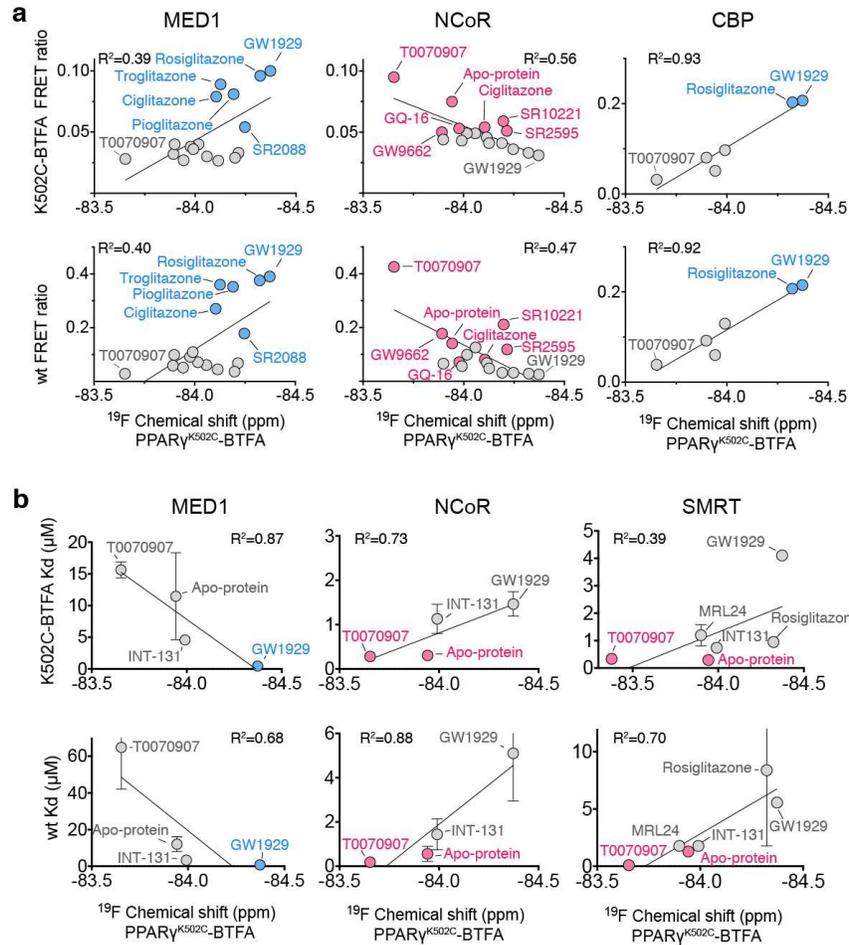
Figure 6 | Ligand-directed helix 12 ensemble dictates PPARγ-coregulator interaction. (a) Plot of mean ^{19}F NMR chemical shift values (PPARγ^{K502C}-BTFA) versus TR-FRET endpoint data for the recruitment of MED1, NCoR and CBP peptides to PPARγ^{K502C}-BTFA (top panels) and wt PPARγ LBD (bottom panels) for the set of 16 pharmacologically distinct synthetic PPARγ ligands and apo-protein (select ligands for CBP). Error bars are relatively small for end point TR-FRET values (**Figure 1** and **Supplementary Figure 6**) and excluded for clarity b) Plot of mean ^{19}F NMR chemical shift values (PPARγ^{K502C}-BTFA) versus MED1, NCoR and SMRT peptide dissociation constant (K_d) for PPARγ^{K502C}-BTFA (top panels) and wt PPARγ LBD (bottom panels) as measured by fluorescence polarization for a subset of the ligands in panel a. Linear regression fit is shown as a solid line and the correlation coefficient (R^2) for the fitted line is indicated. The ligands with the highest efficacy for MED1 and NCoR recruitment for PPARγ^{K502C}-BTFA are highlighted. Error bars represent standard deviation of two (K502C-BTFA K_d and wt SMRT K_d) or three (wt MED1 and NCoR K_d) independent experiments.

2.3.5 Coregulators shift the conformational ensemble

The data above demonstrate that ligands of different pharmacological

activities can stabilize functionally distinct coregulator binding surface

conformations. To determine how coregulator binding influences these ligand

dependent conformational ensembles in solution, we performed ^{19}F NMR with

and without coactivator (MED1 and CBP) and corepressor (SMRT and NCoR)

peptides; and in the absence or presence of the most efficacious agonist

(GW1929) or inverse agonist (T0070907) and a less efficacious inverse

agonist/antagonist (GW9662) (**Figure 7a**). Similar to strong agonists, addition of

coactivators to apo-protein induces an upfield (right) shifting of the spectra. In

contrast, similar to inverse agonists, addition of corepressors to apo-protein

induces a downfield (left) shifting of the spectrum. Notably, corepressor co-

binding to T0070907/PPARγ^{K502C}-BTFA results in smaller perturbations to the

fluorine spectra than coactivator co-binding. The opposite is observed for

coregulator co-binding to GW1929/ PPARγ^{K502C}-BTFA. Apo-protein and

GW9662 bound PPARγ^{K502C}-BTFA are almost equally changed by addition of

coactivators or corepressors (**Figure 7c**). These data suggest that the

GW1929/PPARγ and T0070907/PPARγ coregulator binding surface structural

ensembles are near ideal for MED1/CBP and NCoR//SMRT binding,

respectively. In contrast, the apo-protein and GW9662 structural ensembles are

not ideal for binding any of these coregulators.
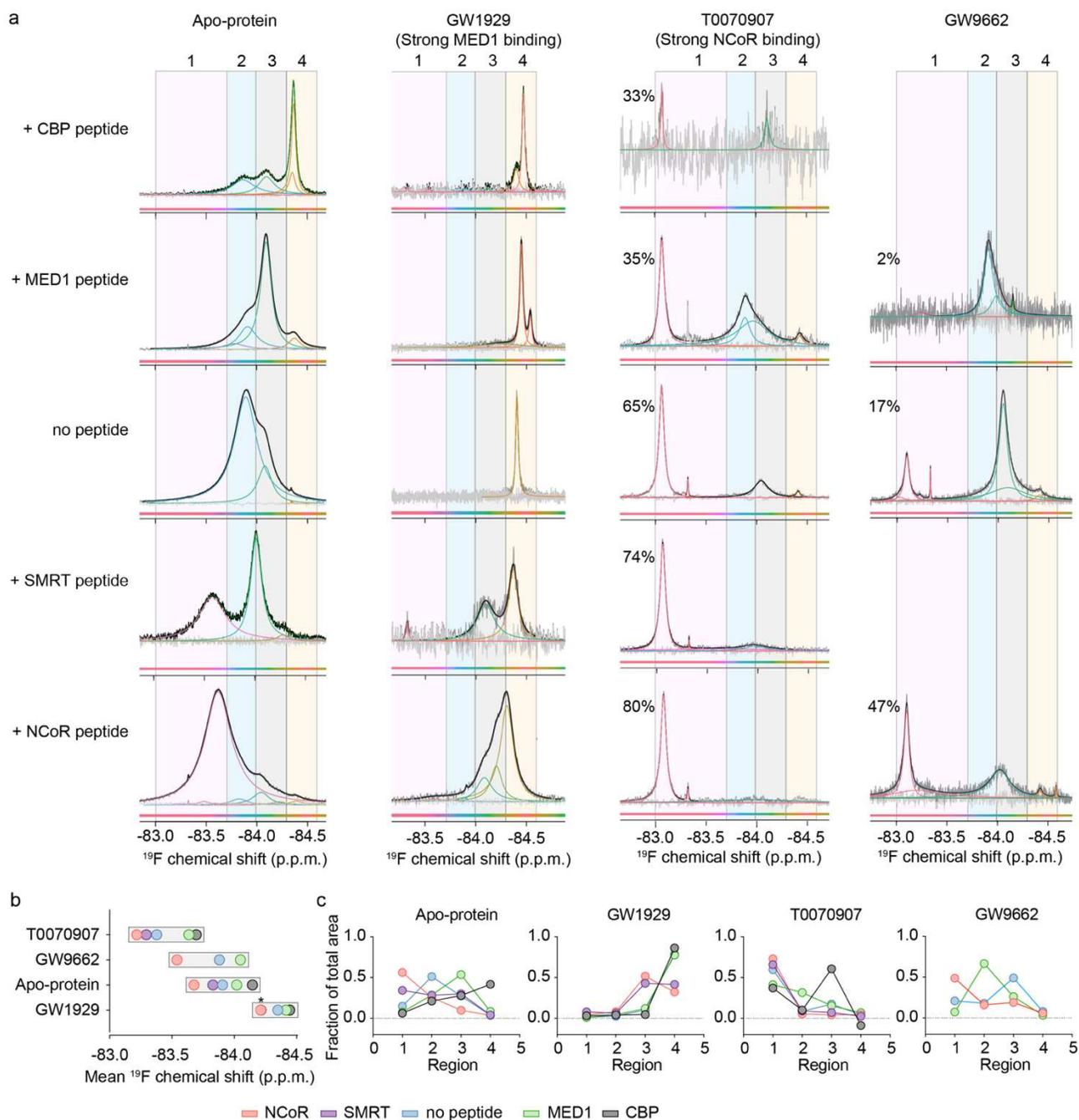
Figure 7 | Coregulator binding shifts the helix 12 conformational ensemble in a manner predicted by ligand induced shifts to the ensemble. (**a**) Deconvoluted ^{19}F NMR spectra of apo-PPARγ^{C313A,K502C}-BTFA and PPARγ^{K502C}-BTFA bound to GW1929 or T0070907 in the absence and presence of MED1 coactivator or NCoR corepressor peptides. The percent of total signal area found in the left sharp peak is shown for T0070907 and GW9662 spectra. The small sharp peak at ~-83.3 ppm is free BTFA. (b) Mean weighted chemical shift values from the plots in (a). (c) Fraction of the total peak areas in the four colored boxed regions in (a), which roughly correspond to the clustered spectral regions from Figure 2; with the two middle regions corresponding to the two peaks observed for apo-protein. Agonist bound PPARγ is changed little by MED1 binding, whereas NCoR binding changes the spectrum drastically and vice versa for inverse agonist

63

(T0070907) bound PPARγ. *SMRT induced mean chemical shift in GW1929 bound PPARγ^{K502C}-BTFA is the same as NCoR. These experiments were performed once.

2.3.6 RXRα heterodimerization shifts the inverse agonist ensemble

PPARγ binds enhancer regions on DNA and affects gene expression primarily as a heterodimer with RXRα[225]. We therefore performed experiments to determine the effect of RXRα heterodimerization on the conformational ensemble and on coregulator recruitment. Other than the expected broadening and consequently decreased signal from the larger molecular weight complex, RXRα heterodimerization has a relatively small effect on the ^{19}F spectra of PPARγ^{C313A,K502C}-BTFA when co-bound to GW1929 (agonist), MRL24 (less efficacious partial agonist) and apo. RXRα binding to PPARγ^{K502C}-BTFA co-bound to T0070907 (inverse agonist) induces a large change in the spectrum, decreasing the relative population found in the left shifted peak (**Figure 8a**), which is expected to disfavor NCoR binding. Consistent with the change in ^{19}F NMR spectrum, NCoR binding is decreased and MED1 recruitment is increased by RXRα heterodimerization with the T0070907/PPARγ complex (**Figure 8**).
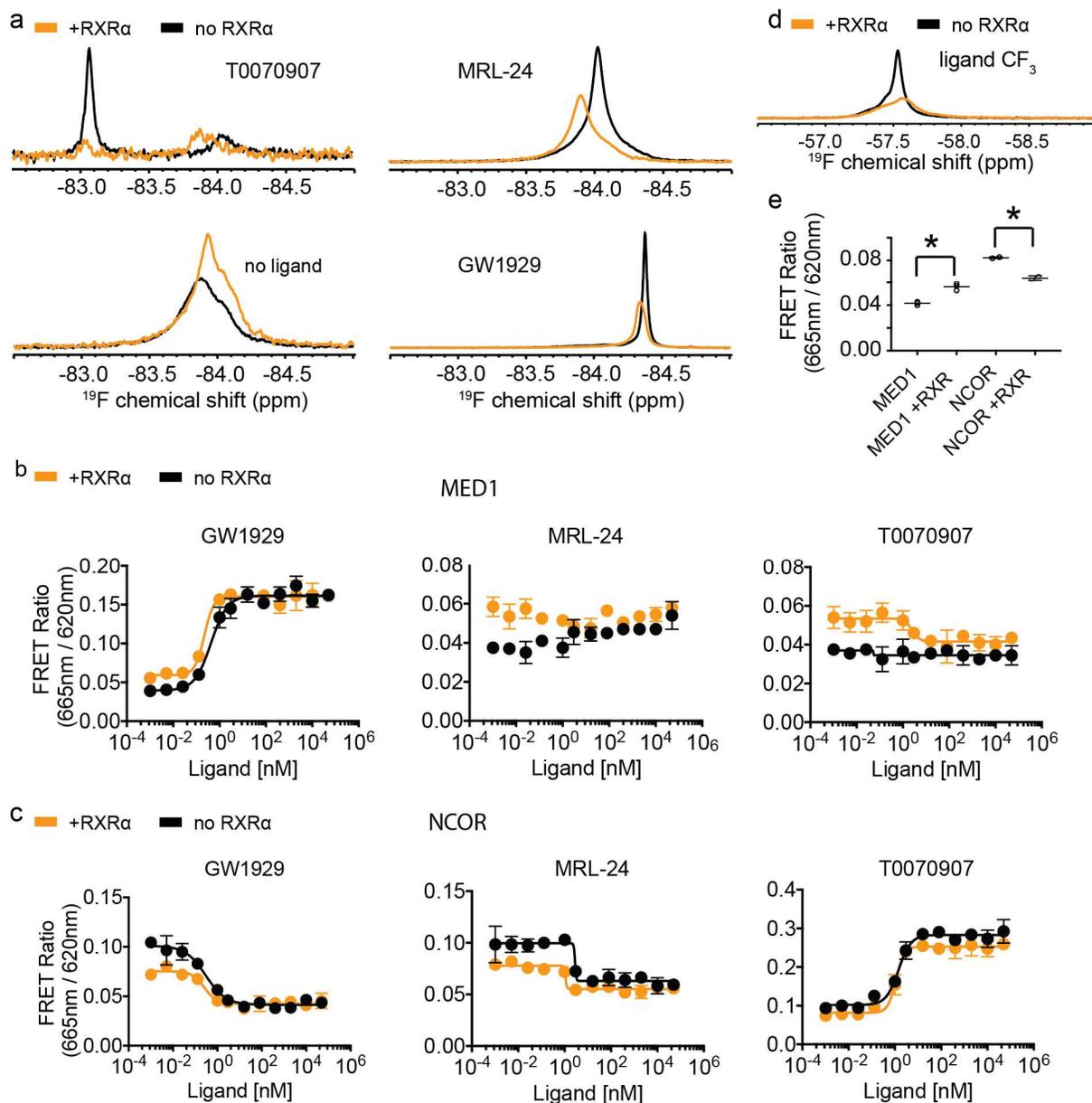
Figure 8. | Heterodimerization of PPARγ LBD with RXRα LBD favors MED1 binding, disfavors NCOR binding. (a) ^{19}F NMR of PPARγ^{K502C}-BTFA bound to T0070907 and PPARγ^{C313A,502C}-BTFA bound to MRL24 or GW1929 or with no ligand bound was performed in the presence (orange) or absence (black) of RXRα LBD. These experiments were performed once. (d) The ^{19}F NMR signal from the MRL24 ligand. Broadening and consequent reduction in signal intensity is expected as a consequence of the increased rotational correlation time of the heterodimer complex. (b-c) TR-FRET was used to measure interaction between wt PPARγ LBD and MED1 or NCOR in the presence or absence of equimolar concentrations of RXRα. Error bars represent standard deviation of two technical replicates within a single experiment. The experiment was repeated twice and gave similar results each time. (e) Heterodimerization favors MED1 binding (p=0.0017) and disfavors NCOR binding (p=0.0076) to apo PPARγ. In addition, visual and statistical comparison of NCOR and MED1 recruitment to PPARγ LBD saturated with

ligand (4 highest concentrations of ligands) indicates that RXRα affects coregulator recruitment to T0070907 bound PPARγ (NCOR p=0.0042; MED1 p=0.0027) more than GW1929 (NCOR p=0.44; MED1 p=0.34) or MRL24 (NCOR p=0.0141; MED1 p=0.061). All p values are derived from a two-tailed t test.

2.3.7 Simulations suggest a diverse ligand dependent ensemble.

The extreme broadening observed in apo PPARγ alone or bound to NCoR peptide originate from multiple conformations exchanging every μs to ms. We reasoned that we could sample some of these conformations in independent long-time scale simulations of apo alone or apo bound to NCoR. We also simulated inverse agonist (T0070907) bound PPARγ LBD with and without co-bound NCoR to gain additional insight into how this ligand changes the coregulator binding surface conformational ensemble. We built simulation models using crystal structures with helix 12 in an "inactive" chain B conformation (**Supplementary Fig. 2d**) to avoid steric clash with bound NCoR peptide. We ran simulations at 37°C (NMR was run at 25°C) and with the TIP3P water model (TIP3P is much less viscous than actual water[226]), to speed relaxation to a local or global energy minimum (i.e. stable conformation) given that this chain B "inactive" conformation may be in a higher energy conformation. These simulations were allowed to run until reaching a structure that remained reasonably stable for at least 5 μs as judged by consistent helix 12 RMSD relative to the starting structure (**Supplementary Fig. 15**). These stable conformations are representative of a local or global free energy minima. Three of four independent simulations of PPARγ LBD alone (apo) relaxed to a cluster of distinct but similar conformations that are somewhat similar to the starting chain B crystal structure, while the fourth does not appear to stabilize to the same degree and goes to a

distinct conformation. These results are consistent with the broad ligand free (apo) PPARγ ^{19}F peaks observed which indicate exchange between two or more conformations (**Figure 9a**) and are also consistent with previous free energy calculations that indicated the "inactive" chain B crystal structure is very similar to low energy apo structures[227]. Addition of NCoR to apo PPARγ LBD results in up to 12 μs of relative instability (**Supplementary Fig. 15b**) before finally relaxing to several different stable conformations that are consistent with the idea that NCoR pushes helix 12 away from the coregulator binding surface, yielding a diverse collection of conformations that are consistent with the observed broad helix 12 ^{19}F NMR peaks observed for this complex (**Figure 9b**). All three simulations of T0070907 bound to PPARγ LBD relax to very distinct conformations (**Figure 9c**). Addition of NCoR to T0070907 bound PPARγ provides a path to one dominant energy minima as all three simulations in this complex reach conformations similar to an "active" conformation but with helix 12 shifted to accommodate the longer LXX I/H IXXX I/L helix of corepressors compared to the LXXLL motif of coactivators[42]. These data are consistent with our ^{19}F NMR data that demonstrate that NCoR binding to the T0070907 PPARγ LBD complex induces a shift towards one particular conformation from several conformations (**Figure 9d**).
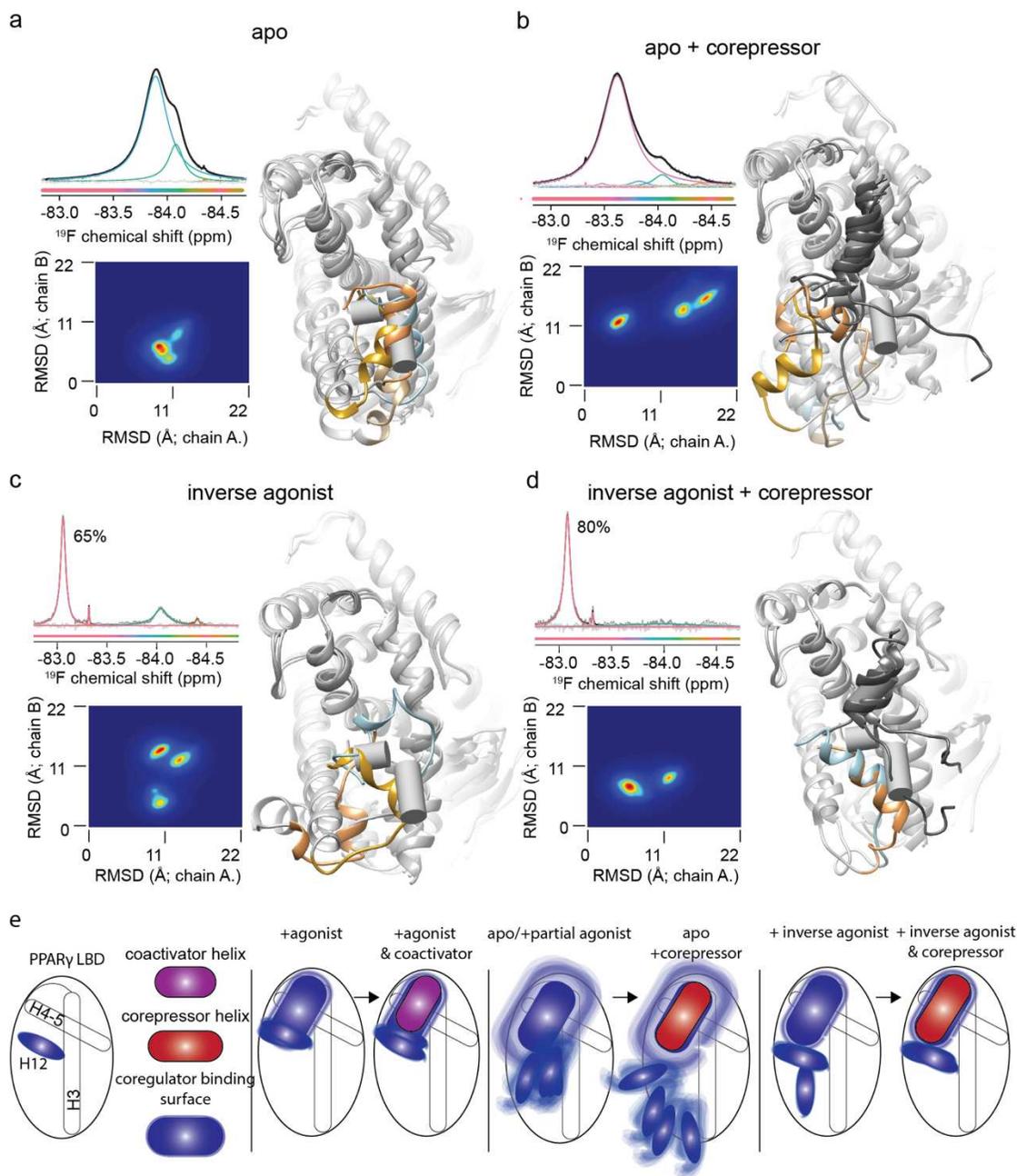
Figure 9 | Microsecond timescale molecular simulations point to a diverse ligand dependent ensemble and a possible NCoR bound structure. (a-d) Root mean square deviation (RMSD) of helix 12 relative to an "inactive" (1PRG chain B) structure and an "active" conformation (1PRG chain A) for helix 12 in four different simulated complexes. Helix 12 from these "inactive" and "active" structures is shown in all panels (grey pipe). SMRT bound to PPARα (PDB code: 1KKQ) was aligned to 1PRG chain A and the SMRT helix is shown in panels b and d (grey pipe). Fourteen independent simulations of (a) apo PPARγ LBD, (b) apo + NCoR corepressor, (c) inverse agonist (T0070907) bound PPARγ LBD, and (d) PPARγ LBD co-bound to T0070907 and NCoR corepressor were carried out for 10 to 30 μs and the last 5 μs were analyzed (4 each for apo simulations and 3 each for T0070907 simulations). Colors on the RMSD plots indicate relative

68

number of frames; red indicates the most prevalent and blue the least. Representative structures from the centroid of the major clusters are shown with helix 12 highlighted in color (4 apo, 4 apo + NCOR, 3 T0070907 and 3 T0070907 + NCoR). Helix 12 disassociates from the LBD in one apo + NCoR simulation. The cluster for this simulation (centered at 21.3Å (x) and 23.8Å (y)) is not shown, however a representative structure is shown in light blue. Corresponding ^{19}F NMR data from **Figure 7** are shown in each panel. (e) A model for helix 12 conformational diversity based on simulation and experiment. Apo or partial agonist bound PPARγ helix 12 is found in many similar conformations of varying helical structure producing broad NMR peaks and rapid hydrogen deuterium exchange, while full agonist is found in a tighter cluster of conformations. Co-binding of the inverse agonist T0070907 and the corepressor NCoR produces one main conformation similar to an "active" conformation, but with helix 12 shifted. Fuzziness implies intermediate exchange (μs to ms) between the conformations.

2.4 Discussion

The data presented here for the first time explicitly reveal the conformational ensemble of the coregulator binding surface, including helix 12, of a nuclear receptor and the effects of ligands on that ensemble. A model of nuclear receptor activation arises from these data and previously published work, which indicates that nuclear receptors and other proteins are found in ensembles of structures and not a single structure[207,211–213,228,229]. The structural variance around this primary structure varies from protein to protein, however for nuclear receptors this variance appears to be considerable[140,207,230]. The data presented here indicate that the coregulator binding surface, including helix 12, exchanges relatively quickly (i.e. μs to ms lifetimes) between many conformations of similar free energy in a broad rough energy well for apo PPARγ, partial agonist bound PPARγ, and apo PPARγ bound to NCoR. Furthermore, agonist and inverse agonist binding reduces the complexity of the apo ensemble and forms distinct, narrow energy well(s) increasing the population found in structurally distinct "active" or "inactive" state(s) which favor coactivator or corepressor binding respectively (**Figure 7 and 9e**). Very slow exchange (seconds) across high kinetic

barriers is observed for some agonists and inverse agonists, indicating that these ligands hold helix 12 in narrow and deep energy wells with rare exchange occurring between sub-ensembles (**Figure 5**). Consistent with these data, free energy calculations indicate that helix 12 in apo-PPARγ is found in two broad energy wells with conformations similar to both "active" and "inactive" apo-PPARγ helix 12 crystal conformations, whereas helix 12 in rosiglitazone-bound PPARγ is found in one deep narrow well with a conformation similar to the "active" chain A conformation[227]. Importantly, these data reveal a correlation between ligand efficacy and the prevalence of at least three distinct structural ensembles using a diverse set of 16 pharmacologically distinct PPARγ ligands.

The long time-scale simulations presented here are qualitatively consistent with the experimental results. Simulations started from an "inactive" chain B structure PPARγ relax to multiple different structures depending on the ligand and coregulator that are bound to PPARγ (**Figure 9**). Apo PPARγ LBD relaxes to conformations similar to the starting "inactive" chain B structure, while addition of NCoR binding pushes helix 12 into various different structures, consistent with the idea that NCoR does not interact productively with helix 12 in apo PPARγ. In contrast, simulations of NCOR and T0070907 co-bound to PPARγ relax to two very similar structural clusters, which are consistent the idea that helix 12 interacts productively with NCoR in this complex. In addition, simulations of GW1929 bound to PPARγ starting with helix 12 in an "active" chain A conformation remain in a similar helix 12 conformation throughout the simulation (**Supplementary Fig. 3**), indicating that as expected this is the dominant

conformation for agonist bound PPARγ in solution. In addition to movement of helix 12 between different positions relative to the rest of the LBD these simulations indicate that the length and helicity of helix 12 can weaken in some conformations (**Figure 9**), which could contribute to broad ^{19}F NMR peaks and increased exchange rates in HDX-MS[161]. Overall it is encouraging that these non-converged simulations qualitatively agree with ^{19}F NMR and provide a glimpse of possible conformations that comprise a portion of the observed ^{19}F NMR spectra, however quantitative comparison between the 19F NMR spectra and converged simulations remains a future challenge.

This view of the coregulator binding surface, including the number and relative populations of conformations and sub-ensembles that comprise the ensemble is made possible by using a single fluorine probe which provides high sensitivity, a single signal which obviates the need to transfer spin between residues, and a total lack of background signal. ^{19}F NMR requires mutation and labeling of PPARγ which perturbs corepressor affinity (especially SMRT) but has less effect on NCoR interaction and little effect on coactivator affinity as measured by TR-FRET peptide recruitment and FP affinities for corepressor (NCOR and SMRT) and coactivator (MED1 and CBP) peptides (**Supplementary Fig. 7**). Based on these observed functional effects, we propose that any effect from the mutations would likely shift inverse agonist bound mean NMR chemical shifts toward the center decreasing the population found in the inverse agonist conformational state (cluster 3). Depending on the chemical shift difference and rate of exchange between the conformations this will result in movement of a

peaks chemical shift and/or a decrease in population of the cluster 3 (left cluster).

For example, a left shifted narrow peak at a chemical shift typical of inverse

agonism (cluster 3) and a broad peak at a chemical shifts typical of

antagonist/partial agonist (cluster 2) are observed for PPARγ^{K502C}-BTFA bound

to the inverse agonist T0070907[156,217] and less efficacious inverse

agonist/antagonist GW9662 (**Figure 2 and 7**). Addition of NCOR or SMRT shifts

the population from cluster 2 to cluster 3, while the mutations and labeling may

shift the equilibrium population in the opposite direction toward the

antagonist/partial agonist cluster (cluster 2). Thus, it may be that helix 12 of WT

PPARγ LBD bound to T0070907 is found in a conformation represented by the

"inverse agonist" cluster (cluster 3; left shifted narrower peak) to a larger degree,

than detected by PPARγ^{K502C}-BTFA.

This work adds detail to how ligands in general control the activity of

nuclear receptors. Our data not only confirm that helix 12 and the coregulator

binding surface exists as a ligand specific dynamic structural ensemble, but also

indicate the relative populations of sub-ensembles that comprise the overall

structural ensemble and correlate function with this ensemble. Further definition

of the conformational ensemble of the entire protein and the kinetics and

thermodynamics of exchange between the members of the ensemble will build an

accurate model of how ligands produce functional outputs via nuclear receptors

and allow greater control of their function via ligands.

2.5 Methods

2.5.1 Protein purification

A pET-46 plasmid carrying the genes for ampicillin resistance and N-terminally 6xHis tagged PPARγ containing a tobacco etch virus nuclear inclusion protease (TEV) recognition site between the His tag and protein of interest was transformed into chemically competent *E. coli* BL21(DE3) Gold cells. Cells were grown in either ZYP-5052 autoinduction media or terrific broth (TB). Cells grown in TB at 37 °C were induced at an OD_{600} of approximately 0.8 by the addition of 0.5mM IPTG and the temperature lowered to 22 °C. Induction proceeded for 16 hours prior to harvesting. Harvested cells were homogenized into 50 mM Phosphate (pH 8.0), 300 mM KCl, 1 mM TCEP and lysed using a C-5 Emulsiflex high-pressure homogenizer (Avestin). Lysates were then clarified and passed through two Histrap FF 5 ml columns in series (GE Healthcare). Protein was eluted using a gradient from 15-500 μM imidazole. Fast protein liquid chromatography (FPLC) was performed on either an NGC Scout system (Biorad) or an ÄKTA Start (GE Healthcare). Eight mg of recombinant 6xHis tagged TEV was added to eluted protein followed by dialysis into 50 mM Tris (pH 8.0), 200 mM NaCl, 1 mM TCEP, and 4 mM EDTA. The protein was again passed through HisTrap FF columns in order to separate cleaved protein from TEV as well as the cleaved 6xHis tag. The cleavage step was only performed on protein which would be used for NMR or fluorescence polarization, protein used for TR-FRET did not have the 6xHis tag removed. The protein was then further purified by gel filtration using a HiLoad 16/600 Superdex 200 PG (GE Healthcare). Size exclusion was performed in 25 mM MOPS (pH 8.0), 300 mM KCl, 1 mM TCEP, and 1 mM EDTA buffer. Protein was then dialyzed into 25 mM MOPS (pH 7.4),

25 mM KCl, and 1 mM EDTA buffer. Protein purity in excess of 95% was determined by gradient 4-20% SDS-PAGE analysis (NuSep). Protein concentration was determined using $\varepsilon_{280} = 12,045$ $M^{-1}cm^{-1}$.

[15]N labeled protein was grown in M9 minimal media containing 99% [15]NH$_4$Cl (Cambridge Isotope Laboratories) as the sole nitrogen source. For this growth cells were grown at 37°C and 180 rpm until an OD$_{600}$ of approximately 1.0 was reached. At this point the temperature was dropped to 22°C for 1 hour. Following cool down period protein expression was induced by the addition of 500 μM IPTG during induction cells remained at 22°C. Protein expression and purification was then accomplished utilizing the same protocol as outlined above.

2.5.2 Delipidation of PPARγ

To delipidate PPARγ LBD, purified protein was diluted to 0.8 mg/ml and batched with Lipidex 1000 (Perkin Elmer) at an equal volume. This mixture was batched for 1 hr at 37 °C and 100 rpm. Immediately following this treatment, protein was pulled through a gravity column by syringe. To increase yield, it was found that speed of elution was important; protein could not remain on the resin at room temperature in excess of 3 min. Two more column volumes of pre-warmed 25 mM MOPS, 25 mM KCl, and 1 mM EDTA were also pulled through in the same manner. Quality of delipidation was then estimated by ^{19}F NMR, loss of lipid can be most easily detected by a reduction in the peak at -84.1 ppm. When non-delipidated protein is used in this report it is labeled as bound to *E. coli* lipids.

2.5.3 Site-directed mutagenesis

Mutations in PPARγ LBD were generated using the Quikchange Lightning site-directed mutagenesis kit (Agilent). To generate each mutation the following primers were used: K502C Forward 5'-aggttaattattagtacaagtcgcagtagatctcctgcaggagcgg-3', Reverse 5'-ccgctcctgcaggagatctactgcgacttgtactaataattaacct-3'; C313A Forward 5'-ccacggagcgaaactgagcgccctgaaagatgcgg-3', Reverse 5'-ccgcatctttcagggcgctcagtttcgctccgtgg-3', Q322C Forward 5'-catactctgtgatctcgcacacagcctccacggagc-3', Reverse 5'-gctccgtggaggctgtgtgcgagatcacagagtatgc-3'. The presence of expected mutations and absence of spurious mutations was confirmed by Sanger sequencing (Eurofins).

2.5.4 Preparation of NMR samples

NMR samples were prepared to a final concentration of 150 µM protein in 470 µL volume containing 10% D_2O. Addition of ligand was done in two separate injections of compound to reduce precipitation. Injections were spaced 30-60 min apart to allow time for binding. All ligands were dissolved in D_6-dimethylsulfoxide (DMSO) with the exception of GQ-16 which was dissolved in D_7-dimethylformamide (DMF). Deuterated solvents were obtained from Cambridge Isotope Laboratories Inc. and were at least 99% isotopically pure. Final concentrations of ligand for samples of PPARγ^{K502C}-BTFA were 1.25x ligand to protein (187.5 µM) with the exception of troglitazone, pioglitazone, and ciglitazone, which were loaded to 2.0x (300 µM) due to poor binding affinity. In samples of PPARγ^{C313A,K502C}-BTFA, ligand concentration was 1.1x to protein

75

(165 μM) with the exception of troglitazone, pioglitazone, and ciglitazone, which were loaded to 1.5x (225 μM). To decrease the likelihood of labeling the single native cysteine (C313) in PPARγ^{K502C}-BTFA, we first loaded ligands into the protein and then labeled with BTFA since bound ligands would restrict access to C313 for all experiments involving PPARγ^{K502C}-BTFA. In both cases, the ligand concentration in DMSO was controlled to maintain a constant volume of DMSO or DMF addition to the sample (8.80 μL for PPARγ^{K502C}-BTFA and 7.76 μL for PPARγ^{C313A,K502C}-BTFA) including for apo and *E. coli* lipid samples. For peptide studies, 1 mM peptide in identical buffer to protein was added at a 2:1 molar ratio (final concentration 300 μM peptide and 150 μM protein). For non-covalent ligands, samples were labeled with 2.0x BTFA after addition of ligand. For covalent ligands, samples were labeled with 10x BTFA following preincubation of T0070907 or GW9662; and addition of 1.5 and 2 ligand molar ratios yielded very similar spectra, indicating complete covalent modification of C313 and likely no bonding to K502C (**Supplementary Fig. 5**). After addition of BTFA, protein was incubated for 30-60 minutes and then buffer exchanged at least 100x using 10kDa Amicon Ultra-15 concentrators (Merck Millipore) to remove excess unbound BTFA. Following this buffered D_2O was added.

Variable D_2O NMR samples were prepared by buffer exchanging protein samples greater than 100x into 25mM MOPS 25mM KCl 1mM EDTA pD 7.4 buffer prepared in either 50% or 100% D_2O using Amicon Ultra-4 10kDa centrifugal filters (EMD Millipore). 100% D_2O buffer was adjusted to read pD 7.4 (pH 7.0[231]), and then mixed with appropriate amounts of H_2O buffer that had been

adjusted to read pH 7.4. Following this, ligands were added to the appropriate

concentration as usual. The samples which contained only 10% D_2O were

prepared as usual with the standard addition of 10% buffered D_2O to the final

sample. Samples of PPARγ^{K502C}-BTFA were already loaded with ligand and

labeled appropriately with BTFA prior to exchange into deuterated buffers.

2.5.5 Fluorescence Polarization (FP) assay

Fluorescence Polarization peptide binding assays were performed by

plating a mixture of 50 nM peptide with an N-terminal fluorescein (FITC) tag, 12-

point serial dilutions of PPARγ-LBD (wild-type, PPARγ^{K502C}-BTFA,

PPARγK502C, or PPARγ^{C313A,K502C}-BTFA), and PPARγ ligands from 50μM to

24nM. PPARγ-LBD and PPARγ ligands were added at a 1:1 ratio. This mixture

was added to wells of low-volume 384-well black plates (Grenier Bio-one,

catalogue number 784076) to a final volume of 16 μL. Peptides were synthesized

by Lifetein LLC (Somerset, NJ) for the for Mediator Complex Subunit 1 (MED1)

peptide, sequence: NTKNHPMLMNLLKDNPAQD; and the Nuclear Receptor

Corepressor 1 (NCoR) peptide, sequence: GHSFADPASNLGLEDIIRKALMG

(2251-2273). Other peptides were purchased from ThermoFisher (Waltham, MA)

for Mediator Complex Subunit 1 (MED1) peptide, sequence:

NTKNHPMLMNLLKDNPAQD (catalogue number PV4549); CREB-Binding

Protein (CBP) peptide, sequence: AASKHKQLSELLRGGSGSS (catalogue

number PV4596); and Silencing Mediator for Retinoid and Thyroid Hormone

Receptors (SMRT), sequence: HASTNMGLEAIIRKALMGKYDQW (catalogue

number PV4424). All dilutions were made in 25 mM MOPS (pH 7.4), 25 mM

KCl, 1 mM EDTA, 0.01% fatty-acid free Bovine Serum Albumen (BSA) (EMD Millipore, catalogue number 126575), 0.01% Tween, and 5 mM TCEP. Assay titrations were performed in duplicate. Plates were incubated in the dark at room temperature for 2 hr before being read on a Synergy H1 microplate reader (BioTek). Fluorescence polarization was measured by excitation at 485nm/20nm and emission at 528nm/20nm for FITC. Data was fit using nonlinear regression (agonist vs response – variable slope 4 parameters) in Prism 7.0b. For FP, TR-FRET and Fluormone competitive binding assays we did two technical replicates and repeated these experiments independently in the lab for once (Fluormone) or 2 or more times FP and TR-FRET. We chose these number of technical replicates and independent experiment replicates based on our experience with the limited variability inherent in these biochemical assays

2.5.6 Time-resolved Förster resonance energy transfer (TR-FRET) assay

TR-FRET peptide recruitment assays were performed by plating a mixture of 8 nM 6xHis-PPARγ-LBD (wild-type, PPARγ^{K502C}-BTFA or PPARγ^{C313A,K502C}-BTFA), 0.9 nM LanthaScreen Elite Tb-anti-His Antibody (LifeTechnologies catalogue number PV5863), 200 nM peptide (N-terminally biotinylated and C-terminally amidated), 400 nM streptavidin-d2 (Cisbio, catalogue number 610SADLB), and 12-point serial dilutions of PPARγ ligands from 50 μM to 1 pM. This mixture was added to wells of low-volume 384-well black plates (Grenier Bio-one, catalogue number 784076) to a final volume of 20 μL. Peptides were synthesized by Lifetein LLC (Somerset, NJ) for Mediator Complex Subunit

1 (MED1) peptide; sequence: VSSMAGNTKNHPMLMNLLKDNPAQ; and

Nuclear Receptor Corepressor 1 (NCoR) peptide, sequence:

GHSFADPASNLGLEDIIRKALMG (2251-2273). All dilutions were made in

25 mM MOPS (pH 7.4), 25 mM KCl, 1 mM EDTA, 0.01% fatty-acid free Bovine

Serum Albumen (BSA) (EMD Millipore, catalogue number 126575), 0.01%

Tween, and 5 mM TCEP. Assay titrations were performed in duplicate. Plates

were incubated in the dark at room temperature for 2 hr before being read on a

Synergy H1 microplate reader (BioTek). TR-FRET was measured by excitation at

330nm/80nm and emission at 620nm/10nm for terbium and 665nm/8nm for d2.

Change in TR-FRET was calculated by 665nm/620nm ratio.

Data was fit using nonlinear regression (agonist vs response – variable

slope 4 parameters) in Prism 7.0b. Outliers were automatically detected using

Prism 7.0b's implementation of the ROUT method[232] and excluded from the

curve fitting (20 out of 2260 total data points were flagged as outliers). However,

all data points, including detected outliers were included in figure graphs. In cases

where curve fitting failed the TR-FRET value nearest the calculated free ligand

concentration in the NMR experiment was used, otherwise the TR-FRET value at

the calculated free NMR ligand concentration was calculated using the fitted

curve. These TR-FRET ratio values were used to correlate with NMR chemical

shift values. One ligand, nTZDpa, showed a biphasic TR-FRET curve. This could

be due to several factors including ligand aggregation[233], absorption interference

in the TR-FRET assay (nTZDpa contains an indole group that could in principle

cause interference), or alternate site binding effects[208]; we therefore utilized the

value of the TR-FRET ratio for nTZDpa based on a fit that did not include the last two concentrations (10 and 50 μM), so as to exclude most effects of the second transition. In addition, the 50 μM SR2088 point was not run because we did not have sufficient ligand. All error bars are standard deviation.

2.5.7 Fluormone competitive binding assays

PPARγ ligand inhibition constants (K_i) were measured using a protocol adapted from LanthaScreen TR-FRET PPARγ Competitive Binding Assay (Invitrogen, catalogue number PV4894). Assay was performed by plating a mixture of 8 nM 6xHis-PPARγ-LBD, 2.5 nM LanthaScreen Elite Tb-anti-His Antibody, 5 nM LanthaScreen Fluormone Pan-PPAR Green (Invitrogen, catalogue number PV4896), and 12-point serial dilutions of PPARγ ligands from 50 μM to 140 fM. This mixture was added to wells of low-volume 384-well black plates (Grenier Bio-one) to a final volume of 16 μL. All dilutions were made in 25 mM MOPS (pH 7.4), 25 mM KCl, 1 mM EDTA, 0.01% fatty-acid free BSA (EMD Millipore), 0.01% Tween, and 5 mM TCEP. Assay titrations were performed in duplicate. Plates were incubated in the dark for 2 hr at room temperature before being read on a Synergy H1 microplate reader (BioTek). TR-FRET was measured by excitation at 330nm/80nm and emission at 495nm/10nm for terbium and 520nm/25nm for Fluormone. Change in TR-FRET was calculated by 520nm/495nm ratio. Nonlinear curve fitting was performed using Prism 7.0b (Graphpad Software, Inc.) as described above for the TR-FRET data, including manual exclusion of highest two concentrations for nTZDpa. 30 of the 1224 total

data points for all three proteins (WT, PPARγ^{K502C}-BTFA and PPARγ^{C313A,K502C}-BTFA) were automatically excluded by Prism in the fits.

2.5.8 K_i calculation

The inhibition constant for each PPARγ ligand was calculated by applying a corrected Cheng-Prusoff (Kenakin, TP 1993 in *Pharmacologic analysis of drug/receptor interaction*, 2nd ed., New York:Raven p. 483.)

$$K_i = \frac{(Lb)\,(IC_{50})(Kd)}{(Lo)(Ro) + Lb(Ro - Lo + Lb - Kd)}$$

where IC_{50} is the concentration of the ligand that produces 50% displacement of the Fluormone tracer, Lo is the concentration of Fluormone in the assay (5 nM), and K_D is the binding constant of Fluormone to wild-type or the two BTFA labeled mutants, Ro is the total receptor concentration and Lb is the concentration of bound Fluormone in the assay with no addition of test ligand. The affinity of Fluormone for the two BTFA labeled mutant proteins was determined via TR-FRET by titration of fluormone into each mutant bound to Elite Tb-anti-His Antibody. Dissociation constants of fluormone for wild-type was measured as PPARγ LBD is 7.9 ±0.2 the variants were measured as 26 ±3 nM for PPARγ^{K502C}-BTFA and 44 ±4 nM for PPARγ^{C313A,K502C}-BTFA and PPARγC313A 12 ±1 nM.

2.5.9 NMR spectroscopy

Acquisition of spectra was performed using a Bruker 700 MHz NMR system equipped with a QCI-F cryoprobe. Chemical shifts were calibrated using an internal separated KF reference in 20 mM KPO$_4$ (pH 7.4) and 50 mM KCl contained in a coaxial tube inserted into the NMR sample tube. KF was set to

be -119.522 ppm, which is the shift of the KF signal with respect to the ^{19}F basic transmitter frequency for the instrument (658.8462650 MHz) at 298.2K, the temperature at which samples were run. Routine 1D fluorine spectra were acquired utilizing the zgfhigqn.2 pulse program (Bruker Topspin 3.5), which consists of a 90° pulse followed by acquisition with proton decoupling (acquisition=0.7 s). Settings were D1=1.2 s, AQ= 0.82 s. Approximately 500 to 4000 transients were collected. Saturation transfer experiments were carried out using the stddiff pulse program. Settings were D1=1.6 s, AQ=0.6 s. For some experiments the total duration of the saturating pulse (Gaus1.1000, 54.52 dB, 50ms) was 1.6 s (D$_2$O) and the location of the saturating pulse was varied. In other experiments, the duration of the saturating pulse was varied, with the saturating pulse location held constant and the rate of exchange was fit using equation 50 found in a previous publication[234]. An off resonance selective saturating pulse was used to determine the peak intensity at time t=0 and R$_1$ was determined experimentally (**Table 2**). Fits of the saturation transfer data were accomplished with a single free parameter using these experimentally determined values for initial intensity and R$_1$ (2.7s^{-1} used for all fits except GW9662 which used 2.4s^{-1}). Transverse and longitudinal relaxation lifetimes (T$_1$ and T$_2$) were determined by fitting data acquired using an inversion recovery experiment (Bruker pulse program t1ir) and cpmg pulse sequence (Bruker cpmg) in Prism 7.0b (Graphpad Software, Inc.) using standard formulas. Spectra were deconvoluted in an objective manner with models chosen statistically by a fitting program[219]. All fits were carried out in the same manner with the same settings in

the fitting program, except where noted. Relative phase of fitted peaks was allowed to vary slightly ($\pi/50$ radians) to accommodate imperfect phasing of these broad signals. The fitting algorithm[219] assumes Lorentzian lineshapes of similar phase. Intermediate exchange effects and field inhomogeneity are likely present in some of these spectra, which will result in inaccuracies in the fitted models; however, notwithstanding these limitations, the deconvolution method provides an objective view of the possible underlying spectral structure and populations. Two-dimensional [^{1}H,^{15}N]-TROSY-HSQC NMR data were obtained using the trosyf3gpphsi19.2 pulse program. Select NMR spectra were replicated in two different ways. 1) Some NMR samples were measured via NMR initially and then days to weeks later to determine if certain parts of the spectrum changed. Cs would indicate that non-reversible processes contribute to that part of the signal, such as unfolding or degradation of the protein. 2) Some spectra were run twice utilizing protein from the same batch as utilized for the first spectra or from an entirely different protein preparation.

2.5.10 Molecular dynamics simulations

Residues in our physical protein constructs used in NMR that were missing in crystal structures (except the N-terminal glycine, which is an unnatural vestige of the cleaved His tag) were added using the modeller[235] extension within Chimera[236] and a PDB file was saved. This PDB file was then submitted to the h++ server[237] (http://biophysics.cs.vt.edu/H++) to determine the state of titratable protons at pH 7.4, along with more realistic rotamers for some residues. This h++ PDB file was then given AMBER names for the various protonation states of

histidine determined by h++ using pdb4amber (AmberTools14[238]). PDB files of

Cysteine-BTFA residue was created through modification of a cysteine residue in

Chimera. This PDB file was then submitted to the RED server[239] for RESP[240]

charge derivation and geometry optimization. RESP values for the cysteine

backbone of these modified residues were constrained to match AMBER cysteine

residue values as part of the input to the RED server. In a manner similar to that

outlined in sections 2-3 in tutorial 5 on the ambermd.org site

([http://ambermd.org/tutorials/basic/tutorial5/)](http://ambermd.org/tutorials/basic/tutorial5/), the output mol2 file was then used

to prepare an ac file and then a prepin file containing the same RESP derived

charges (**Supplementary Methods**) and two force modification files. The

AMBER parameter database derived frcmod file (**Supplementary Methods**) was

loaded after the GAFF[241] parameter database derived frcmod file

(**Supplementary Methods**) within Tleap in order to use AMBER parameters

where possible for the Cysteine-BTFA residue. Tleap was then used to generate

parameter and coordinate files using both ff14SB[242] and GAFF values. A

truncated octahedron solvation cell with boundaries at least 10 Angstroms from

any protein atom was built with TIP3P[243] water. The system was neutralized with

Na+ ions and K+ and Cl- atoms were added to 50 mM. Joung and Cheatham ion

parameters[244] were used. Minimization (imin=1) and equilibration was carried out

in nine steps with non-bonded cutoff (cut) set to 8 angstroms and with the

equilibrations carried out at 310 Kelvin. First, steepest descent minimization

(ntmin = 2) with strong restraints (restraint_WT = 5 kcal/mol Å^2) on protein

heavy atoms for 2000 steps was used followed by NTV MD with shake, the same

restraints and 1 fs steps for 15 ps. Next, two rounds of 2000 steps of steepest

descent minimization with progressively relaxed restraints (restraint_WT = 2 and

0.1 kcal/mol Å^2) followed by a round without restraints. This was followed by

three rounds of NTP MD with shake (5ps, 10ps and 10ps in duration), and protein

heavy atom restraints of 1, 0.5 and 0.5 kcal/mol Å^2. A final unrestrained NTP MD

simulation was then run for 200 ps with 2 fs steps. The final restart file from this

process was used along with a hydrogen mass repartitioned parameter file

(modified using parmed) to run new simulations with new randomized atomic

velocities using 4 fs steps at 310 Kelvin. Analysis was carried out using

CPPTRAJ[245]. All production simulations were carried out using pmemd.cuda or

pmemd.cuda.MPI.

The S enantiomer of GW1929, which was used in NMR and TR-FRET,

was built using Maestro (Schrödinger, LLC). The pyridine ring nitrogen of

GW1929 was the only atom with a predicted pKa near 7.4 (calculated 7.56 +/-

1.12) using the EPIK module (Schrödinger, LLC). We chose to model this

nitrogen as deprotonated. There is no crystal structure for GW1929 bound to

PPARγ, however there is for GI262570, which is bound to PPARγ LBD in the

1FM9 crystal structure. GI262570 is identical to GW1929 for about 2/3 of the

molecule. GW1929 was docked into the 1FM9 crystal structure with GI262570

removed using AutoDock Vina[246]. The best scoring docked binding mode

overlaid well with GI262570. In this docked model in the helix 12 interacting

region the two ligands themselves are identical. RESP charges for GW1929 were

derived using the RED server and force modification files generated using GAFF

parameters. 1FM9 was modified (or not) to incorporate cysteine-BTFA (parameterized as described above) in place of K502 and docked with GW1929 in a similar way to that described above to build PPARγK502C-BTFA and PPARγ bound to GW1929.

1PRG chain B was used to create the build for apo and 3BOR chain B was used to build T0070907 bound PPARγ LBD. 3BOR contains GW9662 which differs from T0070907 by one atom. T0070907 has a nitrogen in place of a carbon atom in one of the ligand rings. This change was made in Chimera and the ligand parameterized and incorporated into the structure as described above for BTFA. NCoR (same sequence as used in NMR and TR-FRET including N-terminal acetylation and C-terminal amidation) was added to these builds utilizing Chimera using the following procedure. The core helix structure from NCoR (from 2OVM) was aligned to SMRT on a PPARα SMRT structure (1KKQ), apo PPARγ chain B (1PRG) was then aligned to PPARα and the PPARα/SMRT structure deleted leaving the aligned NCoR on apo PPARγ chain B. A similar procedure was used with the 3BOR structure to create T0070907 co-bound with NCoR on PPARγ. These PDBs were then used to create the final solvated, minimized and equilibrated structure in a manner similar to that described above.

All simulations were run with settings shown in Supplementary Methods including utilization of SHAKE[247], with variability in the frequency of writing various files to disk. K-means clustering was performed and representative structures from these structures were output by CPPTRAJ[245].

2.5.11 Data availability

Data files for TR-FRET, Fluormone competitive binding assays, and FP are publicly available at https://osf.io/rqdpz/. Any other datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

2.6 Acknowledgements

Author Contribution statement

T.S.H. and D.J.K. conceived the study and designed experiments. I.M.C., M.D.N. prepared samples and performed experiments. T.S.H., I.M.S.dV., J.S., and J.F. prepared samples and performed preliminary experiments. Z.H. and

T.S.H. built, performed and analyzed the molecular simulations. Y.L. assisted with preliminary experiments. A.-L.B., Y.S., and T.M.K. synthesized ligands. H.R.-C. cloned and verified K502C plasmid. P.R.G. contributed to conception, provided key reagents and characterization of ligand pharmacology. R.G.-M. and T.E.C. III assisted T.S.H. with molecular dynamics simulations and ligand parameterization. T.S.H and D.J.K. analyzed data and wrote the manuscript with input from all authors.

Chapter 3: Helix 4 determines the specificity of coactivator interactions in nuclear receptors

Ian M. Chrisman, Zahra Heidari, Trey Patton, Tung-Chung Mou and Travis S. Hughes

3.1 Abstract

Nuclear receptors are a diverse superfamily of transcription factors which are the target of 13% of FDA approved drugs. These proteins function through the recruitment of coregulator complexes which modulate transcription. These coregulator interactions are exceptionally specific and regulate very different pathways. The mechanism by which this selectivity arises is poorly understood. In this work a helix 4 motif on the LBD of nuclear receptors is identified. Through this motif specific interactions are made with coactivators which convey a large degree of specificity. We find that the frequency of interaction between these residues and coactivators is proportional to coactivator affinity in the nuclear receptor peroxisome proliferator-activated receptor γ (PPARγ). The significance of this motif is also verified via molecular dynamics simulations in the nuclear receptor estrogen receptor α (ERα). Additionally, the mechanism by which full agonists display differential efficacy in PPARγ is presented.

3.2 Introduction

The nuclear receptor (NR) superfamily of transcription factors regulates a variety of important pathways in metabolism and development. This superfamily is made up of 48 proteins in the human genome which makes up the largest group of transcription factors[1,2,248]. Dysfunction in these transcription factors is linked to a variety of disease states including cancers[153,249–251], metabolic syndromes[46,49,53,146,252,253] and developmental disorders[254–256]. Thus far 20 of these

proteins when modified at the genetic level have been linked directly to disease states[257–260]. Due to this the NR family has been a focus of drug development with approximately 13% of all FDA approved drugs targeting members of this family[261,262].

The activity of NR is controlled through the recruitment of coregulator proteins. These coregulators are classified into either coactivators which through their interaction promote transcription of regulated genes or corepressors which reduce the expression of regulated genes[263,264]. The interaction of NR and coregulators is highly specific often to the degree of interacting only with a single motif inside a coregulator. These interaction motifs are the well characterized LxxLL motif, NR box (referred to as LxxLL box in this work), seen in coactivators[16,265,266] or the less well understood motif in corepressors which contain either (L/I)xxI(/V)I or Lxxx(I/L)xxx(I/L), referred to as the CoRNR box[17,267,268]. These LxxLL or CoRNR boxes appear to be necessary and sufficient for the interaction and in many cases only a single motif is needed for high affinity binding to the LBD of an NR[269,270].

NR-coregulator interactions can occur in one of two regions of the NR, either the unstructured and very poorly conserved activation factor 1 (AF-1) region of the protein[29,271], or more often through the LBD of the NR via the activation factor 2 (AF-2) region[272,273]. Current knowledge of the mechanism of binding focuses on two major interactions. The first is a hydrophobic cleft present in the NR LBD which allows for interaction with the conserved regions of either the LxxLL or CoRNR box[14,20,274]. The second which is most significant for the interaction of coactivators is the so-called charge clamp. This consists of a conserved lysine at the C-terminal of helix 3 and a conserved glutamate on helix 12[23,275]. These two residues serve to anchor the coactivator through interactions with the peptide backbone of the LxxLL motif in the mousetrap model of coactivator binding[138]. However, neither of these interactions are specific to the coactivator being bound as they interact with conserved motifs.

In this work we identify a highly conserved motif on helix 4 of the NR ligand binding domain (LBD) which makes direct coregulator interactions in the large majority of NR-coregulator crystal structures. This eight-residue motif makes highly selective coregulator interactions through the side chains of the first and last residues in the motif. We propose that this motif leads to a large degree of the observed selectivity in NR coregulator interactions. This hypothesis is verified experimentally in the NR, peroxisome proliferator-activated receptor gamma (PPARγ). From this we observe that the degree of helix 4-coactivator interactions are directly related to the affinity of the complex. Additionally, we identify a structural mechanism which gives rise to differential efficacy among full agonists of this protein. Molecular dynamics driven analysis of NR-coactivator interactions in both PPARγ and estrogen receptor alpha (ERα) indicate that this structural mechanism of coregulator selectivity appears to be conserved among this superfamily of proteins. This provides valuable insight into the currently unknown mechanism by which NR selectively recruit specific coregulators and better defines the functionally important AF-2 region of the NR LBD.

3.3 Results

3.3.1 Coregulator helix 4 interactions in NRs

A query of the protein data bank revealed that of the 48 NR in the human genome 42 had crystal structures of the LBD. Of these, 27 had structures that were solved bound to a coregulator peptide. From these structures we were able to find 23 NR-coregulator complexes that resolved direct interactions between helix 4 residues on the NR and side chains of the coregulator peptide. Almost all of these interactions arose from two residues on helix 4 spaced eight residues apart. Examining the residues in this region we were able to identify a conserved motif. Three of the residues in this motif were almost 100% conserved across all NR. These are an aspartate at position two, a glutamine at position three and a leucine in position six. A consensus logo of this motif is shown in **Figure 1**, other consensus logos can be found in **Figure S1**. Additional interactions were seen in some NR-coactivator structures originating from the residue immediately prior to

this conserved motif (seen with FXR, LXRβ and all ROR proteins), as well as two
with interactions from an arginine located five prior to this motif in helix 3' (in
structures of AR, CAR and GR). However, due to the scarcity of these
interactions it is not realistic to make inferences about their role, but they likely
are significant in selectivity of coactivator recruitment.
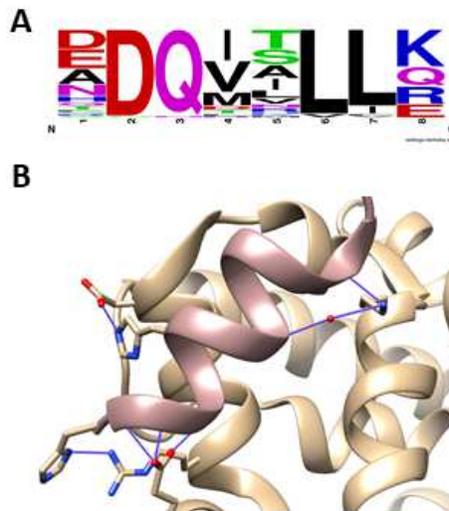
Figure 1. Conservation of the helix 4 motif sequence. A: Consensus logo of the primary
sequence for all nuclear receptors for the helix 4 motif. This omits three proteins (DAX1, GCNF
and TR2) of the 48 human nuclear receptors. This motif is degenerate and cannot be identified in
GCNF and TR2. In DAX1 this motif has expanded to nine residues and cannot be aligned due to
this. B: Example crystal structure of HNFα bound to SRC1 (PDB 1PZL) which shows direct
hydrogen bonding from the side chain of both helix 4 residues, D205 and R212. The consensus
logo was generated using WebLogo[276].

Next, the 15 remaining NR with crystal structures were examined. Among
these this motif could be identified in all but one, Testicular receptor 4 (TR4). In
this case there is a similar sequence but among the three conserved residues in this
motif only the position 6 leucine is present. Further examination of the remaining
NRs based purely on primary sequence was performed. This identified a further
two proteins, testicular receptor 2 (TR2) and germ cell nuclear factor (GCNF),
which do not contain this motif. However, literature on these proteins indicate that

such a lack may not be surprising as they are not known to bind coregulator proteins on the AF-2 surface. For the purposes of examining this motif these three proteins will not be considered due to the lack of consensus. The presence of this motif gives insight into a mechanism through which NR LBDs preferentially bind certain coactivators. Currently this is not well understood as the characterized interactions, the hydrophobic cleft which interacts with the LxxLL box and the charge clamp are conserved among NR.

Examining these sequences, we were able to determine several characteristics of this eight-residue motif, this motif is shown for all 46 NR in which it could be identified in **Figure S2**. The first residue is not heavily conserved but is most frequently an acidic residue, 19 out of 39 cases, and is a residue capable of hydrogen bonding in 79.5% of NR. When this residue is capable of hydrogen bonding it is frequently observed to make interactions with side chains of the coregulator peptide, 14 of 22 proteins crystallized in the presence of a coregulator. The second residue is the fully conserved aspartate. It is not clear what the purpose of this residue is as it is not properly aligned to interact with a coregulator, but it may serve a structural role in the conserved fold seen with NRs. The next residue is the fully conserved glutamine in position three. This residue appears to predominantly make water mediated hydrogen bonds to the carbonyl of one of the leucine residues in the LxxLL motif. It can however hydrogen bond directly to the side chain of a residue in the peptide, but this is less common. The fourth residue is poorly conserved but is a non-aromatic hydrophobic residue (isoleucine, methionine or valine) in 89% percent of the considered cases. Two proteins where this residue is changed are the NR0B family (DAX1 and SHP) which function more like coregulator proteins than as NRs. The other proteins where this residue has changed are the NR4A family (Nur77, Nurr1 and Nor1) which have been shown to lack a normal coregulator binding surface[29,271] and have been found to bind corepressor proteins in a cleft between helix 11 and helix 12, which is the opposite side of H12 from the normal coregulator binding surface[30]. The fifth residue is very poorly conserved, but this is unsurprising as it is oriented directly away from the peptide and does not appear

93

to have any role in peptide recognition or binding. The sixth residue seems to play no role in peptide binding but makes a large amount of hydrophobic contacts which connect helices 4, 8 and 9. This suggests the conservation of this residue has nothing to do with peptide binding and instead is important for proper folding of NRs. The seventh residue is almost exclusively a leucine, in three cases is an isoleucine (AR, MR and PGR) and appears to make non-specific hydrophobic interactions to the leucine repeat in the LxxLL motif of the coactivator peptide. The final residue again seems to make specific hydrogen bonds or salt bridge interactions to the side chains of the peptide. This residue is not very conserved but by majority is a basic residue, 68%, but can also be a glutamine, 18%, or a glutamate, 15%.

Separating out these proteins by family reveals patterns relevant to coregulator specificity profiles within subfamilies. The final residue seems to be heavily conserved within NR families. In NR1 the motif exclusively ends with a lysine while in NR2 all proteins have a motif ending in arginine with the exception of PNR and TLX. However, both PNR and TLX are not well understood proteins so this may indicate that these proteins share their coregulator recruitment pattern more with the NR3 family than with the NR2 family. In the NR3 family the final residue of this motif is less conserved and is just as likely to be a glutamine or a glutamate. This may indicate that in these subfamilies there are differences in preferential recruitment of coactivators. Also interesting is that in the NR1 family the motif sequence is 100% conserved within subfamilies, with the exception of NR1I where the vitamin D receptor (VDR) differs from the consensus sequence of the other two. This is not seen in either NR2 or NR3 and may indicate a greater similarity in coactivator recruitment for proteins from the NR1 family. We further examined the conservation of this motif across several nuclear receptors and species. From this it appears that this motif is 100% conserved in mammals and begins to deviate slightly when more distantly related species are examined. However, this deviation was only found in those nuclear receptors which have a somewhat degenerate motif, TR4 and Nurr1 (**Figure S3**).

3.3.2 Experimental confirmation of coactivator selectivity

There are relatively few NR proteins where the relative affinity of many different coregulator peptides are known. However, there are three, PPARγ[36], TLX[136] and ERα[134], wherein the binding of a wide variety of coregulator peptides have been examined using the PAMGENE system. As well there are some data for relative coregulator affinities for VDR[131] and THRβ[132] but significantly fewer coregulator interactions were examined. Due to a robust pool of synthetic ligands and experience in this system we examined the significance of this motif in PPARγ, which will serve to experimentally verify the importance of these helix 4 interactions and to determine if they are involved in the selective recruitment of coactivator LxxLL boxes to the receptor. There are no published crystal structures of PPARγ which show both helix 4 interactions expected from this motif, (i.e. involving N312 and K319). However, both interactions can be seen independently, the N312 interaction is seen in crystal structures with PPARγ-coactivator 1 alpha (PGC1α) and the K319 interaction can be seen in crystal structures with either steroid receptor coactivator 1 or 2 (SRC1 or SRC2). We hypothesized that co-crystallization of PPARγ with the most efficacious available full agonist and a high affinity coactivator peptide would reveal both these interactions.

3.3.3 Agonist efficacy in PPARγ

There exist many full agonists for PPARγ including two which are currently FDA approved for human use, rosiglitazone and pioglitazone, as well as several experimental ligands of the glitazar or glitazone classes. However, it is not clear which of these is the most efficacious in coactivator recruitment. Based on in cell transcriptional reporter assays it is frequently assumed that these ligands promote the same level of coactivator recruitment and the main differences observed are due to ligand affinity or bioavailability[277,278]. However, this is inconsistent with biophysical data obtained from fluorescence polarization (FP) assays. As such we wished to look at a variety of structurally and functionally diverse ligands, including inverse agonists, partial/non-agonists and full agonists, to determine if there are reproducible differences. This was performed with three

distinct biologically relevant high affinity peptides, PGC1α, CREB-binding protein box 1 (CBP-1) and MED1-2 as well as a corepressor peptide of nuclear receptor corepressor 1 box 2 (NCoR1-2). This will determine if differences in full agonist efficacy could exist and if differences in coregulator peptide affinity are reproducible. These assays were performed by means of FP whereby protein pre-loaded with ligand is titrated into a fixed concentration of fluorescently labeled peptide, the increase in polarization upon binding can be fit to a non-linear regression to determine a K_d of interaction. To simplify interpretation data are represented as a heatmap representing the fold change in affinity with ligand relative to the apo state (**Figure 2a and b**). Actual K_d values are shown in table S1 and S2, curve fits and experimental replicates are shown in **Figures S4 and S5**.
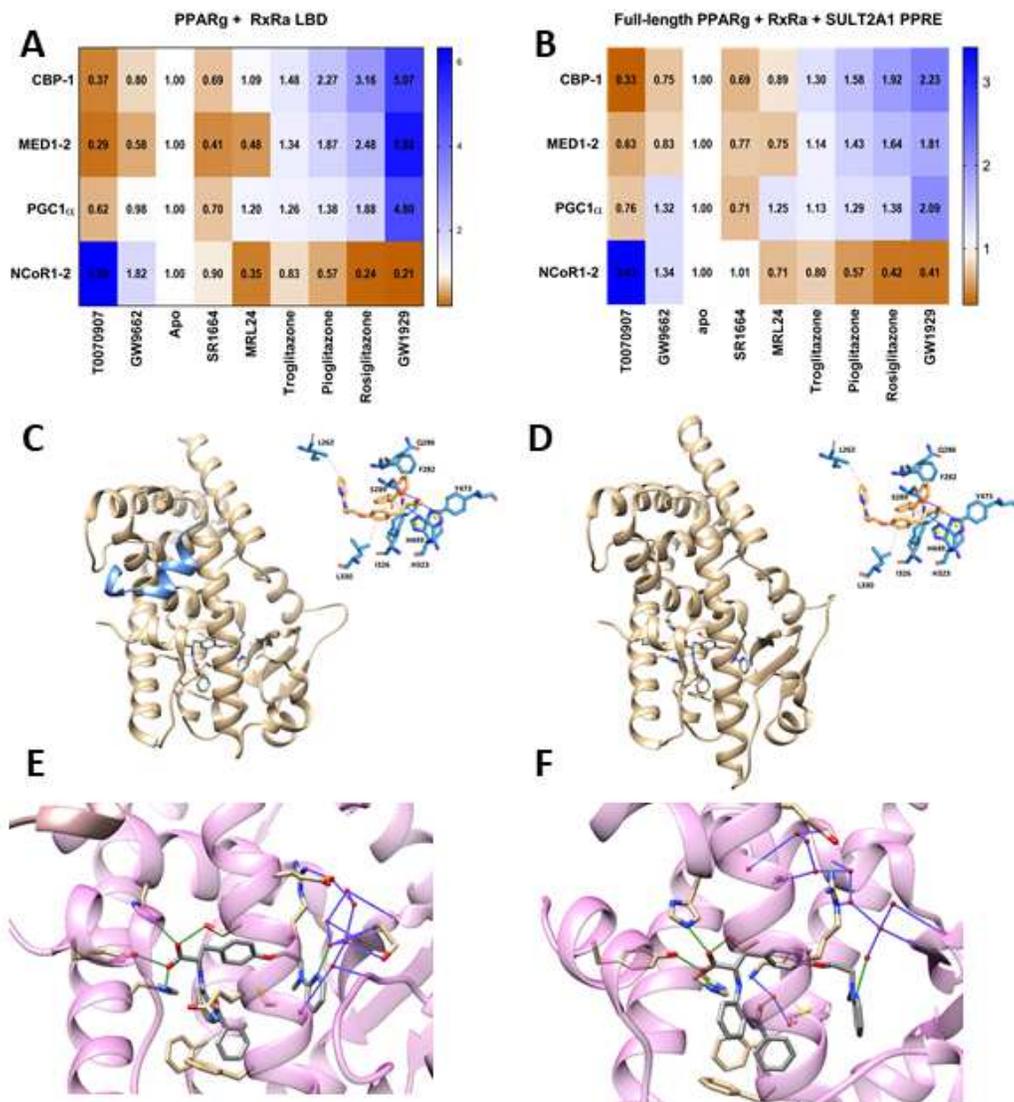
Figure 2. Identification of the most efficacious full agonist of PPARγ. A-B: Fluorescence polarization of either the heterodimer of the PPARγ-RxRα LBD or the full-length complex of PPARγ2-RxRα-SULT2A1 PPRE. Data are shown as fold-change relative to the apo state. All data shown represents two technical replicates from a single experiment. The experiment was replicated independently a second time and those results are shown in figure S4. Curve fits for these data are shown in figure S4. C: The crystal structure of the PPARγ LBD complexed with GW1929 and the MED1-2 peptide. D: Crystal structure of the PPARγ LBD complexed with GW1929. For both structures the direct protein-ligand interactions are shown in the inset. Protein ligand interactions were determined using the Protein Ligand Interaction Profiler[279]. E-F: Blowup of the ligand binding pocket for the crystal structures in C and D showing the large system of highly ordered water molecules which form water-mediated hydrogen bonds between the ligand

97

and the protein. The bonds originating from the ligand are shown in green while the fully water mediated are in purple.

This assay was performed both in a heterodimer of the PPARγ-LBD and the RxRα-LBD as well as the more biologically relevant form of the full-length PPARγ2 and full-length RxRα bound to a PPAR response element from the SULT2A1 gene. In both cases the same pattern was seen wherein the full agonist GW1929 binding led to higher affinity for all studied coactivators than did the other ligands. This demonstrates that full agonists of PPARγ do have different levels of efficacy in coactivator recruitment contrary to the relative affinities implied from in cell transcriptional assays. It is not completely clear why this is, but we hypothesize that in transcriptional assays such as GAL4 the endogenous transcriptional machinery is limiting, preventing differentiation of full agonist behavior. This idea is supported by the phenomenon of squelching wherein coregulator recruitment to reporter plasmids decreases expression from other genes due to limiting supplies of endogenous coregulators such as components of the mediator complex.

3.3.4 Crystallization of PPARγ

We selected mediator of RNA polymerase II transcription subunit 1 LxxLL box 2 (MED1-2) to attempt crystallization of a PPARγ structure where both putative selective helix 4 interactions could be resolved as MED1-2 has not previously been crystallized with PPARγ and has high affinity for PPARγ. This is significant as existing structures of PPARγ include SRC1-1, SRC1-2, SRC2-2 and PGC1α which do not have high affinity for PPARγ. Since GW1929 has not previously been crystallized bound to PPARγ and induces the highest affinity for MED1-2 of any tested ligand we set up trials to co-crystallize the PPARγ-LBD with GW1929 as well as with GW1929 and the MED1-2 peptide, **Figure 2c-f**. Both yielded high-quality crystals which diffracted to 1.9Å, crystallographic data is tabulated in table S3. Initial analysis of the GW1929 bound PPARγ-LBD did not reveal anything significantly different than seen previously in other full

98

agonist bound PPARγ crystal structures. The ligand makes the same four hydrogen bonds to the protein that are seen in many full agonist bound structures. Hydrogen bonds are observed between GW1929 and protein residues S289, H323, H449 and Y473[280,281]. Hydrophobic contacts are also similar to those seen with farglitazar, a closely structurally related ligand, barring only two novel contacts to helices 11 and 12[281]. This indicates the difference in efficacy with GW1929 and other full agonists is likely not due to direct ligand-protein interactions.

What is unique in this crystal is that even in the absence of coregulator peptide it is monomeric which is uncommon for PPARγ structures; only three other WT PPARγ structures are monomeric (3BC5, 3R5N and 4XLD). Additionally, there are no gaps of electron density in the polypeptide chain, including a fully resolved Ω-loop which is also somewhat unusual for a structure of PPARγ. Previous literature suggests that stabilization of the Ω-loop is important for agonism[107], which would be consistent with observations here. The most striking characteristic is the presence of a large system of water-mediated hydrogen bonds which originate from the ligand. These bonds appear to anchor together helices 3, 5 and the β-sheet region. Previously it has been shown that rigidification of PPARγ is related to agonism[140], so this large stabilizing network may contribute to the increase in ligand efficacy seen with GW1929.

Examining the structure bound with MED1-2 almost no structural change is observed with peptide binding, **Figure 3a**. The ligand still makes the same interactions to the protein and the water-mediated hydrogen bond network is still present, although the network of water-mediated hydrogen bonds has changed slightly, contacting the Ω-loop in a stabilizing manner. The Ω-loop undergoes aconformational change upon peptide binding. This may be because the structure without peptide helix 3 contains one additional turn. No crystal contacts were seen in the unit cell to the Ω-loop indicating that the stabilization observed here is an effect of the ligand and not a crystallographic artifact. The one other region of the protein where significant change relative to the peptide free structure is seen is

the phosphorylation loop, residues 238-246, but this region does appear to have significant contacts in the unit cell suggesting these differences may be due to crystal packing forces. As hypothesized both singularly observed helix 4 hydrogen bonds, involving residues N312 and K319, were observed in the GW1929 bound MED1-2 structure simultaneously. This is consistent with the idea that these interactions are important in high affinity binding interactions and are more likely to be observed in crystals representing such structures.

Due to the lack of observed conformational change in the PPARγ-GW1929 complex upon MED1-2 binding, it was of interest to compare it to previously published structures of rosiglitazone with and without coactivator peptide, **Figure 3b**. We attempted to co-crystallize PPARγ with rosiglitazone and MED1-2 for a more useful comparison but were unsuccessful. For this comparison the monomeric crystal structure of PPARγ bound to rosiglitazone, PDB ID:4XLD, and a structure bound to rosiglitazone and a peptide of the reasonably high affinity coactivator PGC1α were selected, PDB ID:3CS8. Pairwise alignment and superposition of the peptide bound and free structures of GW1929 and rosiglitazone reveal that the lack of conformational change observed with GW1929 appears to be unique. In the rosiglitazone bound structures change is observed in helices 3 and 4 with significant changes seen in helix 11. In addition, the Ω-loop is not fully resolved in either structure, however the resolved portion indicates different conformations in this region. We hypothesized that the similarity of the GW1929 and GW1929+MED1 structures and the dissimilarity of the rosiglitazone structures indicates that GW1929 binding restricts the protein to a conformational state that is nearly identical to the coactivator bound state and thereby would reduce any energetic penalties arising from conformational change upon coactivator binding.
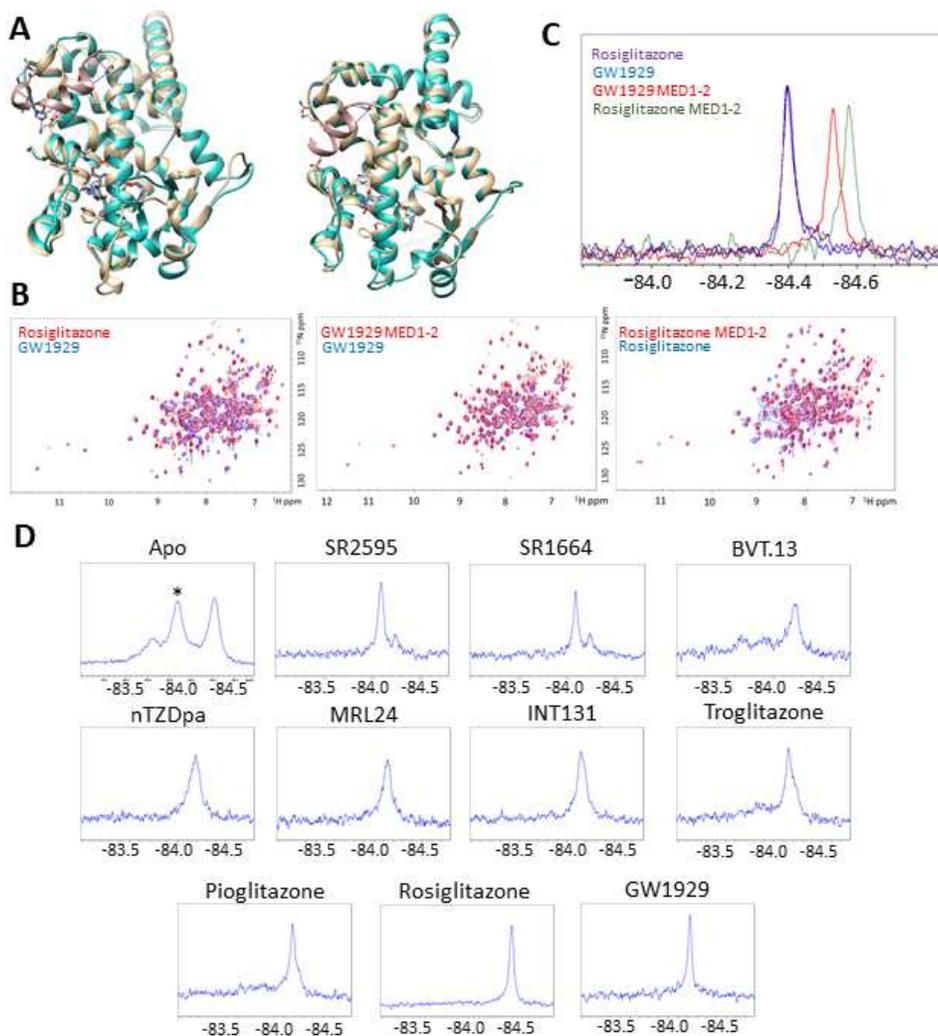
Figure 3. GW1929 stabilizes PPARγ in a state similar to the coactivator bound state. A: Right, crystal overlays of the GW1929 bound PPARγ-LBD in the presence (blue) and absence of MED1-2 (tan). Left, crystal overlays of the rosiglitazone bound PPARγ-LBD in the presence (tan, 4XLD) and absence of PGC1α (blue, 3CS8). B: TROSY-HSQC of the PPARγ-LBD. Left, overlay of the GW1929 and rosiglitazone bound states. Center, overlay of the GW1929-MED1 and the GW1929 bound state. Right, overlay of the rosiglitazone-PGC1α and the rosiglitazone bound states. C: ^{19}F NMR of PPARγ[K474C]-BTFA. D:^{19}F NMR of PPARγ[Q294C]-BTFA with 13 structurally and functionally distinct ligands. An * denotes labeling of the native cysteine, C285. All NMR samples were run once. Labels above indicate the synthetic ligand present in the NMR sample. In panels B and C the text color matches the color of the corresponding spectrum.

3.3.5 ^{15}N HSQC of PPARγ in the presence of coactivators

Because a crystal structure represents only a single low energy state of a protein and we have previously observed PPARγ to be highly plastic we wished to confirm these observations by NMR. Towards this end, we analyzed ^{15}N labeled PPARγ bound to either rosiglitazone or GW1929 in the presence and absence of MED1-2 using TROSY-HSQC NMR, **Figure 3c**. This allows for the observation of significant changes in the backbone conformation of PPARγ upon peptide binding. Comparison of the peptide free rosiglitazone and GW1929 bound states showed only minor differences. A few more resonances are observed in the GW1929 bound state consistent with greater structural rigidity. While the GW1929 bound state spectrum remains virtually unchanged upon MED-1 addition the rosiglitazone bound spectrum changes. Several new resonances are observed upon MED1-2 addition which indicates that the protein is rigidifying. Additionally, almost all peaks have an upfield shift in the ^{1}H dimension. Taken together this indicates that rosiglitazone does not induce or select a conformation as conducive to coactivator binding as GW1929 does. MED1-2 binding to the rosiglitazone PPARγ complex requires significant conformational change/selection which requires energy and lowers affinity.

3.3.6 Probing the PPARγ AF2 conformational ensemble by ^{19}F NMR

Using ^{19}F NMR we focused on specific AF2 regions that contain the charge clamps, which are crucial to coactivator recruitment to determine any structural differences in these regions between full agonists. To accomplish this PPARγ-LBD was irreversibly modified with a trifluoro group (3-bromo-1,1,1-trifluoroacetone; BTFA), via an introduced cysteine as we have done previously[282]. In this work we will utilize the helix 12 probe, PPARγ-BTFA^{K474C}, and a helix 3 probe near the helix 3 charge clamp, PPARγ-BTFA^{Q294C}. The PPARγ-LBD contains only a single native cysteine C285 which points into the ligand binding domain and is largely blocked when ligands are bound to the orthosteric site[282], thus we BTFA label ligand bound PPARγ to avoid labeling the native cysteine. In addition, we confirm all spectra seen in the PPARγ-BTFA^{Q294C} variant with a double mutant which removes the native cysteine, PPARγ-

BTFA^{C285A,Q294C}. Both the single PPARγ-BTFA^{Q294C} and PPARγ-BTFA^{C285A,Q294C} were tested for protein function by FP assays and found to have minimal effects on peptide affinity (Table S4). Both labeled mutants had similar affinity for coregulator peptides as wt PPARγ and appropriate ligand induced changes in affinity for coregulator peptides (Table S4).

The helix 12 probe (PPARγ-BTFA^{K474C}) produces nearly identical spectra when bound to GW1929 or rosiglitazone (**Figure 3d**). Upon addition of MED1-2 both spectra shift upfield, although rosiglitazone shifts more. Thus, the differential affinity for coactivators induced by these ligands may not arise primarily due to helix 12 conformation. This is perhaps not surprising as both ligands make direct hydrogen bonds to Y473 on helix 12 based on crystal structures. Additionally, both ligands induce a similar loss in affinity for NCoR1-2, whose affinity is highly sensitive to the conformation of helix 12[282].

Previously it was shown by hydrogen deuterium exchange mass spectrometry (HDX-MS) that all analyzed PPARγ ligands stabilized helix 3 to some extent, however the C-terminus of helix 3, which contains the charge clamp, was not resolved[164]. More recently it was found that rosiglitazone did not significantly stabilize the charge clamp region, however this was performed in the PPARγ/RxRα heterodimer[124]. We performed fluorine NMR of PPARγ-BTFA^{Q294C}, which contains a probe near the helix 3 charge clamp (K301), bound to a wide variety of 11 apo or ligand-bound states ranging from inverse agonists to full agonists to determine any correlation between spectra and ligand activity (**Figure 3e**). All strong agonists induced a single narrow peak, indicating that they are reducing the conformational flexibility in this region. Less efficacious agonists and non-agonists produced broader or multiple peaks and apo state produced three peaks. Chemical exchange saturation transfer (CEST) NMR and apo WT PPARγ-BTFA spectra confirmed that the middle apo peak arises from labeling of the native cysteine (**Figure S6**). Interestingly, all the highly efficacious agonist ligands (rosiglitazone and pioglitazone) except GW1929 produced the same shift and peak structure. The GW1929 bound state of PPARγ-

BTFA^{Q294C} also had a single narrow peak but a unique shift. These data suggest that some of the difference in efficacy arises from a unique conformation of helix 3 when bound to GW1929. NMR with all ligands was replicated in PPARγ-BTFA^{C285A,Q294C} which produced nearly identical spectra with some differences observed for lower affinity ligands. These differences likely arise from labeling of the native cysteine in the single mutant, which may be more accessible in low affinity ligand PPARγ complexes (**Figure S7**).

3.3.7 Mutagenic analysis of PPARγ

Since the interactions from this motif seem to be related to high affinity interactions and specificity of coactivator recruitment, we decided to test this through mutagenesis. A series of mutations in PPARγ-LBD were made including K301A, N312A, K319A and E499L which are the two charge clamp residues and the two residues on helix 4 which interact with coactivators. We determined the dissociation constant of MED1-2, CBP-1 and PGC1α derived peptides for WT PPARγ-LBD and the point mutants with the full agonists rosiglitazone and GW1929 using FP. These peptides vary in affinity for the PPARγ full agonist complex. Due to anticipated low affinity due to mutation only full agonist ligands were examined. CBP-1 has the highest affinity, MED1-2 has high affinity, while PGC1α has moderate affinity. We attempted to measure affinity for SRC1-2 using FP, however the affinity for WT PPARγ-LBD was too low to allow for comparison with mutants (~37µM affinity for GW1929 bound PPARγ-LBD).
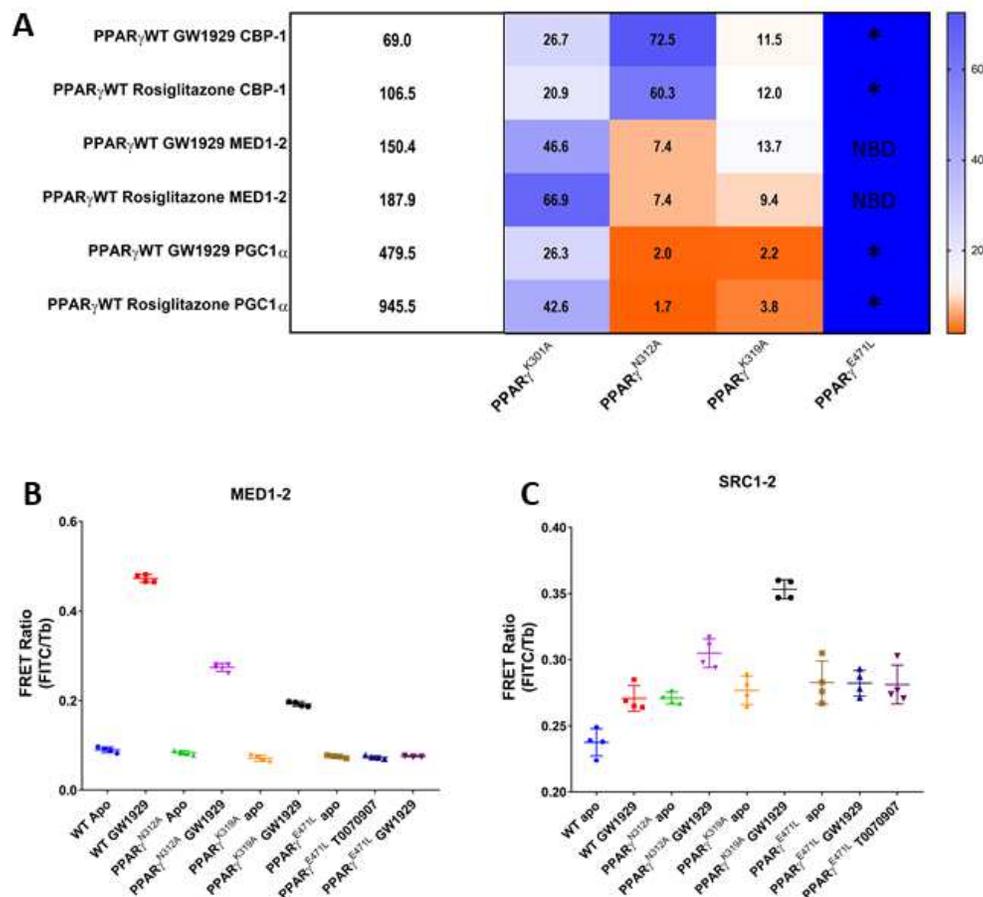
Figure 4. The effect of removing helix 4 residues on coactivator affinity is proportional to the affinity of the peptide. A: Fold change in peptide affinity relative to WT for either the GW1929 or rosiglitazone bound PPARγ LBD point mutants relative to WT. Boxes denoted with an * indicate that the affinity is too low to be accurately determined, >50,000nM, while those marked with NBD, no binding detected, indicates that there was no indication of binding at any concentration used. Data shown is the average of two independent experimental replicates. The PPARγ WT LBD affinity is shown for each complex in the leftmost column in nanomolar. Curve fits and exact K_d values are shown in figure S8 and table S4 respectively. B: TR-FRET of WT and point mutant PPARγ LBD with the MED1-2 peptide. C: TR-FRET of WT and point mutant PPARγ LBD with the SRC1-2 peptide. Data shown in B and C represents four technical replicates run in a single experiment. An experimental replicate is shown in figure S9 for each peptide as well as results with other ligands for the MED1-2 data.

As expected, mutation of the charge clamp significantly effects all peptide binding. Removal of the helix 3 charge clamp residue induces at least a 20-fold

loss in affinity for all coactivators while the removal of the helix 12 charge clamp almost completely ablates coactivator binding (**Figure 4a**). The effect of mutations on helix 4 are more interesting. In the case of CBP-1 we see approximately a 10-fold loss in affinity with the K319 mutation and over a 60-fold loss in affinity with the N312 mutation. This indicates that for the very high affinity peptide both these residues are highly important for binding. With the somewhat lower affinity MED1-2 there is approximately a 10-fold loss in affinity with either helix 4 mutation (**Figure 4a)**. This indicates that both these residues are still significant for this interaction but that the N312 interaction is less important than for CBP-1. When examining the moderate affinity PGC1α either helix 4 mutation has only a 2-fold loss in affinity despite all crystal structures of PPARγ-LBD with this coactivator showing interaction with N312.

Since it was not feasible to examine SRC1-2 binding through FP we utilized time resolved (TR)-FRET based binding assays to examine mutational effects on the binding of this low affinity peptide. In this assay the FRET ratio, which reflects the amount of SRC1-2 bound to PPARγ, was used as a proxy for relative affinity. This assay was run at various SRC1-2 concentrations to ensure reproducibility. Ligand induced changes in TR-FRET were measured for MED1-2 to determine if the results correlate with dissociation constants obtained using FP (**Figure 4b and c**). TR-FRET data for MED1-2 with additional ligands and experimental replicates are shown in **Figure S9**.

The MED1-2 results were consistent with the FP results for the E499L mutation, which completely abolished binding for all complexes. Both helix 4 mutations reduced binding significantly. The PPARγ^{K319A} mutant had more of an effect on FRET than PPARγ^{N312A} (**Figure 4b**), consistent with FP (**Figure 4a**). The E499L mutation completely abolished binding of SRC1-2 similar to the other coactivators tested using FP. In contrast, the helix 4 mutations had smaller or even opposite effects on SRC1-2 binding when compared to the higher affinity coactivators. The PPARγ^{N312A} mutant did not affect SRC1-2 binding and the PPARγ^{K319A} mutant had increased binding of SRC1-2 compared to WT. This

supports our hypothesis that these helix 4 interactions are important for high affinity binding between nuclear receptors and coactivators and consequently for coactivator specificity, that is they determine the coactivator binding profile for a given nuclear receptor.

3.3.8 Molecular dynamics indicates this mechanism in shared in nuclear receptors

After confirming the role the helix 4 residues plays in the selective recruitment of coactivators to PPARγ, we questioned whether this mechanism is shared across nuclear receptors. Towards this end molecular dynamics simulations were employed in another NR with known relative coregulator binding affinities, ERα[134]. Simulations were performed from crystal structures deposited to the PDB, 3UUD and 5DX3, which were bound with low affinity coactivator peptides, SRC1-2 and SRC2-2 respectively. Unfortunately, no crystal structures existed which were bound to a high affinity peptide so the relatively high affinity peptide SRC1-3 was docked *in silico* using existing SRC1 containing structures. These simulations were performed in both the point mutant utilized in crystallography, Y537S, as well as a WT protein where the mutation was reversed. The Y537S mutant and WT ERα showed significantly different behavior. It is not immediately clear why this is, but this mutation is proximal to helix 12 and likely exerts a significant effect on the helix 12 conformational ensemble. Thus we focus here on the results using WT ERα. Protein stability and proper binding modes of the peptide were confirmed for through whole protein RMSD and analysis of hydrogen bonding between the charge clamp residues and the peptide, **Figure S10**.
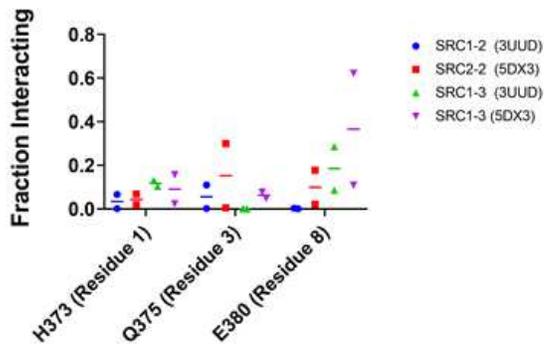
Figure 5. Molecular dynamics indicate the significance of this motif in ERα. A: Significantly greater interaction is seen between the terminal residues of the helix 4 motif and a high affinity coactivator, SRC1-3, peptide with ERα than is seen with two lower affinity coactivator peptides, SRC1-2 and SRC2-2. Each point represents the fraction of hydrogen bonding or salt bridge interactions between the indicated residue and coactivator peptide from a single monomer within the homodimer of ERα. For each simulation the two data points along with the mean are shown.

These simulations showed higher prevalence of hydrogen bonding between residues 1 and 8 of the helix 4 motif and the peptide for the high affinity peptide (SRC1-3) than for the low affinity peptides (SRC1-2 and SRC2-2; **Figure 5**). The non-selective central residue in the helix 4 motif, which makes primarily non-specific backbone interactions, had similar prevalence of protein-peptide interactions for all complexes. These results support the idea that these helix 4 residues control coactivator specificity for the 40 nuclear receptors which contain this conserved motif.

3.4 Conclusions

The data presented herein identifies a mechanism for the selective nature of NR-coregulator interactions, wherein a specific LxxLL or CoRNR box of a coregulator makes discriminating interactions with conserved helix 4 residues, controlling relative affinity. The charge clamp residues exert the greatest effect on binding energy, but do not discriminate between coactivators, as they bond to the backbone of the peptide. The hydrophobic interactions of the leucines, isoleucines

and other hydrophobic residues in the LxxLL motifs also impart affinity but do not discriminate between the known coactivators, many of which have been identified by their possession of this hydrophobic motif[15,283]. We found that the degree to which helix 4 residues interact with an LxxLL box is proportional to the affinity of the coactivator complex for both ERα and PPARγ.

We determined that different full agonists induce distinct affinities for coactivators, in contrast to cell-based transactivation assays, which may suffer from transcriptional machinery squelching. The most efficacious ligand identified was the tyrosine derivative GW1929[142]. In addition, we found that the difference between GW1929 and another highly efficacious full agonist, rosiglitazone, is independent of helix 12. Rather, it appears that this ligand better stabilizes the whole of the LBD of PPARγ in a conformation which is almost identical to that seen when bound with coactivator. This removes the requirement of major rearrangement of the protein to allow for coactivator binding thereby increasing the net binding energy. This effect appears to arise primarily through a large system of ordered water molecules which allow the ligand to stabilize a much larger area of the protein than would be expected through water-mediated hydrogen bonds.

This work reveals the physical mechanism by which NR preferentially bind certain coregulator proteins and better defines the AF-2 region of NR proteins. These data indicates that the AF-2 is made up of two distinct regions which serve different purposes in coactivator binding events. The first comprised of the C-terminal of helix 3 and helix 12 provides the majority of the binding energy but none of the selectivity. While the second region, helix 3' and helix 4, makes highly specific interactions which determine coactivator specificity. This could lead to design of drugs which alter the conformation of the second region specifically and drive the preferential recruitment of desired coactivators to lead to a more specific functional outcome.

3.5 Methods

3.5.1 Materials Used

Isopropy-beta-D-thiogalactopyranoside (IPTG) was obtained from Goldbio (I2481C50). Tris-(2-carboxyethyl)phosphine hydrochloride (TCEP) was obtained from Biosynth (C-1818). Trizma base (T1503), MOPS (RDD003), Potassium Chloride (746436), Ethylenediaminetetraacetic acid (EDTA, 798681), Monobasic potassium phosphate (795488), Dibasic potassium phosphate (795496), 3-bromo 1,1,1-trifluoroacetone (BTFA, 374059), and Tween 20 (P1379) were obtained from Sigma Aldrich. Sodium Chloride (0241) was obtained from Amresco. Fatty acid free bovine serum albumin (126575) was obtained from EMD Millipore. Pierce protease inhibitors (88265) were obtained from Fisher Scientific.

3.5.2 Protein Purification

A pET-45b plasmid carrying the genes for ampicillin resistance and N-terminally 6xHis tagged PPARγ LBD or containing a tobacco etch virus nuclear inclusion protease (TEV) recognition site between the His tag and protein of interest was transformed into chemically competent *E. coli* BL21(DE3) Gold cells. Cells were grown in terrific broth (TB). Cells grown in TB were induced at an OD_{600} of approximately 0.8 by the addition of 0.5mM IPTG. Induction proceeded for 16 hours prior to harvesting. Harvested cells were homogenized into 50mM Phosphate, 300mM KCl, 1mM TCEP pH 8.0 and lysed using a C-5 Emulsiflex high-pressure homogenizer (Avestin). Lysates were then clarified and passed through two Histrap FF 5ml columns in series (GE Healthcare). Protein was eluted using a gradient from 15-500μM imidazole. Fast protein liquid chromatography (FPLC) was performed on either an NGC Scout system (Biorad) or an ÄKTA Start (GE Healthcare). To eluted protein 2-8mg of recombinant 6xHis tagged TEV was added, approximately 1:40 w/w TEV to PPARγ, followed by dialysis into 50mM Tris, 200mM NaCl, 1mM TCEP and 1mM EDTA pH 8.0. To separate cleaved protein from TEV as well as cleaved 6xHis tag the protein was again passed through HisTrap FF columns. The cleavage step was only performed on protein which would be used for NMR or FP, protein used for TR-FRET did not have the 6xHis tag removed. The protein was then further purified

by gel filtration using a HiLoad 16/600 Superdex 200 PG or HiLoad 16/600 Superdex 75 PG (GE Healthcare). Size exclusion was performed in 25mM MOPS, 300mM KCl, 1mM TCEP, 1mM EDTA buffer pH 8.0. Protein was then dialyzed into 25mM MOPS, 25mM KCl, 1mM EDTA buffer pH 7.4. RxRα-LBD was purified by the same protocol as PPARγ-LBD with the only difference being that in all buffers for RxRα purification the TCEP concentration was increased to 5mM to reduce tetramerization. FL PPARγ2 or RxRα was purified in the same manner with the caveat that the 6xHis tag was not cleaved for this protein as the cleavage negatively affected protein activity. Protein purity in excess of 95% was determined by gradient 4-20% SDS-PAGE analysis (NuSep) for the PPARγ-LBD or the RxRα-LBD. In the case of the FL PPARγ2 or RxRα the presence of degradation products made assessing purity more difficult.

3.5.3 Expression of ^{15}N-labeled PPARγ LBD

To express ^{15}N-labeled PPARγ LBD *E. coli* BL21(DE3) gold cells carrying the same pET-45b vector as above were grown in modified M9 media containing ^{15}NH$_4$Cl as the sole nitrogen source were grown to an OD of approximately 1.0 at 37°C and 180 rpm. M9 media was made following the protocol of Justine M. Hill and amended to include a vitamin extraction. This vitamin extraction was made by dissolving a centrum for men vitamin tablet in 20ml of 18MΩ water as much as possible. This solution was then centrifuged at 4,000g for 20 minutes and the supernatant was removed. 1.0ml of this supernatant was added to each liter of M9 medium. Following this the temperature was dropped to 22°C for one hour prior to induction. Cells were induced with 1mM IPTG and induction proceeded for 16 hours. PPARγ LBD was then purified as described above.

3.5.4 Delipidation of PPARγ

To delipidate PPARγ LBD purified protein was diluted to 0.8mg/ml and batched with Lipidex 1000 (Perkin Elmer) at an equal volume. This mixture was batched for 1 hour at 37°C and 100 rpm in an incubator. Immediately following this treatment, protein was pulled through a gravity column by syringe. To

increase yield, it was found that speed of elution was important, protein could not remain on the resin at room temperature in excess of 3 minutes. Two more column volumes of pre-warmed 25mM MOPS, 25mM KCl and 1mM EDTA were also pulled through in the same manner.

3.5.5 Site-Directed Mutagenesis

Mutations in PPARγ LBD were generated using the Quikchange Lightning site-directed mutagenesis kit (Agilent). To generate each mutation the following primers were used: K474C Forward 5'-aggttaattattagtacaagtcgcagtagatctcctgcaggagcgg-3', Reverse 5'-ccgctcctgcaggagatctactgcgacttgtactaataattaacct-3', C285A Forward 5'-ccacggagcgaaactgagcgccctgaaagatgcgg-3', Reverse 5'-ccgcatctttcagggcgctcagtttcgctccgtgg-3'., Q294C Forward 5'-gcatactctgtgatctcgcacacagcctccacggagc -3', Reverse 5'-gctccgtggaggctgtgtgcgagatcacagagtatgc -3'The presence of expected mutations and absence of spurious mutations was confirmed by Sanger sequencing (Eurofins).

3.5.6 Preparation of NMR Samples

NMR samples were prepared to a final concentration of 150μM protein in 470μL volume containing 10% D_2O. Addition of drug was done in two separate injections of compound to reduce precipitation. Injections were spaced 20-30 minutes apart to allow time for binding. All drugs were dissolved in D_6-dimethylsulfoxide (DMSO). Deuterated solvents were obtained from Cambridge Isotope Laboratories Inc. and were at least 99% isotopically pure. Final concentrations of drug for samples of PPARγ-BTFA were 1.25x drug to protein, 187.5μM, with the exception of Troglitazone, Pioglitazone, and BVT.13 which were loaded to 2.0x, 300μM, due to poor binding affinity. In both cases the concentration of the drug was controlled to maintain a constant volume of DMSO, 8.80μL for PPARγ-BTFA. Peptide addition to NMR samples was done to 2.0x peptide to protein, 300μM, of peptide dissolved to 1-1.5mM in buffer.

Following drug addition to PPARγ-BTFA, samples were labeled with 1.0x molar equivalents BTFA for non-covalent drugs, the covalent drugs, T0070907 and GW9662, were labeled with 10x BTFA. After labeling the protein was incubated overnight and then buffer exchanged approximately 100x using 10kDa Amicon Ultra-15 concentrators (Merck Millipore) to remove excess unbound BTFA. Following this buffered D_2O was added.

3.5.7 Fluorine-19 (^{19}F) NMR spectroscopy

Acquisition of spectra was done on a 700 Mhz Bruker magnet with a QCI-F cryoprobe. Various pulse programs were used. Chemical shift was calibrated using an internal separated KF reference (in 20 mM KPO4 pH 7.4 50 mM KCl) contained in a coaxial tube inserted into the NMR sample tube. KF was set to be -119.522 ppm, which is the shift of the KF signal with respect to the fluorine basic transmitter frequency for the instrument (658.8462650 MHz) at 298.2K, the temperature at which samples were run. Routine 1D fluorine spectra were acquired utilizing the zgfhigqn.2 pulse program (topspin 3.5), which consists of a 90-degree pulse followed by acquisition with proton decoupling (acquisition=0.7 s). Settings were D1=1.2 s, AQ= 0.82 s. Approximately 500 to 4000 transients were collected. Saturation transfer experiments were carried out using the stddiff pulse program. Settings were D1=1.6 s, AQ=0.6 s. For some experiments the total duration of the saturating pulse (Gaus1.1000, 54.52 dB, 50ms) was 1.6 s (D20) and the location of the saturating pulse was varied. In other experiments D20 was varied, the saturating pulse was constant and the rate of exchange was fit using equation 50 found here[234]. An off resonance selective saturating pulse was used to determine the peak intensity at time t=0 and R_1 was determined experimentally. Fits of the saturation transfer data were accomplished with a single free parameter using these experimentally determined values for initial intensity and R_1. 2D TROSY amide proton nitrogen correlation data was obtained using the trosyf3gpphsi19.2 pulse program. Transverse and longitudinal relaxation lifetimes (T_1 and T_2) were determined by fitting data acquired using the cpmg and t1ir pulse programs in Prism 7.0b (Graphpad Software, Inc.) using standard formulas.

We deconvoluted the spectra in an objective manner with models chosen statistically by a fitting program. All fits were carried out in the same manner with the same settings in the fitting program, except where noted. Relative phase of fitted peaks was allowed to vary slightly ($\pi/50$ radians) to accommodate imperfect phasing of these broad signals. The fitting algorithm[219] assumes Lorentzian lineshapes of similar phase. Intermediate exchange effects and field inhomogeneity are likely present in some of these spectra, which will result in inaccuracies in the fitted models; however, notwithstanding these limitations, the deconvolution method provides an objective view of the possible underlying spectral structure and populations.

3.5.8 Crystallization of PPARγ-LBD

Crystals of PPARγ-LBD bound to GW1929 were grown by sitting drop with a 1:1 mix of protein to 200mM tri-lithium citrate, 20% w/v PEG 3350 (pH not fixed). Protein concentration was 10mg/ml (318μM) premixed with a 1.05x molar excess of GW1929 in D_6-DMSO (334μM). Large rock/needle crystals grew from the mother liquor at 4°C in 3 days. Crystals of PPARγ-LBD complexed with both MED1-2 and GW1929 were grown by sitting drop with a 1:1 mix of protein to 100mM HEPES, 25% w/v PEG 2000 MME, pH 7.5. Protein concentration was 10mg/ml premixed with a 1.05x molar excess of both MED1-2 and GW1929. Large plate crystals grew from the mother liquor at 4°C in three days. The MED1-2 peptide used in crystallography was comprised of residues 631-655, primary sequence is Ac-VSSMAGNTKNHPM**LMNLL**KDNPAQ-NH$_2$, the binding LxxLL motif is in bold. This peptide was synthesized by Lifetein LLC.

3.5.9 Crystallographic analysis

The PPARγ crystal co-crystallized GW1929 or GW1929 with MED1-2 peptide crystals were cryoprotected by transferred into reservoir solution containing 15-20% (v/v) glycerol and flash-frozen in a 100K nitrogen gas stream.

Diffraction data were collected at 100K and a wavelength of 0.979 Å at the Advanced Photon System SBC-CAT 19-BM beamlines equipped with an

ADSC Q210r CCD detector. Each crystal was collected with an oscillation range (0.2-0.5° per image) for a total 180-360° per dataset Images were indexed, integrated, and scaled using HKL2000[284].The initial phasing map were determined by molecular replacement method with Phaser program, integrated in PHENIX software suite[285] using coordinates of the previous published PPARγ structure (PDB code: 5TTO) as a searching model, against 10-2.5Å experimental data. The initial models were placed into a likelihood-weighted $2mF_o$-DF_c map with COOT[286] and first subjected to one cycle of rigid-body refinement using PHENIX. Subsequently, the models were further refined by iterative model rebuilding with COOT and refinement of atomic positions, real space, occupancy, and isotropic B-factor parameters with PHENIX using a small set of reflections (<10%) for calculation of R_{free}. Ligand in each structure was located in a SIGMA-A weighted mF_o-DF_c difference omit map[287] computed with phases from the refined model. The models were further refined multiple rounds until R_{free} converged. For Gw1929/MED1-2-bound PPARγ structure, additional TLS refinement was included in the final round of refinement. The final refined coordinates and structure factors for GW1929 and Gw1929/MED1-2-bound PPARγ protein have been deposited in Protein Data Bank (PDB) under ID code of 6D8X and 6D94, respectively. Data collection and refinement statistics of a final model are shown in Table 1.

3.5.10 Fluorescence polarization

Fluorescence polarization was performed in 25mM MOPS, 25mM KCl, 1mM EDTA, 5mM TCEP, 0.1mg/ml fraction V BSA, 0.01% Tween 20, pH 7.4. In this assay into 50nM of N-terminally FITC labeled peptide PPARγ loaded stoichiometrically with ligand was titrated over 12 points covering a range of 24nM-50μM in a 384-well plate. Solutions were incubated for two hours in the dark at room temperature and then the fluorescence polarization at 528nm was measured using a Synergy H1 plate reader (Biotek). Fluorescence polarization data was plotted in GraphPad Prism 7.0 (GraphPad software) and the dissociation

constant, K_d, was determined by a non-linear regression. All assay conditions were run in duplicate.

Peptides used in fluorescence polarization assays are as follows, CBP-1, residues 62-81, primary sequence is AASKHKQLSELLRGGSGSS, MED1-2, residues 638-656, primary sequence is NTKNHPMLMNLLKPAQD, PGC1α, residues 136-155, primary sequence is EAEEPSLLKKLLLAPANTQ, NCoR1-2, residues 2251-2273, primary sequence is GHSFADPASNLGLEDIIRKALMG. FITC conjugated peptides of CBP-1 and PGC1α were purchased from Thermo Fisher, FITC conjugated peptides of MED1-2 and NCoR1-2 were synthesized by Lifetein LLC.

3.5.11 Time-Resolved Förster Resonance Energy Transfer

TR-FRET was performed in the same buffer and plate conditions as fluorescence polarization. In this assay 8nM of 6xHis tagged PPARγ-LBD and 0.9nM of LanthaScreen Elite Tb Anti His antibody (LifeTechnologies catalogue number PV5863) was added to either 200nM of FITC-MED1-2 or 1,000nM of FITC-SRC1-2 in the presence or absence of 1μM ligand. Following addition of all components plates were incubated in the dark at room temperature for two hours. TR-FRET was measured with excitation at 330/80nm and emission at 495/10nm for terbium cryptate and emission of 520/20nm for FITC. Change in TR-FRET was measured by the ratio of the 520/495 emission.

Chapter 4: Inverse agonists induce drug specific structure and function via distinct molecular mechanisms in the nuclear receptor PPARγ.

Zahra Heidari, Ian M. Chrisman, Scott Novick, Michelle Nemetchek, Theodore Kamenecka, Patrick Griffin, and Travis Hughes

4.1 Abstract

The off state of PPARγ in particular and nuclear receptors in general is not well defined. While there are several published structures of PPARγ in an active/on state (i.e. bound to a coactivator) to date there is no published structure of PPARγ in an inactive/off state (i.e. bound to a corepressor). Furthermore, most antagonist and agonist bound PPARγ structures are very similar to the apo structure. In this study, we use NMR, hydrogen deuterium exchange mass spectrometry, and extensive accelerated molecular dynamics simulations of PPARγ alone or bound to several inverse agonists and antagonists as well as these complexes cobound to a nuclear receptor corepressor 1 (NCoR1) peptide to demonstrate that PPARγ can adopt distinct inverse agonist structural states. We demonstrate that the helix 3 conformational ensemble is ligand and inverse agonist dependent, including the helix 3 charge clamp. One inverse agonist (T0070907) appears to select/induce just two primary helix 3 and helix 12 structures, while the other (SR10221) selects/induces a more conformationally diverse ensemble. One mechanism of inverse agonism involves disruption of a tripartite salt-bridge which destabilizes the active conformation via disruption of interactions between the helix 2-3 loop and helix 3 which aids formation of non-active helix 12 states with high affinity for corepressors. Finally, we show that these two distinct inverse agonist states induce differential corepressor peptide affinity.

4.2 Introduction

Peroxisome Proliferator-Activated Receptor γ (PPARγ) is a member of the nuclear receptor (NR) superfamily of ligand-dependent transcription factors which

control transcription of a myriad of genes which are implicated in many physiological processes including lipid homeostasis[288,289]. Drugs affect the transcriptional activity of nuclear receptors via recruitment of other proteins, termed co-regulators, which affect transcription through histone modification, post-translational modification of the receptor, or through bridging to promoter associated transcriptional machinery[290,291]. The vast majority of nuclear receptor drugs bind deep in a ligand binding pocket within the ligand binding domain (LBD) allosterically changing the nuclear receptor surface. These changes include changes to the surface that interacts with most coregulators, the coregulator binding surface or activation function-2 (AF-2) surface which includes portions of helices 3, 4 and 12. In this way drugs can modify transcription of genes regulated by a given nuclear receptor[282,292]. Surprisingly, the AF-2 surface looks very similar in many nuclear receptor crystal structures bound to antagonists and agonists. Some structures do show very distinct coregulator binding surfaces, although the influence of crystal contacts makes interpretation difficult[293]. Almost 200 crystal structures of PPARγ bound to various drugs and three different coactivators have been published indicating that full agonists stabilize helix 12 in a particular on/active conformation[294]. However, there is no structure of PPARγ bound to any corepressor or corepressor peptide (i.e. PPARγ in the transcriptionally off state) and relatively few studies have investigated the structural mechanism of PPARγ mediated transcriptional repression[161,295,296]. In fact, the idea that nuclear receptor inverse agonists are more than antagonists that displace activating endogenous ligands is not widely accepted. However, our recently published biophysical work indicates that bona fide inverse agonists of PPARγ exist[282,296]. Despite the lack of PPARγ – corepressor crystal structures important insight regarding the off structural state of PPARγ can be inferred from structures of other nuclear receptors bound to corepressors. These structures reveal diversity in the binding modes and structures of both well characterized corepressors such as silencing mediator of retinoic acid and thyroid hormone receptor (SMRT) and nuclear receptor corepressor 1 (NCoR)[20,200] and less characterized corepressors that do not bind to the traditional AF2 coregulator binding surface[297]. Similar to coactivators, each corepressor

contains several amino acid core motifs, or nuclear receptor boxes (NR boxes), of about nine residues that mediate binding to nuclear receptors[20]. SMRT contains two such motifs while NCoR contains three[298]. Most of these core corepressor motifs have been crystalized with various nuclear receptors revealing characteristically longer alpha helices for corepressors than coactivators. This increased helix length[20] requires a different AF2 structure for optimal binding of corepressors that often involves displacement of helix 12, such as in PPARα[200], progesterone receptor[299] and estrogen related receptor-γ[274]. In contrast several nuclear receptor-corepressor crystal structures display helix 12 in a position similar to the canonical active structure with coactivator bound, positioned roughly orthogonal to and making some contact with the corepressor helix n-terminus, but with a shifted helix 12 to accommodate the longer corepressor helix. For example, the crystal structure of the tetrameric form of the RXRα LBD shows interaction between helix 12 of a tetrameric partner and bound SMRT[300] and glucocorticoid receptor (GR) bound to an antagonist shows an interaction between helix 12 and NCoR[301].

We detected active-like corepressor bound helix 12 conformations for PPARγ bound to NCoR using long conventional molecular dynamics simulations (cMD)[282]. Simulations of apo PPARγ show that H12 is pushed away from the coregulatory binding surface. Displacement of helix 12 from the coregulator binding surface has also been observed in numerous structures of nuclear receptors bound to corepressor core motifs[20,200,298,302]. In contrast, simulations of inverse agonist (T0070907[156,282]) bound PPARγ show helix 12 aligned orthogonal to the NCoR peptide with a potentially productive interaction between helix 12 and NCoR[282]. The physiological importance of these conformations is unknown as we observed very slow (μs to second) exchange between different conformations for many of these complexes, including the PPARγ-T0070907 NCoR complex[282]. It is therefore unlikely that these cMD simulations sampled all physiologically relevant conformations. The sampled structures could instead represent structures from a local minimum potential energy well that contributes little to the overall physiologic structural ensemble. We recently presented evidence that two distinct

119

inverse agonists (T0070907 and SR10221) induce ligand specific helix 12 structural states that are different from the ligand free state. We also found that one or a few T0070907 helix 12 states in relatively fast exchange imparts high affinity for NCoR[282,296]. However, the structure(s) that compose this high corepressor affinity structural ensemble remains unknown. In addition, it is not known whether the structural state of other areas of the PPARγ coregulator binding surface contribute to the high affinity of this structural state for NCoR.

Here, we develop and experimentally test simulation derived structural models for these off states of PPARγ utilizing extensive cMD and accelerated MD (aMD) simulations of PPARγ bound to the covalent PPARγ inverse agonist T0070907, SR10221 and co-bound to NCoR. Experimental data and the reconstructed physiological structural ensemble from these more extensive simulations supports the hypotheses that T0070907, SR10221 and apo have distinct structural ensembles that bind corepressors differently. It also supports the idea that T0070907 andSR10221 induce these structural states through different mechanisms. The molecular mechanism by which the inverse agonist T0070907 induces/selects a conformational ensemble with high affinity for corepressor binding is explored using simulations, mutagenesis and NMR. The simulations reveal an n-terminal extension of helix 3 for T0070907 bound PPARγ and disruption of an ionic bond network involving the n-terminus of helix 3. Together these changes prevent interaction of the helix 2-3 loop (i.e. the omega loop) with helix 3 and change the interaction of the helix 11 to 12 loop and helix 12 with the rest of the molecule. Both the salt bridge network and the omega loop helix 3 interaction have been shown to be important for transcriptional agonism[107,303] and disruption of these interactions would be expected to enhance corepressor affinity. In line with this expectation, mutation of a key residue involved in this salt bridge bond network increases NCoR affinity and decreases affinity for a coactivator peptide across many ligands except for T0070907, which independently disrupts this salt bridge network. Previous research indicated interaction between SR10221 and helix 3 that destabilizes helix 3. Here we find evidence of that SR10221 directly interacts with helix 12 to disrupt the active conformation. Thus, we present

evidence for two separate atomically detailed structural mechanisms of inverse agonism in PPARγ. Finally, we demonstrate that the distinct structural ensembles yield distinct functional effects by showing that T0070907 increases affinity for receptor interaction peptides from both NCoR and SMRT as compared to apo while SR10221 selectively increases affinity for SMRT peptide relative to apo.

4.3 Results

4.3.1 Accelerated molecular dynamics reveals a diverse apo PPARγ structural ensemble.

Previously reported adaptive biasing-force simulations of ligand free (apo) and rosiglitazone bound PPARγ ligand binding domain (LBD) explored the helix 12 energy landscape along defined reaction coordinates and found multiple distinct low energy conformations for apo PPARγ but just one for rosiglitazone bound PPARγ[304]. We refer here to the PPARγ LBD as PPARγ. Other forms of PPARγ will be referred to explicitly. In contrast to adaptive biasing-force simulations, accelerated MD simulations (aMD) do not use defined reaction coordinates but instead smooth the native potential energy landscape making sampling more efficient without restricting the conformational space that can be explored[305]. Given sufficient conformational sampling and a force field that reproduces protein structure accurately, the deepest wells in the aMD generated potential energy landscape contain the most prevalent structures found in the physiologic structural ensemble. Ligand free PPARγ and PPARγ bound to many distinct ligands have crystalized as a homodimer with each monomer in a distinct conformation, differentiated principally by the position of helix 12. One helix 12 conformation is considered active, while the other is considered inactive and of uncertain physiologic relevance. However, there are apparent contacts between helix 12 and other crystallographic unit cell members in both conformations in most structures[282]. We ran extensive (**Supplementary Table 1**) aMD simulations of apo and agonist (rosiglitazone) bound PPARγ starting with helix 12 in the active position. Similar to the adaptive biasing-force simulations, the agonist bound simulation produces one relatively narrow energy well, while the apo PPARγ

energy landscape has several energy minima that would be significantly populated at physiologic temperature (**Figure 1a**). Alternatively, we started very extensive aMD simulations (~57 μs) from the inactive apo structure. These simulations sample some of the same structures as the active apo simulation and reveal some similarity in overall energy landscapes, but do not converge completely (**Figure 1a**). Both apo simulations share one low energy well/conformation, which is found in the third (1.06 kcal mol^{-1}) and second (0.07 kcal mol^{-1}) lowest energy well in the inactive and active simulations respectively (**Figure 1c and Supplementary Figure 1**). Overall these simulations indicate that the physiologic structural ensemble of apo PPARγ is very diverse, which is consistent with NMR and hydrogen deuterium exchange mass spectrometry (HDX-MS) data[140,161,166,207,208,221,306]. We observe four major classes of helix 12 conformations in this diverse ensemble, the active conformation, conformations roughly similar to the active conformation, but with a shifted helix 12 that enlarges the coregulator binding surface, conformations where helix 12 appears to bind the coregulator binding surface in a similar manner to a corepressor (which would block all binding to the coregulator binding surface) and finally conformations where helix 12 lies parallel to helix 3 (**Supplementary Figure 1a and b**). Interestingly, in the corepressor like conformation helix 10/11 is disrupted at 484/456 and 477/449 respectively, bending significantly and allowing LXXLLXXXY residues of helix 12 to bind to the coregulator binding surface. These helix 12 residues are oriented similarly to the L/IXXIIXXXF/Y/L motif[20] of corepressors when aligned to a PPARα/SMRT structure (**Figure 1e**). In addition, similar corepressor like binding of helix 12 to the coregulator binding surface have been observed in rat ERβ[307] and between PPARγ unit crystallographic cell members[48,282] (**Supplementary Figure 9**). Together, these helix 12 conformations appear to be a mix of those favorable for coactivator or corepressor binding and some that would favor binding of neither. In addition, both helix 3 and helix 11 show significant structural diversity. In the lowest and fourth lowest energy wells of inactive apo helix landscape to bind to some or all of the coregulator binding surface (**Figure 1e**). Helix 3 is conformationally diverse and in some cases, helicity is absent in the range of

residues 316/288 to 320/29 and in general helix 3 is bent at various angles in this region which leads to diverse positioning of the n-terminus of helix 3 (**Figure 1d**). This is consistent with our previously observed slow exchange (1.0 s^{-1}) between two conformations of a ^{19}F NMR of a probe placed near the location of the bend in helix 3 (Q322/294) in apo PPARγ[282], although this pattern could also be explained by a population of apo PPARγ with helix 12 bound like a corepressor. Finally, there is an n-terminal extension of helix 3 observed in some apo structures (**Supplementary Figure 1a and b**).

4.3.2 The active helix 12 conformation is not the primary component of the physiologic apo ensemble.

Simulations started from the active apo structure predict that almost all of the helix 12 conformations in the physiologic ensemble are similar to the active structure, whereas the inactive simulation suggests that only ~20% of the physiologic ensemble has an active-like helix 12 conformation. Bonding between K347/319 on helix 4 and helix 12 is characteristic of the active helix 12 conformation but not the inactive conformation in PPARγ crystal structures (**Figure 2a**). In order to estimate values characteristic of the physiologic structural ensemble we clustered the structures found near the bottom of the deepest wells in the aMD generated landscape and started 1 μs cMD simulations from multiple representative structures for each well. Population weighted average values for various structural characteristics, including prevalence of the helix 4-12 salt bridge, were then calculated for each well from these cMD simulations and then these individual well values combined into one value representative of the putative physiologic structural ensemble using Boltzmann weighting (see methods). We refer to these values as Boltzmann averages. Simulations predict high prevalence (i.e. high Boltzmann average) for bonding between K347/319 and helix 12 in an active conformation in the PPARγ-rosiglitazone complex. Simulations predict that apo has either high or low prevalence of this interaction depending on whether the apo simulation was started in the active or inactive helix 12 conformation (**Figure 2b**). Simulations indicate that in the absence of helix 12 bonding K347/319 interacts mainly with solvent (**Supplementary Figure 2c**). Therefore, a K347A mutation

should have minimal impact on structure unless there is K347/319 to helix 12 bonding. We compared 2-dimensional protein NMR of wt and K347A mutant bound to various ligands to determine which PPARγ-ligand complexes have helix 12 in the active conformation. Consistent with the inactive apo and the rosiglitazone simulations, these NMR data indicate that helix 12 in apo PPARγ does not interact with K347 while helix 12 of rosiglitazone bound PPARγ does. The K347A mutation induces many changes to the PPARγ-rosiglitazone structure including disappearance of helix 12 residue L496/469 and a shift in a residue near the coregulator binding surface (L429/401) (**Figure 2b**). In contrast, there are few changes to the apo PPARγ structure. This indicates that an active-like helix 12 conformation is not a major component of the apo physiologic ensemble. However, an active-like conformation likely contributes some to the overall ensemble as demonstrated by the inactive apo simulations and changes in a few amide chemical shifts upon mutation of K347 (**Figure 2b**).
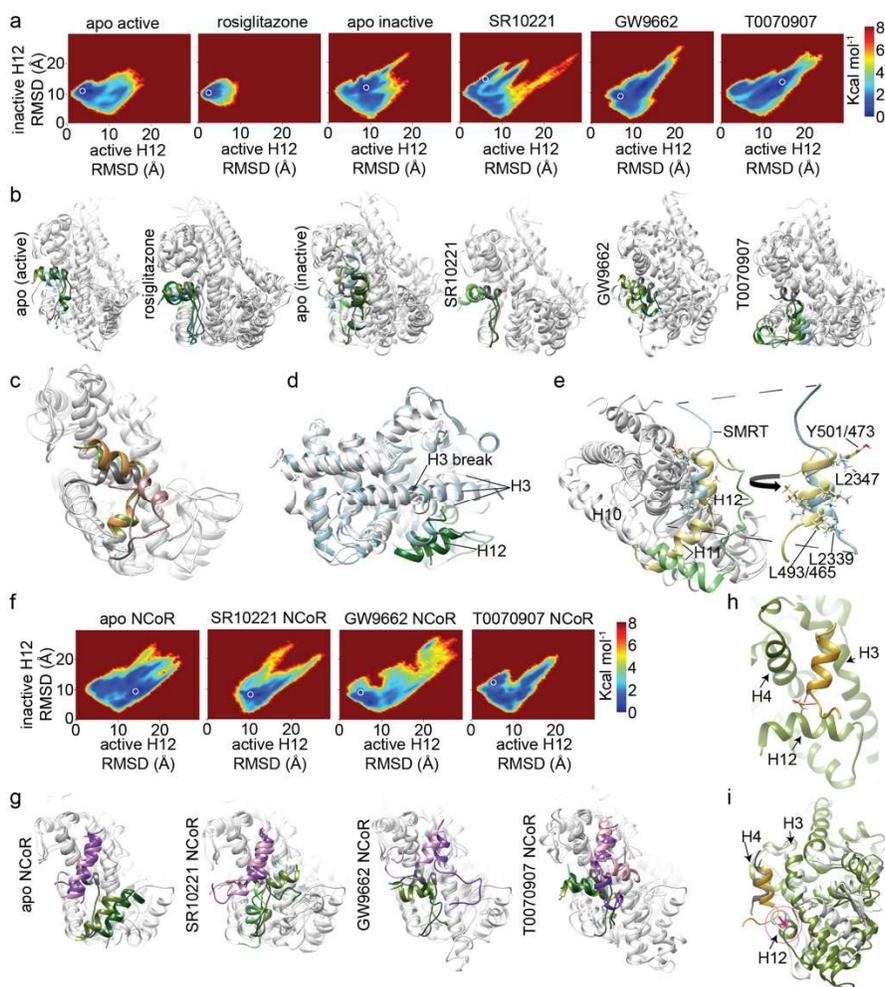
Figure 1. Extensive aMD simulations indicate a diverse PPARγ structural ensemble. Multiple independent simulations were run for the indicated PPARγ ligand binding domain complexes (**see Supplementary Table 1**). a and f) The RMSD of helix 12 was calculated compared to a published apo PPARγ structure (PDB code: 1PRG) active (chain A) and inactive (chain B) structures. The energy of each trajectory snapshot was calculated and overlaid on the 2D RMSD to produce the displayed potential energy landscapes. b and g) The aMD trajectory structures within a 0.2 x 0.2 angstroms RMSD square centered on the lowest energy wells were clustered using k-means clustering into 5 clusters using CPPTRAJ[308]. Arbitrary (i.e. CPPTRAJ chosen) representative structures are shown for the most prevalent clusters. The relative prevalence of the structure is indicated by the color of helix 12 ranging from dark green for the most prevalent to olive drab, light green, light blue and then deep sky blue for the least prevalent. For reference, an active structure (rosiglitazone bound PPARγ; 2PRG chain A) is shown with helix 12 colored dim grey. c) Representative structures from the third and second lowest energy wells in the inactive and active apo simulations respectively are shown in orange and dark green. Representative active and inactive crystal structures (PDB code 1PRG chain a and chain b) are shown in dim grey and light pink respectively. d) Representative structures from active and inactive apo (white and green) are shown compared to an active crystal structure (light blue; PDB code 2PRG chain A). e) A representative structure from a low energy well (well 4 Supplementary Fig. 1; white; PPARγ LBD and gold; helix 11-12) is compared to the crystal structure of PPARα bound to SMRT (PDB code 1KKQ; white; PPARα LBD, green; helix 12 and light blue; SMRT peptide). Comparison of relative position of PPARγ helix 12 LXXLLXXXY residues with the SMRT corepressor motif residues is highlighted. h) Representative structures of PPARγ (green) bound to T0070907 and a peptide from NCoR (gold) from the lowest energy well. The orange lines indicate hydrogen bonds

125

between the helix 12 charge clamp (E499/471) and the NCoR peptide backbone of 1.7 and 2.0 Angstroms as defined by Chimera for this particular trajectory snapshot. i) Comparison of the active structure (white; rosiglitazone; PDB code 2PRG) bound to a coactivator peptide (SRC2; grey) with the most prevalent conformation in the lowest energy well from the aMD simulation of T007-PPARγ-NCoR (green and gold). A shift in helix 12 relative to the active conformation is indicated by the pink arrow.

4.3.3 aMD indicates that inverse agonists bias the structural ensemble towards helix 12 conformations favorable for NCoR binding.

We next performed aMD of PPARγ bound to three inverse agonists/antagonists SR10221[161], GW9662[155] and T0070907 to determine how inverse agonist/antagonist structures differ from apo PPARγ and agonist bound structural ensembles (**Figure 1a**). These simulations were started from structures with helix 12 in the inactive conformation. T0070907 and GW9662 covalently attach to the sole cysteine in the ligand binding domain and differ by one atom; the benzene ring of GW9662 is a pyridine ring in T0070907, making T0070907 a more efficacious inverse agonist[296]. Both covalent ligands increase affinity for a NCoR peptide over apo while SR10221 does not (**Supplementary Table 6**), however only T0070907 and SR10221 act as inverse agonists in a in cell reporter assay[296]. Representative structures from the lowest energy wells demonstrate considerable conformational diversity in helix 3, 11 and 12. Three main low energy helix 12 conformations are detected for the covalent ligands; 1) active-like conformations, 2) active-like conformations with a shifted helix 12 that expands the coregulator binding surface, 3) conformations where helix 12 lies parallel to and near the n-terminus of helix 3, (**Figure 1b and Supplementary Figure 1**). There is also an n-terminal extension of helix 3 observed in some of these structures. Representative structures from the two lowest energy wells for these complexes feature expanded coregulator binding surfaces, in contrast helix 12 would sterically clash with NCoR binding in the lowest energy apo well (**Figure 1b**). The PPARγ-SR10221 complex displays distinct low energy conformations. The lowest energy conformation resembles the active conformation, however helix 12 is shifted along the helical axis toward helix 11 (Figure 1b). Thus, these simulations indicate that the covalent ligands select/induce helix 12 conformations expand the coregulator binding surface in several ways, allowing binding of the longer corepressor peptide. These simulations also indicate distinct helix 12 conformations are induced/selected by

126

SR10221 and GW9662/T0070907, which is supported by fluorine NMR (**Figure 3**).

To experimentally test these simulation-generated structural hypotheses we measured the effect of the K347A mutation on PPARγ bound to SR10221, T0070907, and GW9662. Bonding between K347 and helix 12 is more frequent in active-like conformations (**Supplementary Figure 2e**). The K347A mutation has a large effect on the PPARγ/GW9662 complex, a small effect on the PPARγ-T0070907 complex and no apparent effect on the PPARγ-SR10221 complex (**Supplementary Figure 2a**). Both the GW9662 and T0070907 simulations indicate prevalent low energy conformations with helix 12 in an active-like conformation (**Supplementary Figure 1**) which result in relatively low levels of K347A to E499 bonding (**Figure 2b**). The protein NMR suggests that active-like conformations are more prevalent in the GW9662 complex than the T0070907 complex and that these conformations result in more K347A to helix 12 bonding than observed in the simulations. This raises the possibility that one of the two major helix 12 conformations observed using fluorine[282] and protein[296] NMR (**Supplementary Figure 2b**) is similar to the active state. As shown below, fluorine NMR indicates that 36% of the T0070907 helix 12 structural ensemble is found in this active-like state while 83-88% of helix 12 in the PPARγ-GW9662 structural ensemble is in this active-like state[282,296]. The lack of change in the PPARγ-SR10221 spectrum supports the idea that this ligand induces a distinct helix 12 conformational ensemble from the other two inverse agonist/antagonist ligands. Which is also supported by the distinct PPARγ-SR10221 fluorine NMR spectrum (see below) and distinct low energy conformations observed in simulations (**Supplementary Figure 1**).
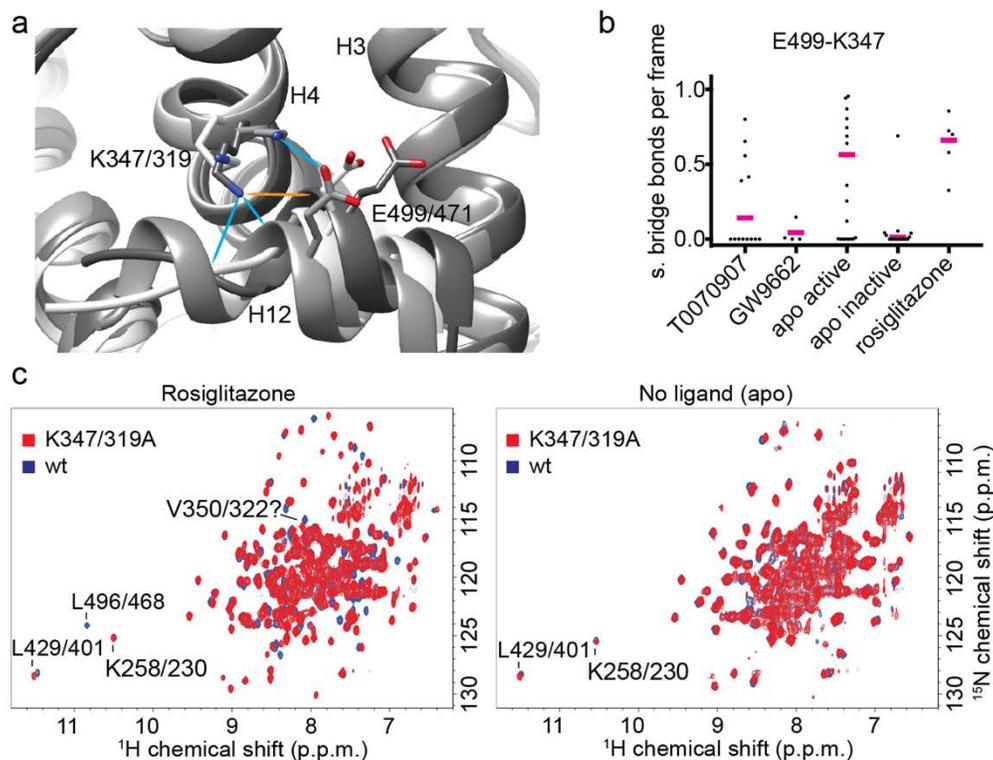
Figure 2. The active apo simulation is inconsistent with NMR data. a) Examples of three classes of helix 4 to helix 12 binding observed in 31 randomly chosen crystal structures with helix 12 in the active conformation (PDB codes 1PRG, 2PRG and 3R5N). PPARγ LBD structures are commonly asymmetric homodimers with helix 12 in an active and inactive conformation. The salt bridge between K347/319 and E499/471 was only seen in 1/31 active structures, while 25/26 active structures showed bonding between K347/319 and the helix 12 backbone and none of the inactive conformations showed bonding between helix 4 and helix 12. b) Prevalence of helix 4 (K347/319) to helix 12 (E499/471) bonding (side chain and backbone), which dominated the helix 4 to helix 12 bonding in active-like structures in simulations. The Boltzmann average is shown as a pink bar while individual values from cMD simulations started from representative structures in the lowest wells in the aMD generated energy landscape are shown as closed circles. Only prevalence of E499 to K347 interaction is shown as this interaction is a good indicator of the active-like conformation (**Supplementary Figure 1g**). c) Comparison of wt and K347/319A mutants bound to rosiglitazone or with no ligand.

4.3.4 Ligands induce/select functionally distinct helix 12 structural ensembles.

We gathered solution structural information for diverse ligand-PPARγ complexes and apo PPARγ using a fluorine NMR probe, 3-bromo-1,1,1-trifluoroacetone (BTFA)[309], attached via an introduced cysteine at three locations; on helix 12 (K502/474C), on Q322/294C between the helix 3 charge clamp lysine (K329/301) and the site of simulation observed helical disruption (316/288-320/292) and on a portion of the omega loop (Q299/271C) near helix 3. We BTFA label PPARγ ligand binding domain in the presence of bound ligand to minimize or

eliminate labeling of the only native cysteine (which points into the ligand binding pocket), alternatively for apo we also use a double mutant that includes a C313A mutation. Fluorescence polarization-based coregulator peptide recruitment assays demonstrated that the K502C-BTFA probe does not have a significant effect on peptide affinity[282], Q299C has no apparent effect on NCoR or MED1 binding (**Supplementary Table 2**), and Q322C causes a 2 fold median increase in affinity for MED1 and no apparent change in affinity for NCoR peptide (**Supplementary Table 3**).

^{19}F NMR of the K502C-BTFA construct confirmed our previous observations[282] that T0070907, GW9662 and the non-covalent inverse agonist SR10221[161,296] induce distinct helix 12 structural ensembles from each other and from apo, although T0070907 and GW9662 differ only in relative populations of the two main structural ensembles (**Figure 3a**). Given that T0070907 induces higher affinity than GW9662 for a nuclear receptor interaction domain peptide from NCoR[296] (**Supplementary Table 2, 3 and 6**) and the T0070907 spectrum has a larger left (i.e. downfield) population than GW9662, this narrower left peak likely represents an ensemble of closely related structures with high affinity for NCoR. This is supported by the fact that addition of SMRT or NCoR increases the population of the left peak in both GW9662 and T0070907 (**Figure 3**) as we previously reported[282]. Objective deconvolution[219] of these spectra indicate that this left peak is relatively narrow (43 Hz) and similar in width to agonist bound PPARγ (rosiglitazone; 32 Hz) and the primary SR10221 bound PPARγ peak (41 Hz) and much narrower than the broad deconvoluted apo peaks which range from 97 to 144 Hz (**Figure 3**). The relatively narrow peaks are consistent with the hypothesis that these peaks represent structures exchanging on relatively fast timescales (e.g. ps to lower microsecond lifetimes) between similar helix 12 conformations while the apo peaks indicate exchange between more distinct helix 12 structures exchanging on the μs-ms time scale[223]. Clearly resolved separate peaks observed in these spectra indicate very slow exchange (ms-s time scale) which we previously confirmed for the two prominent PPARγ-T0070907 peaks[282] and which is consistent with exchange between conformations with larger

structural differences[223]. Together, these data with other NMR and HDX-MS[207,282,296] indicate that agonist (i.e. rosiglitazone) bound PPARγ has one primary narrow structural ensemble, while apo PPARγ and PPARγ bound to T0070907, GW9662 and SR10221 have at least two and very likely more distinct helix 12 structural ensembles. However, these data also support the idea that T0070907 and NCoR co-binding to PPARγ induce/select a primary helix 12 structural state composed of similar structures with high affinity for NCoR and two or more other structural states with lower affinity for NCoR.

4.3.5 The apo helix 3 structural ensemble is composed of structurally and functionally diverse structures.

While the helix 12 conformational ensemble has a large impact on coregulator affinity[282], we hypothesized that the helix 3 structural ensemble is influenced by ligands and also plays a role in coregulator affinity. The fluorine probe on helix 3 of apo PPARγ shows two distinct wide peaks, which chemical exchange saturation transfer (CEST[310]) indicates undergo slow exchange at a rate of less than 0.5 per second for both PPARγ[Q322C]-BTFA (**Figure 3 and Supplementary Figure 3**) and PPARγ[C313A,Q322C]-BTFA[282]. We routinely delipidate PPARγ, which binds *E. coli* lipids[218]. The right (upfield) apo peak originates, at least partially, from the *E. coli* lipid bound state and the observed exchange could be reflective of either lipid exchange or exchange between apo structural states. In contrast to apo, our previous work indicates that the structure and dynamics of PPARγ bound to the ligands utilized here is dominated by the ligands with any residual lipid causing minimal effects[282]. There is no indication that the two states have grossly different affinity for NCoR or CBP as both peaks are in clear slow exchange with the peptide bound peak (**Supplementary Figure 3**). Regardless, the left apo peak is almost certainly lipid free and is wider (~70 Hz) than spectral peaks for PPARγ bound to efficacious ligands (e.g. rosiglitazone and T0070907 primary peaks which are ~30 Hz) indicating that the left apo peak represents two or more distinct structures exchanging on the μs to ms timescale. The ratio of chemical shift separation between the states (350 Hz) to the rate of

exchange between states (<0.5 s^{-1}) is <0.001, thus peak broadening is not due to exchange between the left and right states. Addition of coactivator, corepressor or a heterodimer partner of PPARγ (RXRα) all lead to peak consolidation (**Figure 3 and Supplementary Figure 3**), indicating that the structural ensemble of the portion of helix 3 containing the charge clamp contains a mix of structures with no clear preference for binding coactivators or corepressors.

Ligands induce/select functionally distinct helix 3 structural ensembles from the apo ensemble.

Addition of corepressors (NCoR or SMRT) to PPARγ^{Q322C}-BTFA led to narrowing of primary peaks in inverse agonist (T0070907) bound PPARγ spectra and broadening for labeled PPARγ bound to an agonist (rosiglitazone). Likewise, the binding of a coactivator peptide (MED1 or CBP) has little effect on the spectra from agonist bound labeled PPARγ but induces peak splitting and broadening in inverse agonist/antagonist bound labeled PPARγ (**Figure 3** and **Supplementary Figure 3**). Heterodimerization with RXRα reduces the affinity of the PPARγ T0070907 complex for NCoR[282] and thus would be expected to have a similar structural effect to coactivator binding on the PPARγ^{Q322C}-BTFA spectrum. This is what is observed (**Supplementary Figure 7a**). The MED1/PPARγ^{Q322C}-BTFA/T0070907 complex has the highest dissociation constant of any of these complexes (20 μM; **Supplementary Table 2 and 3**). PPARγ and RXRα have a very low dissociation constant (<10nM). At NMR concentrations (150 μM protein and 300 μM peptide) even the lowest affinity complex should be $>90\%$ bound. Therefore, peak splitting is not likely due to unbound and bound species in the case of the poorest affinity peptides and certainly not in the heterodimer. Rather it appears that binding of coactivators or RXRα to inverse agonist bound PPARγ shifts the equilibrium towards minor conformations. This argument is especially compelling in the case of PPARγ bound to GW9662, where CEST reveals exchange between downfield (left shifted) minor states with the major state and major downfield peaks appear upon addition of peptides or RXRα (**Supplementary Figure 7**). We confirmed

exchange between the two peaks in the PPARγ/GW9662/RXRα complex spectrum using CEST (1.7 s^{-1} 95% CI=1.4-2.0; **Supplementary Figure 7**). Interestingly, in contrast to PPARγ bound to the most efficacious ligands (T0070907 and rosiglitazone), binding of either a coactivator or a corepressor peptide to both apo and a partial agonist (nTZDpa) cause similar changes (**Figure 3 and Supplementary Figure 7**). These data indicate that efficacious ligands select/induce distinct states with a strong preference for coactivators or corepressors, while apo and partial agonist structural ensembles are composed of structures with no strong preference.

4.3.6 Ligands narrow the energy well of the helix 3 c-terminus.

Representative structures from the lowest energy wells of the inactive apo PPARγ simulations, with or without NCoR peptide, show more variability in the position of the c-terminal portion of helix 3, including Q322 and the charge clamp, than PPARγ bound to ligands with or without NCoR peptide (**Figure 3b and c**). The variability in this region is likely functionally significant as it varies the position of the backbone of the K329 charge clamp (**Figure 3d**). Simulations also indicate lower helix 3 helicity for some low energy apo PPARγ structures than ligand bound PPARγ with or without NCoR bound (**Supplementary Figure 3**). Consistent with fluorine NMR and simulations, HDX-MS shows much faster exchange of amide protons in a peptide that spans residues 322 to 326 for the apo PPARγ/NCoR complex than the PPARγ/T0070907/NCoR complex and slightly faster exchange in the PPARγ/SR10221/NCoR complex compared to the T0070907 complex (**Figure 4** and **Supplementary Figure 8**). These data indicate that ligands that induce the highest affinity for coregulators induce/select a c-terminal helix 3 structural ensemble distinct from apo and less efficacious ligands. Together, these data imply the structure of the c-terminus of helix 3 is ligand dependent and functionally important; importance that likely lies in positioning and/or stability of the helix 3 charge clamp.
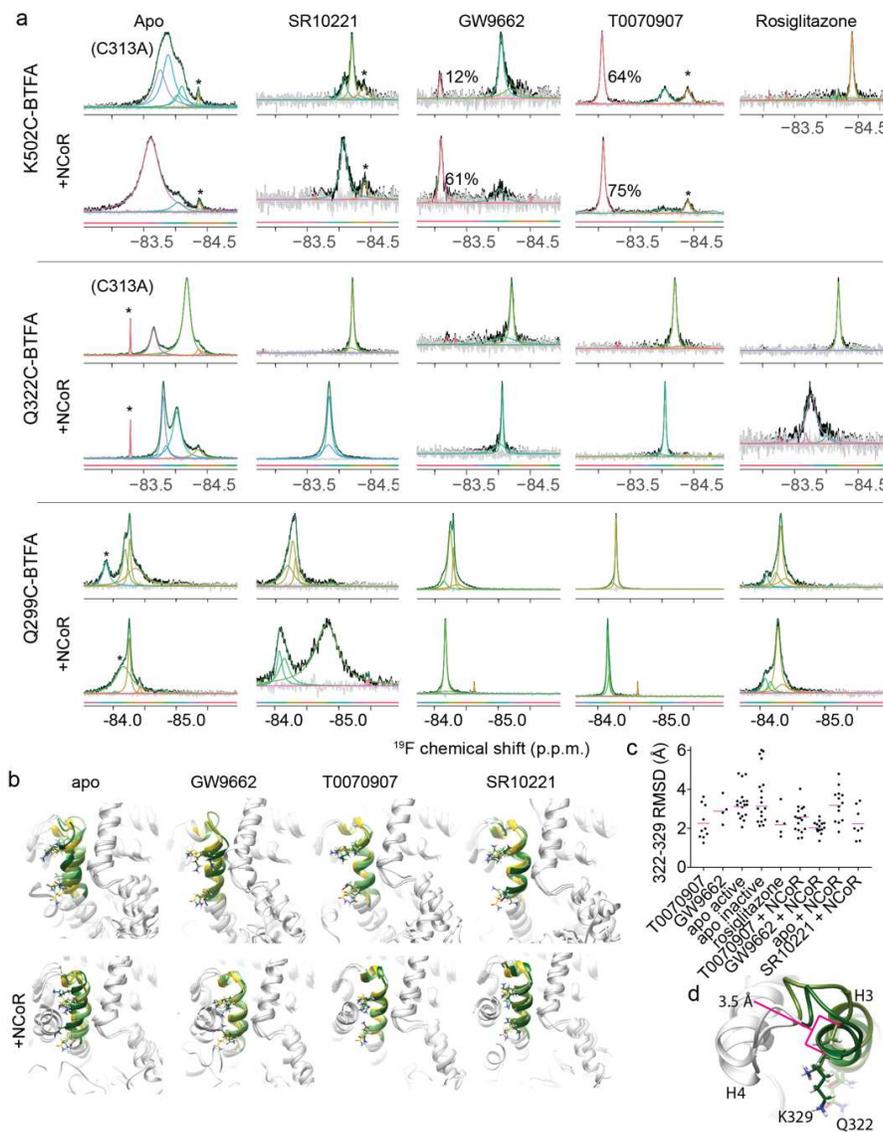
Figure 3. 19F NMR, HDX-MS and simulations indicate that helix 3 has a ligand dependent diverse conformational ensemble that affects the helix 3 charge clamp position. a) ^{19}F NMR spectra of PPARγ labeled at the indicated residues with BTFA and bound to the indicated ligands. The lower rows of spectra included an 2x molar ratio of corepressor peptide derived from NCoR (unlabeled). Asterisk denotes what is likely unfolded PPARγ or BTFA labeled contaminating protein[282]. b) The c-terminal portion of helix 3 is highlighted in color. Dark green, olive drab and light green indicate that the structures are a representative structure from the lowest, second lowest and third lowest energy wells in the aMD energy landscape for the indicated PPARγ-ligand or PPARγ/ligand/NCoR complexes. A representative structure from the rosiglitazone PPARγ simulation is shown in yellow. The side chains of the charge clamp (K329) and the BTFA probe location (Q322) are shown. c) RMSD of residues 322-329 of individual cMD simulations that were started from representative structures from the lowest energy wells for the indicated complexes compared to the same residues in the crystal structure 1PRG (apo PPARγ). Boltzmann averages are shown as magenta bars. The Boltzmann average is shown as a magenta bar. d) Comparison of representative structures from the lowest (dark green) and second lowest (olive drab) energy wells of inactive apo PPARγ. The distance between the alpha carbon atoms for the charge clamp residue K329 for the two structures is indicated.

133
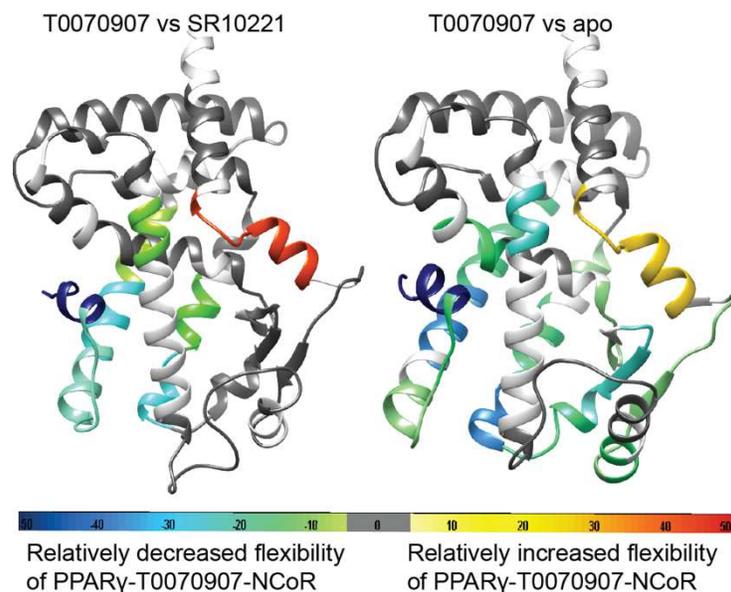
Figure 4. HDX-MS of PPARγ co-bound to T0070907 and a NCoR peptide (2:1 molar ratio NCoR:PPARγ) compared to PPARγ bound to NCoR alone (right) or cobound with SR10221 (left). The difference in hydrogen deuterium exchange for peptides from the indicated regions of the protein is shown by the coloring as indicated. The dark grey regions showed no significant difference, while the white regions were not resolved in the assay.

4.3.7 A single omega loop structure is selected/induced by inverse agonist and corepressor binding.

The spectrum produced by a probe on the omega loop (Q299/272C) near the n-terminus of helix 3 of PPARγ produces a narrow peak (72% of signal) and a wider peak (28% of signal) for the PPARγ-T0070907 complex and more complex spectra for other complexes (**Figure 3a**). The dominant narrow peak indicates that the loop region is stabilized by T0070907 into a single structure in 72% of the population. Fifteen percent of the GW9662 and apo spectra originates from a peak with very similar chemical shift and width to this narrow T0070907 peak. Addition of NCoR or SMRT corepressor peptides drives the omega loop in the GW9662 and T0070907 complexes to a single peak, which is consistent with an ensemble of closely related structures, while the changes to apo are difficult to interpret because of native cysteine labeling (**Figure 3a**). These data indicate that the omega loop is allosterically linked to the coregulator binding surface in these complexes and suggest that a specific omega loop conformation is characteristic of a structural ensemble with the highest affinity for NCoR. In contrast, the

spectrum of Q299-BTFA PPARγ bound to the efficacious agonists rosiglitazone or GW1929 does not change upon addition of coactivators or corepressors indicating a lack of allosteric linkage between the omega loop and the coregulator binding surface for this complex (**Figure 3a and Supplementary Figure 4**). Simulations indicate that in some low energy apo, GW9662 and T0070907 structures (with or without NCoR) helix 3 extends n-terminally to near the fluorine probe residue (299; **Supplementary figure 4b**), which would be consistent with the narrow PPARγ^{Q299C}-BTFA/T0070907 NMR spectrum (**Figure 3a**). Three of 174 PPARγ crystal structures show a 4-residue helix 3 n-terminal extension; a few others show shorter extensions (**Supplementary Figure 4d**). However overall, there is not a good correlation between the Boltzmann average and the population of the narrow peak in the ^{19}F NMR spectrum. In addition, Hydrogen deuterium exchange mass spectrometry (HDX-MS) of an omega loop peptide that covers this region does not show differences between PPARγ/T0070907/NCoR and PPARγ/SR10211/NCoR or apo PPARγ/NCoR complexes (**Figure 4, Supplementary Figure 9**). Indicating that helical extension may not be different between these three complexes. The overall helicity of the region included in the HDX-MS omega loop peptides appears different between the complexes in the simulations (**Supplementary Figure 4f**). Either this difference is not detected by HDX-MS or these simulation structures are representative of intermediates along the path to an as yet unsampled NCoR bound structure. One consequence of Helix 3 extension is reduction of the interaction between the omega loop (i.e. the large loop between helix 2 and helix 3) and helix 3 and/or helix 12 (**Supplementary Figure 6**). Interaction of the omega loop with helix 3 residues has been shown to be important for the transcriptionally active state[106,107]. In fact, SR10221 may disrupt this helix 3 omega loop interaction by pushing on Phe310[161], essentially the same mechanism via different means. Thus helix 3 extension may free up helix 12 positioning as one step towards a final structure. Another consequence of helical extension in both the crystal structures and the simulations appears to be disruption of a critical tripartite salt bridge which we detail below.
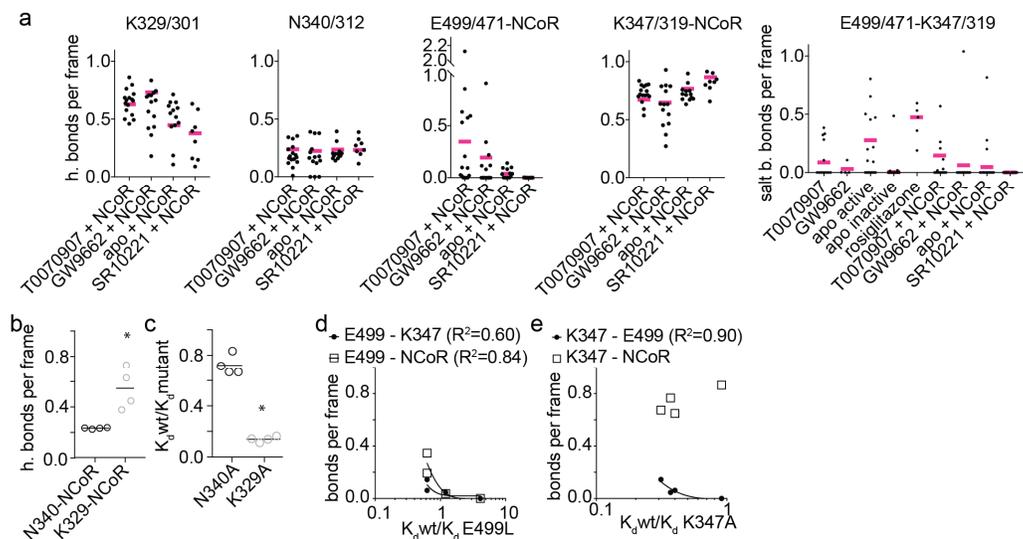
Figure 5. The reconstructed physiologic ensemble (from aMD inspired cMD) indicates that helix 12 bonds to NCoR in the T0070907 complex but not the SR10221 complex. a) The prevalence of the indicated hydrogen and salt bridge bonds in individual 1μs conventional MD runs of representative structures from low energy wells (black circles) and the overall Boltzmann weighted average (pink) for the indicated complexes. b) The Boltzmann average prevalence of hydrogen bonding between the charge clamp residue (K329) and NCoR is significantly higher than between a helix 4 residue (N340) and NCoR. c) The decrease in affinity is greater for a mutation abolishing the helix 3 charge clamp (K329A) than for one abolishing the helix 4-NCoR bonding (N340A). d) Mutation of the helix 12 charge clamp (E499L) leads to ligand dependent decrease in affinity for NCoR, which correlates with simulation derived prevalence of bonding between E499 and both NCoR and helix 4 (K347). e) Abolishing the helix 4 to helix 12 and NCoR interaction through mutation (K347A) reduces affinity for NCoR in a ligand dependent manner, but only the prevalence of the K347-E499 bonding correlates with the decrease. Fits for panels d and e were done with a single exponential as this is the correlation expected between the ratio of dissociation constants and bond prevalence[311].

4.3.8 The structural mechanism of distinct corepressor bound affinities.

The PPARγ-T0070907 complex has the highest affinity for the NCoR peptide out of all the simulated complexes, including the inverse agonist SR10221[296] (**Supplementary Table 6**). To determine the underlying mechanism for how T0070907 induces a high affinity structure for NCoR we analyzed aMD simulations of apo PPARγ, and PPARγ bound to inverse agonist/antagonist (T0070907, GW9662 and SR10221) co-bound to a peptide from NCoR. These simulations were started from structures with helix 12 in the inactive conformation

because docking of NCoR onto PPARγ with helix 12 in the active conformation results in steric clash between NCoR and helix 12. Representative structures from the deepest wells indicate that the structure of these PPARγ NCoR complexes are ligand dependent, with the most prominent differences found in the omega loop, helices 3, 11 and 12 (**Figure 1d and Supplementary Figure 5**). As expected from the PPARα-SMRT structure[200], Boltzmann average values indicate prevalent hydrogen bonding between the NCoR peptide and Lys329/301 of helix 3 and Asn340/312 of helix 4 (**Figure 5a**). Interestingly, low energy ensembles of T0070907/NCoR and GW9662/NCoR, but not SR10221/NCoR, show hydrogen bonding between the NCoR backbone and helix 12 similar to the active conformation, but with helix 12 shifted to allow binding of the longer corepressor helix (**Figure 1 and 5** and **Supplementary Figure 5**). The importance of the charge clamp residues (Lys329/301 and E499/471) for coactivator binding has been shown extensively for several different nuclear receptors, including PPARγ[22], however the helix 12 clamp (E499) has not been proposed before to be important for corepressor binding to nuclear receptors including PPARγ. The helix 12 charge clamp was shown to interact with, but not hydrogen bond to, NCoR in a glucocorticoid receptor antagonist structure[301], where helix 12 is shifted in a similar manner to accommodate the longer compressor helix[42]. In addition, these simulations indicate that low energy helix 12 conformations include disordered and/or coregulator binding surface displaced helix 12 conformations similar to those previously observed in various other corepressor bound crystal structures[20,200,299,300,302].

As noted above, simulations indicate that residues K329/301 (Helix 3) and N340/312 (Helix 4) interact with NCoR, while E499/471 (Helix 12) and K347/319 (Helix 4) can interact both with NCoR and form a salt bridge with each other in a ligand dependent manner. A series of point mutants of the PPARγ-LBD were generated in these four residues to test these simulation-generated structural hypotheses. In addition we used a deletion mutant, L496stop (referred to herein as ΔH12), which deletes helix 12[312,313] and increases affinity for corepressor proteins[314]. The change in affinity upon mutation of K329/301 or N340/312 should correlate (exponentially) with hydrogen bond prevalence as simulations

indicate that the hydrogen bond partners to these residues on NCoR are solvent exposed in the mutants (**Supplementary Table 4**) and these residues do not interact with other PPARγ residues and are thus less likely to affect the native state structure (e.g. are non-disruptive[315]). Simulations indicate that the K329-NCoR hydrogen bond is more prevalent than the N340-NCoR hydrogen bond across the tested complexes (**Figure 5b**). Consistent with the simulations, mutation of K329 has a significantly larger effect on affinity for NCoR than mutation of N340 (**Figure 5c**).

Interpretation of the E499 and K347 mutation data is less straightforward because simulations indicate that these two residues not only hydrogen bond to NCoR but also form a salt bridge with each other in a ligand dependent manner. Mutation of these residues may impact PPARγ structure in addition to NCoR binding. The prevalence of both E499-K347 and E499-NCoR hydrogen bonding alone and summed correlates with the effects of E499L mutation on NCoR affinity indicating that both interactions contribute to NCoR affinity. In contrast, the prevalence of K347-NCoR hydrogen bonding does not correlate with mutagenesis results. However there is the expected exponential correlation[316] between K347-E499 bonding prevalence and affinity changes in the K347A mutant across the four complexes. This suggests that either disruption of the K347-E499 salt bridge affects helix 12 conformation and may obscure the effects of K347-NCoR bonding or that the K347-NCoR bonding observed in simulations is absent or inconsequential in the actual complex. The E499L mutant has increased affinity for NCoR in the ligand free (apo) and SR10221 bound forms which are dominated by structures with no hydrogen bonding between helix4/NCoR and helix 12 (**Figure 5a**). Thus, the increased affinity of the apo and SR10221 bound E499L mutant for NCoR is consistent with the E499L mutation lowering the energetic cost of helix 12 displacement through disruption of the E499-K347 salt bridge. Consistent with this model, deletion of helix 12 had no effect on the affinity of the PPARγ-T0070907 complex for NCoR but increased affinity of NCoR for GW9662 bound PPARγ and dramatically increased affinity for apo and PPARγ bound to SR10221 (**Supplementary Table 5**), which is a

pattern consistent with the prevalence of both the helix 4 to helix 12 (K347 to E499) salt bridge prevalence and the E499 to NCoR hydrogen bonding prevalence in these complexes. These data indicate that both displacement of helix 12 from the coregulator binding surface and positioning of helix 12 that allows direct binding to a corepressor (**Figure 1d and Supplementary Figure 5**) can both increase NCoR affinity.
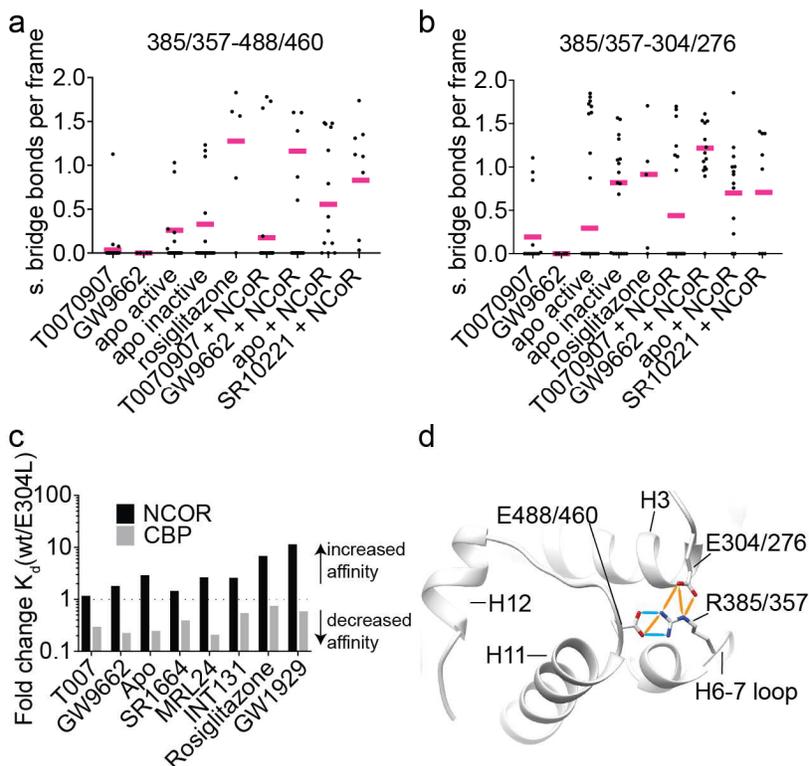
Figure 6. A tripartite salt bridge that biases towards agonism is disrupted in the T0070907 bound structural ensemble. Prevalence of the helix 6-7 loop to a) helix 11-12 loop and b) helix 3 salt-bridges are shown for the indicated complexes for individual cMD simulations started from representative structures from the lowest energy wells in the aMD potential energy landscapes (black circles). The Boltzmann average is shown by a magenta bar. c) The change in affinity induced by mutation of the helix 3 salt bridge residue (E304) is shown for the indicated complexes. d) The tripartite salt bridge as detected by chimera in a PPARγ bound to rosiglitazone crystal structure (PDB code 2PRG). Blue indicate ideal hydrogen bonds, while orange lines indicate relaxed constraint hydrogen bonds (20 degrees and 0.4 angstroms).

4.3.9 A salt bridge switch for inverse agonism.

We next used the simulations to identify the possible mechanisms by which T0070907 induces higher affinity for NCoR peptide than apo PPARγ. A

tripartite salt bridge network was identified in active but not inactive structures in analysis of >200 PPARγ crystal structures[294]. A mutation (F388L) found in familial partial lipodystrophy disrupts this salt bridge network and reduces transcriptional activation potency for drugs[303]. This salt bridge network ties together a very mobile domain of the ligand binding domain[140,207] which includes the n-terminal part of helix 3 (E304/276), the adjacent helix 6-7 loop (R385/357) and the helix 11-12 loop (E488/460; **Figure 6d**). The T0070907 complex with or without NCoR shows the lowest Boltzmann average prevalence for both these salt bridges (**Figure 6 a,b**). We analyzed the simulations and found a correlation between helix 3 extension and interaction of the omega loop with helix 3 and the integrity of this tripartite salt bridge (**Supplementary Figure 6 a,b,e**). In addition, analysis of all available PPARγ crystal structures indicates that this tripartite salt bridge is more prevalent when helix 12 is in the active conformation than when it is in an inactive conformation (**Supplementary Figure 6d**) . It appears that the PPARγ-T0070907 structural ensembles are biased against the helix 11-12 loop to helix 6-7 loop salt bridge formation in ways in addition to helix 3 extension as 3/7 of the PPARγ-T0070907 ensembles without helix 3 extension lack this salt bridge. Thus, T0070907 disrupts this tripartite salt bridge network, which would be expected to allow alternative conformations of the helix 11-12 loop and thereby helix 12, impacting coregulator binding.

To experimentally test the importance of E304/276 to coregulator affinity we generated an E304L mutant and tested the effect on affinity for a FITC-NCoR peptide using fluorescence polarization. These data confirm the importance of the E304/276-385/R357 salt bridge for the active high coactivator affinity state as disruption of this salt bridge increases NCoR affinity and decreases CBP affinity consistently for all complexes except the T0070907 complex (**Figure 6c**). These data are consistent with our simulation generated models which indicate that the inverse agonist T0070907 disrupts this salt bridge network, inducing a unique conformation with increased affinity for NCoR.

4.3.10 The distinct structural ensembles of SR10221 and T0070907 yield distinct function.

^{19}F NMR of helix 12 and helix 3 (**Figure 3** and **Supplementary Figure 4 and 10**), protein NMR (**Supplementary Figure 2**) and simulations (**Figure 1 and 5**) indicate that the inverse agonists SR10221 and T0070907 induce/select distinct NCoR bound conformations. Comparison of PPARγ co-bound to T0070907 and NCoR peptide with the apo NCoR and SR10221-NCoR complexes using hydrogen deuterium exchange mass spectrometry (HDX-MS) also reveals major structural differences (**Figure 4**). The T0070907 complex shows much less exchange in helix 12 and less exchange in helix 10/11 and the 11-12 loop than both apo/NCoR and SR10221/NCoR co-bound to PPARγ. Together these data indicate that SR10221 and T0070907 induce different structures and may achieve inverse agonist effects via distinct mechanisms.

An important functional effect of drug binding to PPARγ is a change in affinity for coregulators. We tested the effect of SR10221, T0070907 and other PPARγ ligands on PPARγ affinity for peptides from NCoR and SMRT using fluorescence polarization. Compared to apo PPARγ, T0070907 binding induces a dramatic increase in affinity for both NCoR (**Supplementary Table 2, 3, 6**) and SMRT (**Supplementary Table 7**) while SR10221 increases affinity for SMRT only (**Supplementary Table 7**). Thus, differences in the structural ensemble of helix 12, 3 and possibly other regions of PPARγ lead to differential recruitment of these corepressor peptides. Differential recruitment of NCoR and SMRT would be expected to produce unique functional effects *in vivo*[317].

4.4 Discussion

These fluorine and protein NMR, HDX-MS, mutagenesis and extensive enhanced sampling simulations reveal a very diverse ligand dependent PPARγ conformational ensemble. These data also demonstrate that helix 3 dynamics and structure is key to coregulator affinity. They also implicate a particular omega loop structural state as important for a structural ensemble with high affinity for the corepressor NCoR. Previous HDX-MS and NMR work indicated that

stabilization of PPARγ helix 3 and the βsheet region, and not just helix 12, is important for transcriptional activity[221,318]. Remarkably, the PPARγ[Q299C]-BTFA (omega loop) spectrum dramatically narrows upon binding the inverse agonist T0070907 or when co-bound to the antagonist GW9662 and NCoR or SMRT peptides. These narrow NCoR bound spectra are consistent with a helix 3/omega loop structural ensemble of similar conformations and indicate that T0070907 induces/selects a helix3/omega loop structural ensemble that is well suited to NCoR binding as we have observed previously for helix 12[282] and other areas of PPARγ[296] (**Figure 3**). These data also indicate that changes to the AF2 surface can induce changes in the omega loop depending on the ligand bound to PPARγ, indicating that the omega loop conformation is important to function. These results, as well as previously obtained data on helix 12 and other areas of the PPARγ ligand binding domain suggest that T0070907 induces/selects a conformational ensemble that accommodates NCoR binding better than GW9662 as the structural ensemble induced/selected by GW9662 changes more than that of T0070907 upon ligand binding[282,296]. This is consistent with the idea that less of the NCoR free energy of binding is used in altering the conformational ensemble of PPARγ-T0070907 than PPARγ-GW9662 resulting in a higher affinity of NCoR peptide for PPARγ-T0070907.

Simulations indicate that Helix 3 extension would disrupt interaction of the omega loop with helix 3. Evidence for the importance of helix 3-omega loop bonding for the transcriptionally active state has been published previously[107]. In addition, mutation of F488/460, involved in the ionic bond network that is also disrupted by helix 3 extension, is associated with familial partial lipodystrophy[303]. In addition, one member of a tripartite salt-bridge network is located on the n-terminus of helix 3 (near the omega loop). Simulations indicate that T0070907 disrupts this network. Disruption of this network favors corepressor binding. These changes appear to free helix 3 and 12 to move to a new conformation with high affinity for NCoR. Together these data indicate that an important consequence of T0070907 binding is disruption of this tripartite salt bridge.

Surprisingly, the helix 12 charge clamp (E499[22]) appears to have some importance for NCoR binding to inverse agonist (T0070907) bound PPARγ. Simulations indicate that the lowest energy PPARγ-T0070907-NCoR structure is similar to the active structure with a shifted helix 12. HDX-MS (**Figure 3**) and mutagenesis (**Figure 4**) are consistent with this model. A similar helix 12 and corepressor position has been observed via crystallography for glucocorticoid receptor bound to an NCoR peptide[301]. This position allows hydrogen bonding between the helix 12 charge clamp (E499) and the NCoR peptide backbone in a similar manner to the coactivator helix 12 interaction. However, given that our aMD simulations are not completely converged, there may be other lower energy unsampled or poorly sampled structures that are also consistent with our experimental data. In that case this conformation could represent a minor helix 12 conformation that composes 25% of the NCoR bound PPARγ[K502C]-BTFA spectrum.

Interestingly, another inverse agonist (SR10221) appears to work as an inverse agonist by inducing a distinct structural ensemble which does not involve interaction of helix 12 with NCoR or a single omega loop structure. The distinct structural ensembles induced by T0070907 and SR10221 produce different corepressor recruitment profiles. Compared to apo PPARγ T0070907 increases affinity for both SMRT and NCoR, while SR10221 increases affinity for only SMRT. This opens the possibility that SR10221 and T0070907 could have distinct repressive effects in animals and possibly selective physiological effects similar to what has been observed in selective PPARγ modulators such as MRL24 and SR1664[166,306].

4.5 Methods

4.5.1 Protein and ligand 3D-structure preparation

The crystal structure of GW9662-bound PPARγ (PDB code 3B0R) was used as the initial structure in all simulations of GW9662 bound PPARγ in this study. This crystal structure was also used to construct the initial 3D structure of

T0070907 bound PPARγ. In this model, GW9662 was transformed to T0070907 by converting benzene ring of GW9662 to the pyridine ring. In both models, the chain B conformation was used. In order to construct the 3D structure for nuclear receptor corepressor 1 (NCoR1) crystal structure with PDB code 2OVM was used. 1KKQ crystal structure which is the only structure with corepressor (SMART) was used as a template to design NCoR bound to PPARγ-T0070907. Both chain A and B of crystal structure with PDB code 1PRG were used for apo and 10221 (chain B) simulations. To build the initial structure for Rosiglitazone simulations chain A of 2PRG PDB code along with S enantiomer of Rosiglitazone at pH 7.2 which has the highest affinity for PPARɣ were used. UCSF Chimera modeler extension was used to model the missing residues in PDB files.

4.5.2 Molecular Dynamics (MD) simulations

Molecular Dynamics (MD) simulations were performed using two different methods: conventional (cMD) and accelerated molecular dynamics simulation (aMD). aMD simulations were performed in order to sample the conformational spaces better. Both cMD and aMD production runs were carried out using Amber 16 molecular modeling package and the AMBER ff14SB force field and general Amber force field (GAFF2) parameters were used to describe protein and ligand[241,319]. Refined structures were submitted to h++ server (http://biophysics.cs.vt.edu/H++)[237] to generate the protonated states of protein at pH 7.4. The resulting PDB files were modified using pdb4amber in AmberTools14[238] for use with tleap. The R.E.D server (http://upjv.q4md-forcefieldtools.org/REDServer-Development/)[239] was used for the ligand parameterization and charge calculations. The structures were immersed in an octahedron box of TIP3P[243] water molecules extended to 10 Å from the protein atoms. Enough Na+ atoms were added to neutralize the structure and KCl (K+ and Cl- ions) was added to 50 mM[320]. The resulting system was equilibrated using a nine-step of minimization and restrained simulations protocol as following. In the first step a force constant of 5 kcal mol^{-1} Å^{-2} was applied on the protein heavy atoms through 2000 steps. Then, the MD simulation was performed for 15 ps with shake under constant volume periodic boundary conditions (NVT). This was

followed by two rounds of 2000 steps of steepest descent minimization with 2 and 0.1 kcal/mol Å2 spring constant. The system was then subjected to a simulation with no restraints followed by three rounds of simulations with 1, 0.5 and 0.5 kcal/mol Å2 force constant on heavy atoms for 5 ps, 10 ps and 10 ps. Finally, a simulation without restraints was performed for 200 ps under NPT condition. *Hydrogen mass repartitioning* along with SHAKE algorithm were *used* to *allow* an integration *time step* of 4 fs. *Production* MD *runs* of constant pressure *replicates were* performed from *randomized initial velocities*. The pressure was controlled by a Monte Carlo barostat with a pressure relaxation time (taup) of 2 ps. The Langevin dynamics with a collision frequency (gamma_ln) of 3 ps^{-1} was used to keep the temperature at 310 K. The particle mesh Ewald41 with an 8.0 Å cuto□ was carried out to treat electrostatic interactions. The time step of cMD simulations was 100 ps and three independent ~15-μs-long cMD production runs were performed on T0070907, GW9662 and SR10221bound PPARγ with and without NCoR. Three independent simulations with the same length were performed on Rosiglitazone bound PPARγ. All of the aMD simulations were started from the equilibrated structures and average dihedral energy and total potential energy obtained from cMD simulations were used to calculate the boost parameters. In this study a dual boosting approach was carried out in which two separate boost potentials are applied to the torsional and the total potential terms. aMD simulations for ~2.5 μs with time step of 3fs were performed on T0070907 and GW9662 bound PPARγ with and without NCoR.

All production simulations were performed using pmemd.cuda or pmemd.cuda.MPI. The simulation results were analyzed using *cpptraj* program in the AmberTools 14 Toolbox[308]. A toolkit of Python scripts "PyReweighting" was used to reweight the biased aMD frames and to calculate free energy profiles[321].

4.5.3 Protein purification

A pET45b plasmid containing the PPARγ-LBD, residues 230-505 as well as an N-terminal 6x His tag and tobacco etch virus nuclear inclusion protease (TEV) recognition site was transformed into BL21 (DE3) gold cells (Invitrogen). Cells were grown in either terrific broth or ZYP-5052 autoinduction media. In the case

of autoinduction media cells were grown for 10 hours at 37°C and then allowed to induce for an additional 12 hours at 22°C and 180 rpm. Cells in terrific broth were grown at 37°C and 170 rpm until OD_{600} of between 0.6-1 was reached. Following this the incubator was dropped to 20°C for one hour and cells were induced overnight by the addition of 500μM IPTG. Cells were pelleted by centrifugation and stored at -20°C until ready for use. Cell pellets were resuspended in 50mM KPO_4, 300mM KCl, 1mM TCEP and 1mM EDTA pH 8.0 and lysed using a C-5 Emulsiflex high pressure homogenizer (Avestin). Initial protein purification was performed on either an AKTA start (GE Healthcare) or a NGC Scout (Bio Rad) FPLC using 2 His Trap FF 5ml columns in series (GE Healthcare). Following this the 6x his tag was removed by the addition of approximately 1:40 w/w 6x his tagged TEV and overnight incubation. Cleaved tag and TEV protease was removed by again passing through His Trap FF columns. Size exclusion chromatography was then performed using a Hiload 16/600 Superdex 75 pg column (GE Healthcare). Protein purity in excess of 95% was confirmed by SDS-PAGE.

4.5.4 Site-directed mutagenesis

Site-directed mutagenesis was performed using the Quikchange Lightning mutagenesis kit (Agilent Technologies). All mutations as well as the absence of spurious mutations were confirmed by Sanger sequencing (Eurofins). Primers used to generate the mutants used in this work are as follows, C313A: 5'-ccacggagcgaaactgagcgccctgaaagatgcgg-3', Q322C: 5'-gcatactctgtgatctcgcacacagcctccacggagc-3', Q314A: 5'-ctccacggagcgaaacgcgcagccctgaaagatg-3', K329A: 5'-tttacaaaaccaggaatgcttgcggcatactctgtgatctcctg-3', N340A: 5'-atttgaggagagttacttggtcggccaagtcaagatttacaaaaccag-3', K347A: 5'-gatctcgtggactccatatgcgaggagagttacttggtcg-3', E499L: 5'-caagtccttgtagatcagctgcaggagcggtga-3', ΔH12: 5'-tccttgtagatctcctgcagttacgggtgaagactcatgtctg-3'

4.5.5 Preparation of NMR samples

All NMR samples were prepared to a final volume of 470μL and a final concentration of 150μM. Samples of PPARγ^{C313A,Q322C} and PPARγ^{Q322C} were loaded with ligands to a final concentration of 165μM for all ligands except Pioglitazone, Troglitazone and BVT.13 which were loaded to a final concentration of 225μM or GW9662 and T0070907 which were added to a final concentration of 300μM. Ligand concentration was varied to fix the DMSO addition to each sample at 7.8μL, 1.66% final v/v. In the case of PPARγ^{C313A,Q322C} the protein was labeled with a 10-fold excess of BTFA during purification. However, the single mutant PPARγ^{Q322C} was labeled with a 2-fold molar excess of BTFA following drug addition, with the following exceptions, samples loaded with the covalent ligands were labeled with a 10-fold molar excess of BTFA while the apo sample was

labeled with a stoichiometric concentration of BTFA to reduce labeling on the native cysteine in the ligand binding pocket. Following labeling samples were incubated for 30 minutes when loaded with ligand or 2 hours for the apo sample and then buffer exchanged >100x using an amicon ultra centrifugal filter (Millipore) with a 10kDa molecular weight cut off to remove excess BTFA label. All peptides were added to NMR samples at a final concentration of 300μM from stocks of approximately 1mM in the same buffer as the NMR samples. Following sample preparation 10% of final volume of buffered D_2O was added to all samples.

Chapter 5: Conclusions

Presented in this work were a wide variety of experiments to increase the understanding of a fundamental property of the significant nuclear receptor PPARγ. All data in this work revolved around the central theme towards elucidating the conformational mechanisms by which very diverse ligands cause their functional outcomes. From this we were able to expand the available knowledge in this area, but the picture is still incomplete. However, a substantial contribution is made towards the understanding of this area of study which will help guide future research.

This work demonstrates that the two-state "on-off" model for PPARγ, and likely other nuclear receptors, does not appear to be correct. We observed that the conformations induced by the majority of ligands were highly complex. Even in the highly efficacious ligands such as the inverse agonist T0070907 or the FDA approved full agonist pioglitazone there appeared to be two or more significantly populated conformational states of the protein. In fact, when probing helix 12, considered the most important region of the AF-2 surface, it was seen that only a small number of ligands, GW1929, rosiglitazone and SR10221, were generating only a single dominant conformation. Results were similar when probing the conformational ensemble of helix 3 which is another portion of the AF-2 surface. In this case all efficacious agonists or inverse agonists selected for a more specific local conformation, but the majority of ligands showed multiple conformations present at high population. Interestingly, it was seen that both agonists and inverse agonists seemed to select for the same conformation in this region. This has not previously been observed but is not completely surprising as the helix 3 charge clamp is significant in the binding of all coregulators.

Additionally, we observed that inverse agonism does not operate through a single mechanism. This has long been thought to occur in NR through a displacement of helix 12 which then allows room for the larger footprint of the CoRNR box to bind on the AF-2 surface. This hypothesis was due to the fact that when nuclear receptors were crystallized with corepressor peptides helix 12 was

always unresolved and appeared to point into space. Such a mechanism does appear to be relevant in PPARγ inverse agonists, namely SR10221 and nTZDpa, but does not appear to be the exclusive conformation. We identified a second inverse agonist induced conformation which appears to be unique to the covalent ligands, T0070907 and GW9662. In this mechanism the small covalent ligand remodels helix 3 which then exerts allosteric effect on helix 12 through the disruption of intra-protein non-covalent interactions. This allows helix 12 to reposition into a new conformation which allows for the binding of a CoRNR box while retaining productive interactions to the corepressor from the helix 12 charge clamp. Such a mechanism has not previously been reported and the presence of these productive helix 12 interactions is likely why T0070907 is the most efficacious known inverse agonist in this protein, approximately three times more efficacious than SR10221 in SMRT or NCoR recruitment. Though it should be noted that in the monomeric PPARγ SR10221 is a selective inverse agonist and only promotes the recruitment of SMRT not NCoR, while in the PPARγ-RxRα heterodimer it increases affinity for both corepressors.

Perhaps the most interesting finding in this work was the discovery of the helix 4 motif that seems to direct specificity in coactivator interactions. This eight-residue motif makes discriminatory hydrogen bonds from the terminal residues to the side chains of residues on the coactivator peptide that seem to be responsible for a large portion of the extreme specificity seen in coactivator interactions. They also seem to have some significance in corepressor interactions but not to nearly as large of a degree. It could be envisioned where this could be exploited towards the design of biased drugs to recruit only specific coactivators and thereby drive specific signaling events. An example of such would be in PPARγ where the moderate affinity PGC1α protein is relatively unaffected by ablation of helix 4 hydrogen bond interactions. This leads to the idea of the creation of a ligand that specifically alters the conformation of helix 4 while retaining other agonist characteristics. It is reasonable to hypothesize that this ligand would than preferentially recruit PGC1α. Such a design could be exploited similarly in other nuclear receptors to select for desired effects. The caveat would

149

be that such a design could lead to negative side effects. For example, it has been shown that overexpression of PGC1α in mice can lead to increased insulin resistance in the liver[322]. As such it would require special care in drug design to avoid such unforeseen effects. Another caveat is that this could only be exploited to preferentially recruit coactivators with lower affinity since the disruption of helix 4 would severely reduce the affinity for the high affinity coactivators, such as MED1 and CBP. So, while this may be a strategy in drug design which would be slightly limited it could still be of very significant value provided proper care was taken.

One interesting question this work was not able to resolve is the mechanism of action in partial agonists. Literature would suggest that the majority of the insulin sensitizing effects seen with these arises from the modulation of phosphorylation on S245/273 but this is not their only effect. Several of these ligands including MRL24[323], INT131[163] and SR2088[282] have been shown reliably to increase coactivator recruitment, either through transactivation assays or *in vitro* peptide recruitment assays. However, it has previously been shown these make no interaction on helix 12[164], consistent with our own data, and herein we show they do not generate a conformation in the helix 3 portion of the AF-2 distinguishable from non-agonists. The only portion of the AF-2 this leaves they could exert an effect on is helix 4 but according to molecular dynamics and crystallography they do not make significant contacts in this region. It should be stated that there may be direct effects on this region which are not captured by these techniques. To verify this NMR studies similar to those performed in chapters 2 or 3 would need to be employed. Other two or three-dimensional NMR based techniques would likely not be successful in ascertaining these effects since partial agonist binding generally results in a very plastic state of the protein leading to a small amount of resolved peaks which makes assignment impractical and without assigned backbone residues it would not be possible to determine what changes are occurring. If as the initial data suggests there are no significant effects on helix 4 conformation induced by partial agonist binding that would imply that there must be a significant allosteric mechanism exerted by partial

agonists. This would be highly interesting to unravel but may be very challenging. The difficulty lies in the fact that partial agonists make only minimal direct contacts to the protein and those contacts are shared with non-agonists. However, a stepwise mutagenic screen may be able to parse these effects and solve the underlying mechanism. This would be of great value as these ligands are the current focus of active research despite having no real understanding of their mechanism. Such an understanding would allow for the more intelligent design of non-agonist ligands such as SR1664[166], this ligand strongly increases insulin sensitivity but has very poor bioavailability making it unsuitable for medicinal use, which could increase insulin sensitivity but have no effect on other protein signaling pathways. Such drugs would be of great use as potential therapeutics which could yield the desired medical effects but avoid the very serious side-effects associated with full agonist activation of PPARγ.

Through this work we have greatly expanded the knowledge base relating to how ligand binding leads to a functional effect in PPARγ. This is of great value in future guided studies to either create novel ligands or modify existing ligands in a more efficient way. Through this the current drug design for this protein which relies primarily on a somewhat scattershot method of creating hundreds if not thousands of ligands could be made more efficient leading to the more rapid release of potential therapeutics which target this important protein for metabolic and neurodegenerative disease.

Chapter 6: References

1.   Evans, R. The steroid and thyroid hormone receptor superfamily. *Science (80-.).* **240,** 889–895 (1988).

2.   Mangelsdorf, D. J. *et al.* The Nuclear Receptor Superfamily: The Second Decade. **83,** 835–839 (1995).

3.   Krust, A. *et al.* The chicken oestrogen receptor sequence: homology with v-erbA and the human oestrogen and glucocorticoid receptors. *EMBO J.* **5,** 891–897 (1986).

4.   Nwachukwu, J. C. *et al.* Resveratrol modulates the inflammatory response via an estrogen receptor-signal integration network. *Elife* **2014,** (2014).

5.   Srinivasan, S. *et al.* Ligand-binding dynamics rewire cellular signaling via Estrogen Receptor-α. *Nat Chem Biol* **9,** 326–332 (2013).

6.   Arao, Y. & Korach, K. S. The F domain of estrogen receptor is involved in species-specific, tamoxifen-mediated transactivation. *J. Biol. Chem.* **293,** 8495–8507 (2018).

7.   Patel, S. R. & Skafar, D. F. Modulation of nuclear receptor activity by the F domain. *Mol. Cell. Endocrinol.* **418,** 298–305 (2015).

8.   Carter, E. L. & Ragsdale, S. W. Modulation of nuclear receptor function by cellular redox poise. *J. Inorg. Biochem.* **133,** 92–103 (2014).

9.   Wang, Z. *et al.* Structure and function of Nurr1 identifies a class of ligand-independent nuclear receptors. *Nature* **423,** 555–560 (2003).

10.   Zhao, Y. & Bruemmer, D. NR4A Orphan Nuclear Receptors: Transcriptional Regulators of Gene Expression in Metabolism and Vascular Biology. *Arter. Thomb Vasc Biol* **30,** 1535–1541 (2010).

11.   de Vera, I. M. S. *et al.* Defining a canonical ligand-binding pocket in the orphan nuclear receptor Nurr1. *bioRxiv* 1–16 (2018).

12.   Zhang, Y. & Dufau, M. L. Gene Silencing by Nuclear Orphan Receptors. *Vitam. Horm.* **68,** 1–48 (2004).

13.   Iyer, A. K. & McCabe, E. R. B. Molecular mechanisms of DAX1 action. *Mol. Genet. Metab.* **83,** 60–73 (2004).

14.   Feng, W. *et al.* Hormone-Dependent Coactiva tor Binding to a surface. Hydrophobic Cleft on Nuclear Receptors. *Science (80-.).* **280,** 1747–1749 (1998).

15.   Heery, D. M., Kalkhoven, E., Hoare, S. & Parker, M. G. A signature motif in transcriptional co-activators mediates binding to nuclear receptors. *Nature* **387,** 733–736 (1997).

16.   Savkur, R. S. & Burris, T. P. The coactivator LXXLL nuclear receptor recognition motif. *J Pept Res* **63,** 207–212 (2004).

17. Hu, X. & Lazar, M. A. The CoRNR motif controls the recruitment of corepressors by nuclear hormone receptors. *Nature* **402,** 93–96 (1999).

18. Hu, X., Li, Y. & Lazar, M. Determinants of CoRNR-dependent repression complex assembly on nuclear hormone receptors. *Mol Cell Biol* **21,** 1747–1758 (2001).

19. Perissi, V. & Rosenfeld, M. G. Controlling nuclear receptors: The circular logic of cofactor cycles. *Nat. Rev. Mol. Cell Biol.* **6,** 542–554 (2005).

20. Phelan, C. A. *et al.* Structure of Rev-erbα bound to N-CoR reveals a unique mechanism of nuclear receptor-co-repressor interaction. *Nat. Struct. Mol. Biol.* **17,** 808–814 (2010).

21. Darimont, B. D. *et al.* Structure and specificity of nuclear receptor – coactivator interactions Structure and specificity of nuclear receptor – coactivator interactions. **1,** 3343–3356 (1998).

22. Nolte, R. T. *et al.* Ligand binding and co-activator assembly of the peroxisome proliferator-activated receptor-gamma. *Nature* **395,** 137–143 (1998).

23. Huang, P., Chandra, V. & Rastinejad, F. Structural Overview of the Nuclear Receptor Superfamily: Insights into Physiology and Therapeutics. *Annu Rev Physiol* **71,** 247–272 (2010).

24. Khorasanizadeh, S. & Rastinejad, F. Visualizing the architectures and interactions of nuclear receptors. *Endocrinology* **157,** 4212–4221 (2016).

25. Bevan, C. L., Hoare, S., Claessens, F., Heery, D. M. & Parker, M. G. The AF1 and AF2 Domains of the Androgen Receptor Interact with Distinct Regions of SRC1. *Mol. Cell. Biol.* **19,** 8383–8392 (1999).

26. Benecke, A., Chambon, P. & Gronemeyer, H. Synergy between estrogen receptor a activation functions AF1 and AF2 mediated by transcription intermediary factor TIF2. *EMBO Rep.* **1,** 151–157 (2000).

27. Lin, S. J. *et al. TR2 and TR4 Orphan Nuclear Receptors: An Overview. Current Topics in Developmental Biology* **125,** (Elsevier Inc., 2017).

28. Lin, S.-J. *et al.* Minireview: Pathophysiological Roles of the TR4 Nuclear Receptor: Lessons Learned From Mice Lacking TR4. *Mol. Endocrinol.* **28,** 805–821 (2014).

29. Senali Abayratna Wansa, K. D., Harris, J. M. & Muscat, G. E. O. The activation function-1 domain of Nur77/NR4A1 mediates trans-activation, cell specificity, and coactivator recruitment. *J. Biol. Chem.* **277,** 33001–33011 (2002).

30. Codina, A. *et al.* Identification of a novel co-regulator interaction surface on the ligand binding domain of Nurr1 using NMR footprinting. *J. Biol. Chem.* **279,** 53338–53345 (2004).

31. Horwitz, K. B. *et al.* Nuclear receptor coactivators and corepressors. *Mol. Endocrinol.* **10,** 1167–1177 (1996).

32. Ruan, X. Z., Varghese, Z., Powis, S. H. & Moorhead, J. F. Nuclear receptors and their coregulators in kidney. *Kidney Int.* **68,** 2444–2461 (2005).

33. Xu, J. & Li, Q. Review of the *in Vivo* Functions of the p160 Steroid Receptor Coactivator Family. *Mol. Endocrinol.* **17,** 1681–1692 (2003).

34. Ngan, V. & Goodman, R. H. CBP and p300 in Transcriptional Regulation. *Biochemistry* **5,** 1–42 (2003).

35. Yao, T.-P., Ku, G., Zhou, N., Scully, R. & Livingston, D. M. The nuclear hormone receptor coactivator SRC-1 is a specific target of p300. *Biochemistry* **93,** 10626–10631 (1996).

36. Koppen, A. *et al.* Nuclear Receptor-Coregulator Interaction Profiling Identifies TRIP3 as a Novel Peroxisome Proliferator-activated Receptor γ Cofactor. *Mol. Cell. Proteomics* **8,** 2212–2226 (2009).

37. Wallberg, A. E., Yamamura, S., Malik, S., Spiegelman, B. M. & Roeder, R. G. Coordination of p300-mediated chromatin remodeling and TRAP/mediator function through coactivator PGC-1α. *Mol. Cell* **12,** 1137–1149 (2003).

38. Charos, A. E. *et al.* A highly integrated and complex PPARGC1A transcription factor binding network in HepG2 cells. *Genome Res.* **22,** 1668–1679 (2012).

39. Puigserver, P. & Spiegelman, B. M. Peroxisome proliferator-activated receptor-γ coactivator 1α (PGC-1α): Transcriptional coactivator and metabolic regulator. *Endocr. Rev.* **24,** 78–90 (2003).

40. Martínez-Jiménez, C. P., Gómez-Lechón, M. J., Castell, J. V. & Jover, R. Underexpressed coactivators PGC1α and SRC1 impair hepatocyte nuclear factor 4α function and promote dedifferentiation in human hepatoma cells. *J. Biol. Chem.* **281,** 29840–29849 (2006).

41. Viswakarma, N. *et al.* The Med1 subunit of the mediator complex induces liver cell proliferation and is phosphorylated by AMP kinase. *J. Biol. Chem.* **288,** 27898–27911 (2013).

42. Perissi, V. *et al.* Molecular determinants of nuclear receptor – corepressor interaction. 3198–3208 (1999).

43. Hörlein, A. J. *et al.* Ligand-independent repression by the thyroid hormone receptor mediated by a nuclear receptor co-repressor. *Nature* **377,** 397 (1995).

44. Chen, J. D. & Evans, R. M. A transcriptional co-repressor that interacts with nuclear hormone receptors. *Nature* **377,** 454 (1995).

45. Li, P. *et al.* Adipocyte NCoR Knockout Decreases PPARγ Phosphorylation and Enhances PPARγ Activity and Insulin Sensitivity. *Cell* **147,** 815–826 (2011).

46. Rosell, M., Jones, M. C. & Parker, M. G. Role of nuclear receptor corepressor RIP140 in metabolic syndrome. *Biochim. Biophys. Acta - Mol. Basis Dis.* **1812,** 919–928 (2011).

47. Nautiyal, J., Christian, M. & Parker, M. G. Distinct functions for RIP140 in development, inflammation, and metabolism. *Trends Endocrinol. Metab.* **24,** 451–459 (2013).

48. Frkic, R. L. *et al.* PPARγ in Complex with an Antagonist and Inverse Agonist: A Tumble and Trap Mechanism of the Activation Helix. *iScience* **5,** 69–79 (2018).

49. Wang, Y. X. PPARs: Diverse regulators in energy metabolism and metabolic diseases. *Cell Res.* **20,** 124–137 (2010).

50. Olivier, B., Aldridge, T. C., Norbert, L. & Stephen, G. PPAR-RXR HETERODIMER ACTIVATES A PEROXISOME PROLIFERATOR RESPONSE ELEMENT UPSTREAM OF THE BIFUNCTIONAL ENZYME GENE. *Biochem. Biophys. Res. Commun.* **192,** 37–45 (1993).

51. Lam, V. Q., Zheng, J. & Griffin, P. R. Unique Interactome Network Signatures for Peroxisome Proliferator-activated Receptor Gamma (PPARγ) Modulation by Functional Selective Ligands. *Mol. Cell. Proteomics* **16,** 2098–2110 (2017).

52. Poulsen, L. la C., Siersbæk, M. & Mandrup, S. PPARs: Fatty acid sensors controlling metabolism. *Semin. Cell Dev. Biol.* **23,** 631–639 (2012).

53. Barish, G. D. *et al.* PPAR d : a dagger in the heart of the metabolic syndrome Find the latest version : Review series PPAR δ : a dagger in the heart of the metabolic syndrome. *J. Clin. Invest.* **116,** 590–597 (2006).

54. Katsiki, N. *et al.* The Role of Fibrate Treatment in Dyslipidemia: An Overview. *Curr. Pharm. Des.* **19,** (2013).

55. Chen, W. *et al.* A metabolomic study of the PPARδ agonist GW501516 for enhancing running endurance in Kunming mice. *Sci. Rep.* **5,** 1–13 (2015).

56. Cox, R. L. Rationally designed PPARδ-specific agonists and their therapeutic potential for metabolic syndrome. *Proc. Natl. Acad. Sci.* **114,** 3284–3285 (2017).

57. Li, L. *et al.* Expression patterns of peroxisome proliferator-activated receptor gamma 1 versus gamma 2, and their association with intramuscular fat in goat tissues. *Gene* **528,** 195–200 (2013).

58. Medina-Gomez, G. *et al.* PPAR gamma 2 prevents lipotoxicity by controlling adipose tissue expandability and peripheral lipid metabolism. *PLoS Genet.* **3,** 0634–0647 (2007).

59. Saraf, N., Sharma, P. K., Mondal, S. C., Garg, V. K. & Singh, A. K. Role of PPARg2 transcription factor in thiazolidinedione-induced insulin sensitization. *J. Pharm. Pharmacol.* **64,** 161–171 (2012).

60. Lee, J. E. & Ge, K. Transcriptional and epigenetic regulation of PPARγ expression during adipogenesis. *Cell Biosci.* **4,** 1–11 (2014).

61. Peter, T., Erding, H. & Spiegelman, B. M. Stimulation of Adipogenesis by PPARγ2, a Lipid-Activated in Fibroblasts Transcription Factor. *Cell* **79,** 1147–1156 (1994).

62. Brun, R. P. *et al.* Differential activation of adipogenesis by multiple PPAR isoforms. *Genes Dev.* **10,** 974–984 (1996).

63. Barak, Y. *et al.* PPARγ is required for placental, cardiac, and adipose tissue development. *Mol. Cell* **4,** 585–595 (1999).

64. He, W. *et al.* Adipose-specific peroxisome proliferator-activated receptor ⬚ knockout causes insulin resistance in fat and liver but not in muscle. *Proc. Natl. Acad. Sci.* **100,** 15712–15717 (2003).

65. Imai, T. *et al.* Peroxisome proliferator-activated receptor gamma is required in mature white and brown adipocytes for their survival in the mouse. Proc Natl Acad Sci U S A 101, 4543-4547 (2004)..pdf. **101,** 4543–4547 (2004).

66. Tang, W. *et al.* White Fat Progenitors Reside in the Adipose Vasculature. *Science (80-.).* **322,** 583–586 (2009).

67. Petrovic, N. *et al.* Chronic peroxisome proliferator-activated receptor γ (PPARγ) activation of epididymally derived white adipocyte cultures reveals a population of thermogenically competent, UCP1-containing adipocytes molecularly distinct from classic brown adipocytes. *J. Biol. Chem.* **285,** 7153–7164 (2010).

68. Nedergaard, J. & Cannon, B. The Changed Metabolic World with Human Brown Adipose Tissue: Therapeutic Visions. *Cell Metab.* **11,** 268–272 (2010).

69. CANNON, B. Brown Adipose Tissue: Function and Physiological Significance. *Physiol. Rev.* **84,** 277–359 (2004).

70. Nedergaard, J., Petrovic, N., Lindgren, E. M., Jacobsson, A. & Cannon, B. PPARγ in the control of brown adipocyte differentiation. *Biochim. Biophys. Acta - Mol. Basis Dis.* **1740,** 293–304 (2005).

71. Ohno, H., Shinoda, K., Spiegelman, B. M. & Kajimura, S. PPAR agonists induce a white-to-brown fat conversion through stabilization of PRDM16 protein. *Cell Metab.* **15,** 395–404 (2012).

72. Qiang, L. *et al.* Brown Remodeling of White Adipose Tissue by SirT1-Dependent Deacetylation of Pparγ. *Cell* **150,** 620–632 (2012).

73. Armoni, M., Harel, C. & Karnieli, E. Transcriptional regulation of the GLUT4 gene: from PPAR-γ and FOXO1 to FFA and inflammation. *Trends Endocrinol. Metab.* **18,** 100–107 (2007).

74. Wu, Z., Xie, Y., Morrison, R. F., Bucher, N. L. R. & Farmer, S. R. PPARγ induces the insulin-dependent glucose transporter GLUT4 in the absence of C/EBPα during the conversion of 3T3 fibroblasts into adipocytes. *J. Clin. Invest.* **101,** 22–32 (1998).

75. Ribon, V., Johnson, J. H., Camp, H. S. & Saltiel, A. R. Thiazolidinediones and insulin resistance: peroxisome proliferatoractivated receptor gamma activation stimulates expression of the CAP gene. *Proc. Natl. Acad. Sci. U. S. A.* **95,** 14751–14756 (1998).

76. Maeda, N. *et al.* PPARγ Ligands Increase Expression and Plasma Concentrations of Adiponectin, an Adipose-Derived Protein. *Diabetes* **50,** 2094–2099 (2001).

77. Hollenberg, A. N. *et al.* Functional antagonism between CCAAT/enhancer binding protein-α and peroxisome proliferator-activated receptor-γ on the leptin promoter. *J. Biol. Chem.* **272,** 5283–5290 (1997).

78. Zhang, Y. *et al.* A noncanonical PPARγ/RXRα-binding sequence regulates leptin expression in response to changes in adipose tissue mass. *Proc. Natl. Acad. Sci.* **115,** E6039–E6047 (2018).

79. Hofmann, C. *et al.* Altered gene expression for tumor necrosis factor-alpha and its receptors during drug and dietary modulation of insulin resistance. *Endocrinology* **134,** 264–270 (1994).

80. Borst, S. E. The role of TNF-α in insulin resistance. *Endocrine* **23,** 177–182 (2004).

81. Kim, J. K. *et al.* Differential effects of rosiglitazone on skeletal muscle and liver insulin resistance in A-ZIP/F-1 fatless mice. *Diabetes* **52,** 1311–1318 (2003).

82. Berkowitz, K. *et al.* Effect of Troglitazone on Insulin Sensitivity and Pancreatic β-Cell Function in Women at High Risk for NIDDM. *Diabetes* **45,** 1572–1579 (1996).

83. Kim, H. *et al.* Peroxisomal Proliferator-Activated Receptor-γ Upregulates Glucokinase Gene Expression in β-Cells. *Diabetes* **51,** 676–685 (2002).

84. Way, J. M. *et al.* Comprehensive messenger ribonucleic acid profiling reveals that peroxisome proliferator-activated receptor γ activation has coordinate effects on gene expression in multiple insulin-sensitive tissues. *Endocrinology* **142,** 1269–1277 (2001).

85. Kallwitz, E. R., McLachlan, A. & Cotler, S. J. Role of peroxisome proliferators-activated receptors in the pathogenesis and treatment of nonalcoholic fatty liver disease. *World J. Gastroenterol.* **14,** 22–28 (2008).

86. Mayerson, A. B. *et al.* The effects of rosiglitazone on insulin sensitivity, lipolysis, and hepatic and skeletal muscle triglyceride content in patients with type 2 diabetes. *Diabetes* **51,** 797–802 (2002).

87. Bruedigam·, C., Drabek, K., van de Peppel, J. & van Leeuwen, H. Peroxisome proliferator-activated receptor gamma controls human osteoblast differentiation celerity and life span – A novel concept for aging-related thiazolidinedione-mediated bone loss. *Bone* **44,** S297 (2009).

88. Patel, J. J., Butters, O. R. & Arnett, T. R. PPAR agonists stimulate adipogenesis at the expense of osteoblast differentiation while inhibiting osteoclast formation and activity. *Cell Biochem. Funct.* **32,** 368–377 (2014).

89. Akune, T. *et al.* PPARγ insufficiency enhances osteogenesis through osteoblast formation from bone marrow progenitors. *J. Clin. Invest.* **113,** 846–855 (2004).

90. Wan, Y., Chong, L. W. & Evans, R. M. PPAR-γ regulates osteoclastogenesis in mice. *Nat. Med.* **13,** 1496–1503 (2007).

91. Liu, L. *et al.* Rosiglitazone Inhibits Bone Regeneration and Causes Significant Accumulation of Fat at Sites of New Bone Formation Lichu. *Calcif Tissue Int.* **91,** 139–148 (2012).

92. Lu, M. *et al.* Brain PPARγ Promotes Obesity and is Required for the Insulin–Sensitizing Effect of Thiazolidinediones. *Nat Med.* **17,** 618–622 (2011).

93. Karen K. Ryan, Bailing Li, Bernadette E. Grayson, Emily K. Matter, Stephen C. Woods, and R. J. S. A role for CNS PPAR γ in the regulation of energy balance. *Nat Med.* **17,** 623–626 (2011).

94. Sarafidis, P. A., Georgianos, P. I. & Lasaridis, A. N. PPAR-γ Agonism for Cardiovascular and Renal Protection. *Cardiovasc. Ther.* **29,** 377–384 (2011).

95. Kuba, K. *et al.* A crucial role of angiotensin converting enzyme 2 (ACE2) in SARS coronavirus-induced lung injury. *Nat. Med.* **11,** 875–879 (2005).

96. Guan, Y. F. *et al.* Thiazolidinediones expand body fluid volume through PPARγ stimulation of ENaC-mediated renal salt absorption. *Nat. Med.* **11,** 861–866 (2005).

97. Zhang, H. *et al.* Collecting duct-specific deletion of peroxisome proliferator-activated receptor gamma blocks thiazolidinedione-induced fluid retention. *Proc. Natl. Acad. Sci. U. S. A.* **102,** 9406–9411 (2005).

98. Son, N. *et al.* Cardiomyocyte expression of PPAR g leads to cardiac dysfunction in mice Find the latest version : Cardiomyocyte expression of PPAR γ leads to cardiac dysfunction in mice. **117,** 2791–2801 (2007).

99. Huang, J. V., Greyson, C. R. & Schwartz, G. G. PPAR-γ as a therapeutic target in cardiovascular disease: evidence and uncertainty. *J. Lipid Res.* **53,** 1738–1754 (2012).

100. Duan, S. Z., Ivashchenko, C. Y., Russell, M. W., Milstone, D. S. & Mortensen, R. M. Cardiomyocyte-specffic knockout and agonist of peroxisome proliferator-activated receptor-γ both induce cardiac hypertrophy in mice. *Circ. Res.* **97,** 372–379 (2005).

101. Ahmadian, M. *et al.* Pparγ signaling and metabolism: The good, the bad and the future. *Nat. Med.* **19,** 557–566 (2013).

102. Waite, K. J., Floyd, Z. E., Arbour-Reily, P. & Stephens, J. M. Interferon-γ-induced Regulation of Peroxisome Proliferator-activated Receptor γ and STATs in Adipocytes. *J. Biol. Chem.* **276,** 7062–7068 (2001).

103. Chang, A. J., Song, D. H. & Wolfe, M. M. Attenuation of peroxisome proliferator-activated receptor γ (PPARγ) mediates gastrin-stimulated colorectal cancer cell proliferation. *J. Biol. Chem.* **281,** 14700–14710 (2006).

104. Kliewer, S. A. *et al.* A prostaglandin J2 metabolite binds peroxisome proliferator-activated receptor γ and promotes adipocyte differentiation. *Cell* **83,** 813–819 (1995).

105. Forman, B. M. *et al.* 15-Deoxy-Δ12,14-Prostaglandin J2is a ligand for the adipocyte determination factor PPARγ. *Cell* **83,** 803–812 (1995).

106. Waku, T., Shiraki, T., Oyama, T. & Morikawa, K. Atomic structure of mutant PPARγ LBD complexed with 15d-PGJ2: Novel modulation mechanism of PPARγ/RXRα function by covalently bound ligands. *FEBS Lett.* **583,** 320–324 (2009).

107. Waku, T. *et al.* Structural Insight into PPARγ Activation Through Covalent Modification with Endogenous Fatty Acids. *J. Mol. Biol.* **385,** 188–199 (2009).

108. Itoh, T. *et al.* Structural basis for the activation of PPARgamma by oxidized fatty acids. *Nat. Struct. Mol. Biol.* **15,** 924–31 (2008).

109. Erlemann, K. R., Rokach, J. & Powell, W. S. Oxidative stress stimulates the synthesis of the eosinophil chemoattractant 5-oxo-6,8,11,14-eicosatetraenoic acid by inflammatory cells. *J. Biol. Chem.* **279,** 40376–40384 (2004).

110. O'Flaherty, J. T. *et al.* 5-Oxo-ETE analogs and the proliferation of cancer cells. *Biochim. Biophys. Acta - Mol. Cell Biol. Lipids* **1736,** 228–236 (2005).

111. Sheu, S. H., Kaya, T., Waxman, D. J. & Vajda, S. Exploring the binding site structure of the PPARγ ligand-binding domain by computational solvent mapping. *Biochemistry* **44,** 1193–1209 (2005).

112. Hughes, T. S. *et al.* An alternate binding site for PPARγ ligands. *Nat Commun* **359,** 1018–1026 (2015).

113. Jang, J. Y. *et al.* Structural basis for differential activities of enantiomeric PPARγ agonists: Binding of S35 to the alternate site. *Biochim. Biophys. Acta - Proteins Proteomics* **1865,** 674–681 (2017).

114. Pan, T. & Coleman, J. E. Structure and function of the Zn(II) binding site within the DNA-binding domain of the GAL4 transcription factor. *Proc. Natl. Acad. Sci. U. S. A.* **86,** 3145–9 (1989).

115. Shao, D. *et al.* Interdomain communication regulating ligand binding by PPAR-gamma. *Nature* **396,** 377–380 (1998).

116. Luckert, C., Hessel, S., Lampen, A. & Braeuning, A. Utility of an appropriate reporter assay: Heliotrine interferes with GAL4/upstream activation sequence-driven reporter gene systems. *Anal. Biochem.* **487,** 45–48 (2015).

117. Nishikawa, J. *et al.* New screening methods for chemicals with hormonal activities using interaction of nuclear hormone receptor with coactivator. *Toxicol. Appl. Pharmacol.* **154,** 76–83 (1999).

118. Luo, Y., Batalao, A., Zhou, H. & Zhu, L. Mammalian two-hybrid system: A complementary approach to the yeast two- hybrid system. *Biotechniques* **22,** 350–352 (1997).

119. Delerive, P. *et al.* Peroxisome Proliferator-activated Receptor α Negatively Regulates the Vascular Inflammatory Gene Response by Negative Cross-talk with Transcription Factors NF- κ Peroxisome Proliferator-activate. *J. Biol. Chem.* **274,** 32048–32054 (1999).

120. IJpenberg, A. *et al.* In vivo activation of PPAR target genes by RXR homodimers. *EMBO J.* **23,** 2083–2091 (2004).

121. Michel, M. C. & Charlton, S. J. Biased agonism in drug discovery - is it too soon to choose a path? *Mol. Pharmacol.* mol.117.110890 (2018).

doi:10.1124/mol.117.110890

122. Rankovic, Z., Brust, T. F. & Bohn, L. M. Biased agonism: An emerging paradigm in GPCR drug discovery. *Bioorganic Med. Chem. Lett.* **26,** 241–250 (2016).

123. Brust, R. *et al.* A structural mechanism for directing inverse agonism of PPARγ. *bioRxiv* (2018). doi:10.1101/245852

124. de Vera, I. M. S. *et al.* Synergistic Regulation of Coregulator/Nuclear Receptor Interaction by Ligand and DNA. *Structure* **25,** 1506–1518.e4 (2017).

125. Bramlett, K. S., Wu, Y. & Burris, T. P. Ligands Specify Coactivator Nuclear Receptor (NR) Box Affinity for Estrogen Receptor Subtypes. *Mol. Endocrinol.* **15,** 909–922 (2001).

126. Igor, G., Flores, A. M. & Aneskievich, B. J. Corepressors of Agonist-Bound Nuclear Receptors. *Toxicol Appl Pharmacol* **223,** 288–298 (2007).

127. Comley, J. TR-FRET based assays - Getting better with age. *Drug Discov. World* **7,** 22–37 (2006).

128. Taosheng, C., Wen, X., Michele, A. & Martin, B. Coactivators in Assay Design for Nuclear Hormone Receptor Drug Discovery. *Assay Drug Dev. Technol.* **1,** (2004).

129. Francois, D. *et al.* HTRF: A Technology Tailored for Drug Discovery –A Review of Theoretical Aspects and Recent Applications. *Curr. Chem. Genomics* **3,** 22–32 (2009).

130. Du, Y. in *Protein-Protein Interactions* 529–544 (Humana Press, 2015).

131. Teichert, A. *et al.* Quantification of the Vitamin D Receptor # Coregulator Interaction Quantification of the Vitamin D Receptor-Coregulator Interaction †. *Society* 1454–1461 (2009). doi:10.1021/bi801874n

132. Moore, J. M. R. *et al.* Quantitative proteomics of the thyroid hormone receptor-coregulator interactions. *J. Biol. Chem.* **279,** 27584–27590 (2004).

133. Zhang, C., Nordeen, S. K. & Shapiro, D. J. Fluorescence anisotropy microplate assay to investigate the interaction of full-length steroid receptor coactivator-1a with steroid receptors. *Methods Mol Biol* **977,** 339–351 (2013).

134. Houtman, R. *et al.* Serine-305 Phosphorylation Modulates Estrogen Receptor Alpha Binding to a Coregulator Peptide Array, with Potential Application in Predicting Responses to Tamoxifen. *Mol. Cancer Ther.* **11,** 805–816 (2012).

135. Broekema, M. F. *et al.* Profiling Of 3696 Nuclear Receptor-Coregulator Interactions: A Resource For Biological And Clinical Discovery. *Endocrinology* **9,** 397 (2018).

136. Corso-Díaz, X. *et al.* Co-activator candidate interactions for orphan nuclear receptor NR2E1. *BMC Genomics* **17,** 1–11 (2016).

137. Rosenfeld, M. G., Lunyak, V. V & Glass, C. K. integrating signal-dependent programs of transcriptional response Sensors and signals : a coactivator /

corepressor / epigenetic code for integrating signal-dependent programs of transcriptional response. *Genes Dev.* **20,** 1405–1428 (2006).

138.   Nagy, L. & Schwabe, J. W. R. Mechanism of the nuclear receptor molecular switch. *Trends Biochem. Sci.* **29,** 317–324 (2004).

139.   Tontonoz, P. & Spiegelman, B. M. Fat and Beyond: The Diverse Biology of PPARγ. *Annu. Rev. Biochem.* **77,** 289–312 (2008).

140.   Johnson, B. A. *et al.* Ligand-induced stabilization of PPARγ monitored by NMR spectroscopy: implications for nuclear receptor activation. *J. Mol. Biol.* **298,** 187–194 (2000).

141.   Day, C. Thiazolidinediones: A new class of antidiabetic drugs. *Diabet. Med.* **16,** 179–192 (1999).

142.   Brown, K. K. *et al.* A novel N-aryl tyrosine activator of peroxisome proliferator-activated receptor-γ reverses the diabetic phenotype of the Zucker diabetic fatty rat. *Diabetes* **48,** 1415–1424 (1999).

143.   Patel, H. *et al.* Preclinical evaluation of saroglitazar magnesium, a dual PPAR-α/γ agonist for treatment of dyslipidemia and metabolic disorders. *Xenobiotica* **0,** 1–10 (2017).

144.   Jaeschke, H. Troglitazone hepatotoxicity: Are we getting closer to understanding idiosyncratic liver injury? *Toxicol. Sci.* **97,** 1–3 (2007).

145.   Ikeda, T. Drug-induced Idiosyncratic Hepatotoxicity: Prevention Strategy Developed after the Troglitazone Case. *Drug Metab. Pharmacokinet.* **26,** 60–70 (2011).

146.   Bach, R. G. *et al.* Rosiglitazone and outcomes for patients with diabetes mellitus and coronary artery disease in the bypass angioplasty revascularization investigation 2 diabetes (BARI 2D) Trial. *Circulation* **128,** 785–794 (2013).

147.   Virk, I. S., Ruby, R. S. & Tepper, D. Rosiglitazone and heart failure: the controversy and clinical implications. *Congest. Heart Fail.* **13,** 293–295 (2007).

148.   Komajda, M. *et al.* Heart failure events with rosiglitazone in type 2 diabetes: Data from the RECORD clinical trial. *Eur. Heart J.* **31,** 824–831 (2010).

149.   Ryder, R. E. J. Pioglitazone has a dubious bladder cancer risk but an undoubted cardiovascular benefit. *Diabet. Med.* **32,** 305–313 (2015).

150.   Korhonen, P. *et al.* Pioglitazone use and risk of bladder cancer in patients with type 2 diabetes: retrospective cohort study using datasets from four European countries. *BMJ* **354,** i3903 (2016).

151.   An, Z., Muthusami, S., Yu, J.-R. & Park, W.-Y. T0070907, a PPAR γ inhibitor induced G2/M Arrest Enhances the Effect of Radiation in Human Cervical Cancer Cells Through Mitotic Catastrophe. *Reprod. Sci.* 1–10 (2014). doi:10.1177/1933719114525265

152.   Zaytseva, Y. Y., Wallis, N. K., Southard, R. C. & Kilgore, M. W. The PPARgamma

antagonist T0070907 suppresses breast cancer cell proliferation and motility via both PPARgamma-dependent and -independent mechanisms. *Anticancer Res.* **31,** 813–823 (2011).

153.　Wang, X., Sun, Y., Wong, J. & Conklin, D. S. PPARγ maintains ERBB2-positive breast cancer stem cells. *Oncogene* **32,** 5512–5521 (2013).

154.　Seargent, J. M., Yates, E. A. & Gill, J. H. GW9662, a potent antagonist of PPAR??, inhibits growth of breast tumour cells and promotes the anticancer effects of the PPAR?? agonist rosiglitazone, independently of PPAR?? activation. *Br. J. Pharmacol.* **143,** 933–937 (2004).

155.　Leesnitzer, L. M. *et al.* Functional consequences of cysteine modification in the ligand binding sites of peroxisome proliferator activated receptors by GW9662. *Biochemistry* **41,** 6640–6650 (2002).

156.　Lee, G. *et al.* T0070907, a selective ligand for peroxisome proliferator-activated receptor γ, functions as an antagonist of biochemical and cellular activities. *J. Biol. Chem.* **277,** 19649–19657 (2002).

157.　Fehlberg, S., Trautwein, S., Göke, A. & Göke, R. Bisphenol A diglycidyl ether induces apoptosis in tumour cells independently of peroxisome proliferator-activated receptor-gamma, in caspase-dependent and -independent manners. *Biochem. J.* **362,** 573–8 (2002).

158.　Marciano, D. P. *et al.* Pharmacological Repression of PPARγ Promotes Osteogenesis. *Nat Commun* **6,** (2015).

159.　Xu, H. E. *et al.* Structural basis for antagonist mediated recruitment of nuclear co-repressors by PPARalpha. *Nature* **415,** 813–817 (2002).

160.　Rangwala, S M; Lazar, M. A. The dawn of the SPPARMs? *Sci STKE* **121,** (2002).

161.　Marciano, D. P. D. P. *et al.* Pharmacological repression of PPARγ promotes osteogenesis. *Nat. Commun.* **6,** 7443 (2015).

162.　Berger, J. P. *et al.* Distinct Properties and Advantages of a Novel Peroxisome Proliferator-Activated Protein γ Selective Modulator. *Mol. Endocrinol.* **17,** 662–676 (2003).

163.　Motani, A. *et al.* INT131: A Selective Modulator of PPARγ. *J. Mol. Biol.* **386,** 1301–1311 (2009).

164.　Bruning, J. B. *et al.* Partial Agonists Activate PPAR?? Using a Helix 12 Independent Mechanism. *Structure* **15,** 1258–1271 (2007).

165.　Kroker, A. J. & Bruning, J. B. Review of the Structural and Dynamic Mechanisms of PPARgamma Partial Agonism. *PPAR Res.* **2015,** 816856 (2015).

166.　Choi, J. H. *et al.* Anti-Diabetic Actions of a Non-Agonist PPARγ Ligand Blocking Cdk5-Mediated Phosphorylation. *Nature* **477,** 477–481 (2012).

167.　Acton, J. J. *et al.* Benzoyl 2-methyl indoles as selective PPARγ modulators. *Bioorganic Med. Chem. Lett.* **15,** 357–362 (2005).

168. Rangwala, S. M. *et al.* Genetic modulation of PPARγ phosphorylation regulates insulin sensitivity. *Dev. Cell* **5,** 657–663 (2003).

169. Burns, K. A. & Vanden Heuvel, J. P. Modulation of PPAR activity via phosphorylation. *Biochim Biophys Acta* **1771,** 952–960 (2007).

170. Ohshima, T., Koga, H. & Shimotahno, K. Transcriptional activity of peroxisome proliferator-activated receptor γ is modulated by SUMO-1 modification. *J. Biol. Chem.* **279,** 29551–29557 (2004).

171. Katafuchi, T. *et al.* PPAR{gamma}-K107 SUMOylation regulates insulin sensitivity but not adiposity in mice [Physiology]. *Proc. Natl. Acad. Sci. U. S. A.* (2018). doi:10.1073/pnas.1814522115

172. Pascual, G. *et al.* A sumoylation-dependent pathway mediating transrepression of inflammatory response genes by PPARγ. *Nature* **437,** 759–763 (2005).

173. Diezko, R. & Suske, G. Ligand Binding Reduces SUMOylation of the Peroxisome Proliferator-activated Receptor γ (PPARγ) Activation Function 1 (AF1) Domain. *PLoS One* **8,** (2013).

174. Choi, J. H. *et al.* Obesity-linked phosphorylation of PPARγ by cdk5 is a direct target of the anti-diabetic PPARγ ligands. *Nature* **466,** 451–456 (2010).

175. Bórquez, D. A. *et al.* Bioinformatic survey for new physiological substrates of Cyclin-dependent kinase 5. *Genomics* **101,** 221–228 (2013).

176. Mottin, M., Souza, P. C. T. & Skaf, M. S. Molecular Recognition of PPAR?? by Kinase Cdk5/p25: Insights from a Combination of Protein-Protein Docking and Adaptive Biasing Force Simulations. *J. Phys. Chem. B* **119,** 8330–8339 (2015).

177. Choi, J. H. *et al.* Thrap 3 docks on phosphoserine 273 of PPAR γ and controls diabetic gene programming. *Genes Dev.* **28,** 2361–2369 (2014).

178. Choi, S. S. *et al.* A novel non-agonist peroxisome proliferator-activated receptor γ (PPARγ) ligand UHC1 blocks PPARγ phosphorylation by cyclin-dependent kinase 5 (CDK5) and improves insulin sensitivity. *J. Biol. Chem.* **289,** 26618–26629 (2014).

179. Eaton, S. L. *et al.* A Guide to Modern Quantitative Fluorescent Western Blotting with Troubleshooting Strategies. *J. Vis. Exp.* 1–10 (2014). doi:10.3791/52099

180. Filho, H. V. R. *et al.* Screening for PPAR non-agonist ligands followed by characterization of a hit, AM-879, with additional no-adipogenic and cdk5-mediated phosphorylation inhibition properties. *Front. Endocrinol. (Lausanne).* **9,** (2018).

181. Prokoph, N. *et al.* Development of an ELISA for High-Throughput Screening of Inhibitors of Cdk5-Mediated PPAR? Phosphorylation. *Assay Drug Dev. Technol.* **14,** 261–272 (2016).

182. Li, X., Xu, B., Wang, Y. & Wei, L. Anti-inflammatory effect of peroxisome proliferator-activated receptor-γ (PPAR-γ) on non-obese diabetic mice with Sjogren's syndrome. *Int. J. Clin. Exp. Pathol.* **7,** 4886–4894 (2014).

183. Yang, S. *et al.* Pioglitazone ameliorates Aβ42 deposition in rats with diet-induced insulin resistance associated with AKT/GSK3β activation. *Mol. Med. Rep.* **15,** 2588–2594 (2017).

184. Kapadia, R. Mechanisms of anti-inflammatory and neuroprotective actions of PPAR-gamma agonists. *Front. Biosci.* **13,** 1813 (2008).

185. Chung, S. W. *et al.* Oxidized Low Density Lipoprotein Inhibits Interleukin-12 Production in Lipopolysaccharide-activated Mouse Macrophages via Direct Interactions between Peroxisome Proliferator-activated Receptor-γ and Nuclear Factor-κB. *J. Biol. Chem.* **275,** 32681–32687 (2000).

186. Annese, V., Rogai, F., Settesoldi, A. & Bagnoli, S. PPAR γ in inflammatory bowel disease. *PPAR Res.* **2012,** (2012).

187. Dammann, K. *et al.* PAK1 modulates a PPARγ/NF-κB cascade in intestinal inflammation. *Biochim. Biophys. Acta - Mol. Cell Res.* **1853,** 2349–2360 (2015).

188. Scirpo, R. *et al.* Stimulation of nuclear receptor PPAR- γ limits NF-kB-dependent inflammation in mouse cystic fibrosis biliary epithelium. *Hepatology* **62,** 1551–1562 (2015).

189. Karmakar, S., Jin, Y. & Nagaich, A. K. Interaction of glucocorticoid receptor (GR) with estrogen receptor (ER) ?? and activator protein 1 (AP1) in dexamethasone-mediated interference of ER?? activity. *J. Biol. Chem.* **288,** 24020–24034 (2013).

190. Ohshima, K., Mogi, M. & Horiuchi, M. Role of Peroxisome Proliferator-Activated Receptor-gamma in Vascular Inflammation. *Int J Vasc Med* **2012,** 508416 (2012).

191. Carta, A. R. PPAR-γ: Therapeutic Prospects in Parkinson's Disease. *Curr. Drug Targets* **14,** 743–751 (2013).

192. Breidert, T. *et al.* Protective action of the peroxisome proliferator-activated receptor-γ agonist pioglitazone in a mouse model of Parkinson's disease. *J. Neurochem.* **82,** 615–624 (2002).

193. Nicolakakis, N. *et al.* Complete Rescue of Cerebrovascular Function in Aged Alzheimer's Disease Transgenic Mice by Antioxidants and Pioglitazone, a Peroxisome Proliferator-Activated Receptor Agonist. *J. Neurosci.* **28,** 9287–9296 (2008).

194. Steinmetz, A. C., Renaud, J. P. & Moras, D. Binding of ligands and activation of transcription by nuclear receptors. *Annu. Rev. Biophys. Biomol. Struct.* **30,** 329–359 (2001).

195. Overington, J. P., Al-Lazikani, B. & Hopkins, A. L. How many drug targets are there? *Nat. Rev. Drug Discov.* **5,** 993–996 (2006).

196. Nettles, K. W. & Greene, G. L. Ligand Control of Coregulator Recruitment To Nuclear Receptors. *Annu. Rev. Physiol.* **67,** 309–333 (2005).

197. Brzozowski, a M. *et al.* Molecular basis of agonism and antagonism in the oestrogen receptor. *Nature* **389,** 753–758 (1997).

198.   Shiau, A. K. *et al.* The structural basis of estrogen receptor/coactivator recognition and the antagonism of this interaction by tamoxifen. *Cell* **95,** 927–937 (1998).

199.   Bourguet, W., Ruff, M., Chambon, P., Gronemeyer, H. & Moras, D. Crystal structure of the ligand-binding domain of the human nuclear receptor RXR-alpha. *Nature* **375,** 377–382 (1995).

200.   Xu, H. E. *et al.* Structural basis for antagonist-mediated recruitment of nuclear co-repressors by PPARα. *Nature* **415,** 813–817 (2002).

201.   Bain, D. L., Heneghan, A. F., Connaghan-Jones, K. D. & Miura, M. T. Nuclear Receptor Structure: Implications for Function. *Annu. Rev. Physiol* **69,** 201–20 (2007).

202.   Kroker, A. J. & Bruning, J. B. Review of the Structural and Dynamic Mechanisms of PPARγ Partial Agonism. *PPAR Res.* **2015,** 816856 (2015).

203.   Kojetin, D. J. & Burris, T. P. Small molecule modulation of nuclear receptor conformational dynamics: implications for function and drug discovery. *Mol. Pharmacol.* **83,** 1–8 (2013).

204.   Choi, J. H. *et al.* Anti-diabetic drugs inhibit obesity-linked phosphorylation of PPARgamma by Cdk5. *Nature* **466,** 451–456 (2010).

205.   Wright, E. *et al.* Helix 11 dynamics is critical for constitutive androstane receptor activity. *Structure* **19,** 37–44 (2011).

206.   Zhang, J. *et al.* DNA binding alters coactivator interaction surfaces of the intact VDR-RXR complex. *Nat. Struct. Mol. Biol.* **18,** 556–63 (2011).

207.   Hughes, T. S. *et al.* Ligand and receptor dynamics contribute to the mechanism of graded PPARγ agonism. *Structure* **20,** 139–150 (2012).

208.   Hughes, T. S. *et al.* An alternate binding site for PPARγ ligands. *Nat. Commun.* **5,** 3571 (2014).

209.   Berger, J. P. *et al.* Distinct properties and advantages of a novel peroxisome proliferator-activated protein [gamma] selective modulator. *Mol Endocrinol* **17,** 662–676 (2003).

210.   Sharaf, N. G. & Gronenborn, A. M. in *Methods in Enzymology* **565,** 67–95 (2015).

211.   Kim, T. H. *et al.* The role of ligands on the equilibria between functional states of a G protein-coupled receptor. *J. Am. Chem. Soc.* **135,** 9465–9474 (2013).

212.   Nygaard, R. *et al.* The dynamic process of ??2-adrenergic receptor activation. *Cell* **152,** 532–542 (2013).

213.   Manglik, A. *et al.* Structural Insights into the Dynamic Process of β2-Adrenergic Receptor Signaling. *Cell* **161,** 1101–11 (2015).

214.   Liu, J. J., Horst, R., Katritch, V., Stevens, R. C. & Wüthrich, K. Biased signaling pathways in β2-adrenergic receptor characterized by 19F-NMR. *Science* **335,**

1106–10 (2012).

215. Kitevski-LeBlanc, J. L. & Prosser, R. S. Current applications of 19F NMR to studies of protein structure and dynamics. *Prog. Nucl. Magn. Reson. Spectrosc.* **62,** 1–33 (2012).

216. Lehmann, J. M. *et al.* An antidiabetic thiazolidinedione is a high affinity ligand for peroxisome proliferator-activated receptor gamma (PPAR gamma). *J. Biol. Chem.* **270,** 12953–12956 (1995).

217. Brust, R. *et al.* Modification of the Orthosteric PPARγ Covalent Antagonist Scaffold Yields an Improved Dual-Site Allosteric Inhibitor. *ACS Chem. Biol.* **12,** 969–978 (2017).

218. Liberato, M. V. *et al.* Medium chain fatty acids are selective peroxisome proliferator activated receptor (PPAR) gamma activators and Pan-PPAR partial agonists. *PLoS One* **7,** e36297 (2012).

219. Hughes, T. S., Wilson, H. D., De Vera, I. M. S. & Kojetin, D. J. Deconvolution of complex 1D NMR spectra using objective model selection. *PLoS One* **10,** e0134474 (2015).

220. Evanics, F. *et al.* Tryptophan solvent exposure in folded and unfolded states of an SH3 domain by 19F and 1H NMR. *Biochemistry* **45,** 14120–14128 (2006).

221. Bruning, J. B. *et al.* Partial Agonists Activate PPARγ Using a Helix 12 Independent Mechanism. *Structure* **15,** 1258–1271 (2007).

222. Evanics, F., Kitevski, J. L., Bezsonova, I., Forman-Kay, J. & Prosser, R. S. 19F NMR studies of solvent exposure and peptide binding to an SH3 domain. *Biochim. Biophys. Acta - Gen. Subj.* **1770,** 221–230 (2007).

223. Henzler-Wildman, K. & Kern, D. Dynamic personalities of proteins. *Nature* **450,** 964–972 (2007).

224. Hwan Bae, Jun Young Jang, Sun-Sil Choi, Jae-Jin Lee, Heejun Kim, Ala Jo, Kong-Joo Lee, Jang Hyun Choi, S. W. S. and S. B. P. Mechanistic elucidation guided by covalent inhibitors for the development of anti-diabetic PPARγ ligands. *Chem. Sci.* **7,** 5523–5529 (2016).

225. Nielsen, R. *et al.* Genome-wide profiling of PPARγ:RXR and RNA polymerase II occupancy reveals temporal activation of distinct metabolic pathways and changes in RXR dimer composition during adipogenesis. *Genes Dev.* **22,** 2953–2967 (2008).

226. Song, Y. & Dai, L. L. The shear viscosities of common water models by non-equilibrium molecular dynamics simulations. *Mol. Simul.* **36,** 560–567 (2010).

227. Batista, M. R. B. & Martínez, L. Conformational Diversity of the Helix 12 of the Ligand Binding Domain of PPARγ and Functional Implications. *J. Phys. Chem. B* **119,** 15418–15429 (2015).

228. Ye, L., Van Eps, N., Zimmer, M., Ernst, O. P. & Scott Prosser, R. Activation of the

A2A adenosine G-protein-coupled receptor by conformational selection. *Nature* **533,** 1–15 (2016).

229. Staus, D. P. *et al.* Allosteric nanobodies reveal the dynamic range and diverse mechanisms of G-protein-coupled receptor activation. *Nature* **535,** 448–452 (2016).

230. Kojetin, D. J. *et al.* Structural mechanism for signal transduction in RXR nuclear receptor heterodimers. *Nat Commun* **6,** 8013 (2015).

231. Glasoe, P. K. & Long, F. A. USE OF GLASS ELECTRODES TO MEASURE ACIDITIES IN DEUTERIUM OXIDE1,2. *J. Phys. Chem.* **64,** 188–190 (1960).

232. Motulsky, H. J. & Brown, R. E. Detecting outliers when fitting data with nonlinear regression – a new method based on robust nonlinear regression and the false discovery rate. *BMC Bioinformatics* **7,** 123 (2006).

233. Owen, S. C. *et al.* Colloidal drug formulations can explain 'bell-shaped' concentration-response curves. *ACS Chem. Biol.* **9,** 777–784 (2014).

234. Helgstrand, M., Härd, T. & Allard, P. Simulation of NMR pulse sequences during equilibrium and non-equilibrium chemical exchange. *J. Biomol. NMR* **18,** 49–63 (2000).

235. Fernandez-Fuentes, N. & Fiser, A. Modeling Loops in Protein Structures. *Introd. to Protein Struct. Predict.* 279–298 (2010). doi:10.1002/9780470882207.ch13

236. Pettersen, E. F. *et al.* UCSF Chimera—A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25,** 1605–1612 (2004).

237. Anandakrishnan, R., Aguilar, B. & Onufriev, A. V. H++ 3.0: Automating pK prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulations. *Nucleic Acids Res.* **40,** 537–541 (2012).

238. D.A. Case, V. Babin, J.T. Berryman, R.M. Betz, Q. Cai, D.S. Cerutti, T.E. Cheatham, III, T.A. Darden, R. E., Duke, H. Gohlke, A.W. Goetz, S. Gusarov, N. Homeyer, P. Janowski, J. Kaus, I. Kolossváry, A. K., T.S. Lee, S. LeGrand, T. Luchko, R. Luo, B. Madej, K.M. Merz, F. Paesani, D.R. Roe, A. Roitberg, C. S., R. Salomon-Ferrer, G. Seabra, C.L. Simmerling, W. Smith, J. Swails, R.C. Walker, J. Wang, R.M. Wolf, X. & Kollman, W. and P. A. AMBER 14. (2014).

239. Vanquelef, E. *et al.* R.E.D. Server: A web service for deriving RESP and ESP charges and building force field libraries for new molecules and molecular fragments. *Nucleic Acids Res.* **39,** W511-7 (2011).

240. Cornell, W. D., Cieplak, P., Bayly, C. I. & Kollman, P. A. Application of RESP Charges To Calculate Conformational Energies, Hydrogen Bond Energies, and Free Energies of Solvation. *J. Am. Chem. Soc.* **115,** 9620–9631 (1993).

241. Wang, J. M., Wolf, R. M., Caldwell, J. W., Kollman, P. a & Case, D. a. Development and testing of a general amber force field. *J. Comput. Chem.* **25,** 1157–1174 (2004).

242. Maier, J. A. *et al.* ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* **11,** 3696–3713 (2015).

243. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79,** 926–935 (1983).

244. Joung, I. S. & Cheatham, T. E. Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations. *J. Phys. Chem. B* **112,** 9020–9041 (2008).

245. Roe, D. R. & Cheatham, T. E. PTRAJ and CPPTRAJ: Software for processing and analysis of molecular dynamics trajectory data. *J. Chem. Theory Comput.* **9,** 3084–3095 (2013).

246. Trott, O. & Olson, A. AutoDock Vina: inproving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. *J. Comput. Chem.* **31,** 455–461 (2010).

247. Ryckaert, J. P., Ciccotti, G. & Berendsen, H. J. C. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* **23,** 327–341 (1977).

248. Govindan, M. V, Devic, M., Green, S., Gronemeyer, H. & Chambon, P. Cloning of the human glucocorticoid receptor cDNA. *Nucleic Acids Res.* **13,** (1985).

249. Chan, J. S. K. *et al.* Targeting nuclear receptors in cancer-associated fibroblasts as concurrent therapy to inhibit development of chemoresistant tumors. *Oncogene* **37,** 160–173 (2018).

250. Dhiman, V. K., Bolt, M. J. & White, K. P. Nuclear receptors in cancer - Uncovering new and evolving roles through genomic analysis. *Nat. Rev. Genet.* **19,** 160–174 (2018).

251. Jin, H. S., Kim, T. S. & Jo, E. K. Emerging roles of orphan nuclear receptors in regulation of innate immunity. *Arch. Pharm. Res.* **39,** 1491–1502 (2016).

252. Sonoda, J., Pei, L. & Evans, R. M. Nuclear receptors: Decoding metabolic disease. *FEBS Lett.* **582,** 2–9 (2008).

253. Hong, C. & Tontonoz, P. Coordination of inflammation and metabolism by PPAR and LXR nuclear receptors. *Curr. Opin. Genet. Dev.* **18,** 461–467 (2008).

254. Bernal, J. Thyroid hormone receptors in brain development and function. *Nat. Clin. Pract. Endocrinol. &Amp; Metab.* **3,** 249 (2007).

255. Zanaria, E. *et al.* An unusual member of the nuclear hormone receptor superfamily responsible for X-linked adrenal hypoplasia congenita. *Nature* **372,** 635 (1994).

256. Haider, N. B. *et al.* Mutation of a nuclear receptor gene, NR2E3, causes enhanced S cone syndrHaider, N. B., Jacobson, S. G., Cideciyan, a V, Swiderski, R., Streb, L.

M., Searby, C., … Sheffield, V. C. (2000). Mutation of a nuclear receptor gene, NR2E3, causes enhanced S cone . *Nat. Genet.* **24,** 127–131 (2000).

257. Burris, T. P. & McCabe, E. R. B. *Nuclear Receptors and Genetic Disease*. (2000).

258. Achermann, J. C., Schwabe, J., Fairall, L. & Chatterjee, K. Genetic disorders of nuclear receptors. *J. Clin. Invest.* **127,** 1181–1192 (2017).

259. Gurnell, M., Krishna, V. & Chatterjee, K. Nuclear receptors in disease: thyroid receptor beta, peroxisome-proliferator-activated receptor gamma and orphan receptors. *Essays Biochem.* **40,** 169–189 (2004).

260. Petrosino, M. *et al.* Single-nucleotide polymorphism of PPARγ, a protein at the crossroads of physiological and pathological processes. *Int. J. Mol. Sci.* **18,** 1–20 (2017).

261. Overington, J. P., Al-lazikani, B. & Hopkins, A. L. How many drug targets are there? *Nat. Rev. Drug Discov.* **5,** 993–996 (2006).

262. Santos, R. *et al.* A comprehensive map of molecular drug targets. *Nat. Rev. Drug Discov.* **16,** 19–34 (2016).

263. Lonard, D. M. & O'Malley, B. W. Nuclear receptor coregulators: modulators of pathology and therapeutic targets. *Nat. Rev. Endocrinol.* **8,** 598 (2012).

264. Dasgupta, S., Lonard, D. M. & O'Malley, B. W. Nuclear Receptor Coactivators: Master Regulators of Human Health and Disease. *Annu. Rev. Med.* **65,** 279–292 (2014).

265. Plevin, M. J., Mills, M. M. & Ikura, M. The LxxLL motif: a multifunctional binding sequence in transcriptional regulation. *Trends Biochem. Sci.* **30,** 66–69 (2005).

266. Chang, C. *et al.* Dissection of the LXXLL Nuclear Receptor-Coactivator Interaction Motif Using Combinatorial Peptide Libraries: Discovery of Peptide Antagonists of Estrogen Receptors α and β. *Mol. Cell. Biol.* **19,** 8226–8239 (1999).

267. Webb, P. *et al.* The nuclear receptor corepressor (N-CoR) contains three isoleucine motifs (I/LXXII) that serve as receptor interaction domains (IDs). *Mol. Endocrinol.* **14,** 1976–1985 (2000).

268. Nagy, L. *et al.* Mechanism of corepressor binding and release from nuclear hormone receptors. *Genes Dev.* **13,** 3209–3216 (1999).

269. Bunce, C. M. & Campbell, M. J. *Nuclear Receptors: Current Concepts and Future Challenges*. (Springer Science and Business Media, 2010).

270. Leers, J., Treuter, E. & Gustafsson, J. A. Mechanistic principles in NR box-dependent interaction between nuclear hormone receptors and the coactivator TIF2. *Mol. Cell. Biol.* **18,** 6001–13 (1998).

271. Wansa, K. D. S. A., Harris, J. M., Yan, G., Ordentlich, P. & Muscat, G. E. O. The AF-1 Domain of the Orphan Nuclear Receptor NOR-1 Mediates Trans-activation, Coactivator Recruitment, and Activation by the Purine Anti-metabolite 6-Mercaptopurine. *J. Biol. Chem.* **278,** 24776–24790 (2003).

272. Weatherman, R. V, Fletterick, R. J. & Scanlan, T. S. NUCLEAR-RECEPTOR LIGANDS AND LIGAND-BINDING DOMAINS Ross. *Annu. Rev. Biochem.* **68,** 559–581 (1999).

273. Moras, D. & Gronemeyer, H. The nuclear receptor ligand-binding domain: Structure and function. *Curr. Opin. Cell Biol.* **10,** 384–391 (1998).

274. Wang, L. *et al.* X-ray crystal structures of the estrogen-related receptor-γ ligand binding domain in three functional states reveal the molecular basis of small molecule regulation. *J. Biol. Chem.* **281,** 37773–37781 (2006).

275. Wu, Y., Chin, W. W., Wang, Y. & Burris, T. P. Ligand and coactivator identity determines the requirement of the charge clamp for coactivation of the peroxisome proliferator-activated receptor ?? *J. Biol. Chem.* **278,** 8637–8644 (2003).

276. Crooks, G., Hon, G., Chandonia, J. & Brenner, S. WebLogo: a sequence logo generator. *Genome Res* **14,** 1188–1190 (2004).

277. Holt, J. A. *et al.* Helix 1/8 interactions influence the activity of nuclear receptor ligand-binding domains. *J. Control. Release* **17,** 1704–1714 (2003).

278. Agostini, M. *et al.* Tyrosine agonists reverse the molecular defects associated with dominant-negative mutations in human peroxisome proliferator-activated receptor γ. *Endocrinology* **145,** 1527–1538 (2004).

279. Salentin, S., Schreiber, S., Haupt, V. J., Adasme, M. F. & Schroeder, M. PLIP: Fully automated protein-ligand interaction profiler. *Nucleic Acids Res.* **43,** W443–W447 (2015).

280. Gelin, M. *et al.* Combining 'dry' co-crystallization and in situ diffraction to facilitate ligand screening by X-ray crystallography. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **71,** 1777–1787 (2015).

281. Haffner, C. D. *et al.* Structure-Based Design of Potent Retinoid X Receptor α Agonists. *J. Med. Chem.* **47,** 2010–2029 (2004).

282. Chrisman, I. M. *et al.* Defining a conformational ensemble that directs activation of PPARγ. *Nat. Commun.* **9,** 1–16 (2018).

283. Torchia, J. *et al.* The transcriptional co-activator p/CIP binds CBP and mediates nuclear- receptor function. *Nature* **387,** 677–684 (1997).

284. Otwinoski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276,** 307–326 (1997).

285. Adams, P. D. *et al.* PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66,** 213–221 (2010).

286. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66,** 486–501 (2010).

287. Read, R. J. Improved Fourier Coefficients for Maps Using Phases from Partial Structures with Errors By. *Acta Cryst.* **42,** 140–149 (1986).

288. Kawai, M. & Rosen, C. J. PPARγ: a circadian transcription factor in adipogenesis and osteogenesis. *Nat. Rev. Endocrinol.* **6,** 629–36 (2010).

289. Bookout, A. L. *et al.* Anatomical profiling of nuclear receptor expression reveals a hierarchical transcriptional network. *Cell* **126,** 789–799 (2006).

290. Holzer, G., Markov, G. V. & Laudet, V. *Evolution of Nuclear Receptors and Ligand Signaling: Toward a Soft Key–Lock Model? Current Topics in Developmental Biology* **125,** (Elsevier Inc., 2017).

291. Millard, C. J., Watson, P. J., Fairall, L. & Schwabe, J. W. R. An evolving understanding of nuclear receptor coregulator proteins. *J. Mol. Endocrinol.* **51,** (2013).

292. Burris, T. P. *et al.* Nuclear Receptors and Their Selective Pharmacologic Modulators. *Pharmacol. Rev. Pharmacol Rev* **65,** 710–778 (2013).

293. Huang, P., Chandra, V. & Rastinejad, F. Structural overview of the nuclear receptor superfamily: insights into physiology and therapeutics. *Annu. Rev. Physiol.* **72,** 247–272 (2010).

294. Kaupang, Å., Laitinen, T., Poso, A. & Hansen, T. V. Structural review of PPARγ in complex with ligands: Cartesian- and dihedral angle principal component analyses of X-ray crystallographic data. *Proteins Struct. Funct. Bioinforma.* **85,** 1684–1698 (2017).

295. Trump, R. P. *et al.* Co-crystal structure guided array synthesis of PPARgamma inverse agonists. *Bioorg. Med. Chem. Lett.* **17,** 3916–20 (2007).

296. Brust, R. *et al.* A structural mechanism for directing corepressor-selective inverse agonism of PPARγ. *Nat. Commun.* **9,** 4687 (2018).

297. Zhi, X. *et al.* Structural basis for corepressor assembly by the orphan nuclear receptor TLX. *Genes Dev.* **29,** 440–450 (2015).

298. Kim, J. Y., Son, Y. L., Kim, J. S. & Lee, Y. C. Molecular Determinants Required for Selective Interactions between the Thyroid Hormone Receptor Homodimer and the Nuclear Receptor Corepressor N-CoR. *J. Mol. Biol.* **396,** 747–760 (2010).

299. Madauss, K. P. *et al.* A Structural and in Vitro Characterization of Asoprisnil: A Selective Progesterone Receptor Modulator. *Mol. Endocrinol.* **21,** 1066–1081 (2007).

300. Zhang, H., Chen, L., Chen, J., Jiang, H. & Shen, X. Structural basis for retinoic X receptor repression on the tetramer. *J. Biol. Chem.* **286,** 24593–24598 (2011).

301. Schoch, G. A. *et al.* Molecular Switch in the Glucocorticoid Receptor: Active and Passive Antagonist Conformations. *J. Mol. Biol.* **395,** 568–577 (2010).

302. Noguchi, M. *et al.* Ternary complex of human RORγ ligand-binding domain, inverse agonist and SMRT peptide shows a unique mechanism of corepressor recruitment. *Genes to Cells* **22,** 535–551 (2017).

303. Lori, C. *et al.* Structural basis of the transactivation deficiency of the human

PPARγ F360L mutant associated with familial partial lipodystrophy. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **70,** 1965–1976 (2014).

304. Batista, M. R. B., Mart??nez, L. & Martínez, L. Conformational Diversity of the Helix 12 of the Ligand Binding Domain of PPARγ and Functional Implications. *J. Phys. Chem. B* **119,** 15418–15429 (2015).

305. Pierce, L. C. T., Salomon-Ferrer, R., Augusto F. de Oliveira, C., McCammon, J. A. & Walker, R. C. Routine Access to Millisecond Time Scale Events with Accelerated Molecular Dynamics. *J. Chem. Theory Comput.* **8,** 2997–3002 (2012).

306. Choi, J. H. *et al.* Anti-diabetic drugs inhibit obesity-linked phosphorylation of PPARγ by Cdk5. *Nature* **466,** 451–456 (2010).

307. Pike, A. C. W. *et al.* Structure of the ligand-binding domain of oestrogen receptor beta in the presence of a partial agonist and a full antagonist. *EMBO J.* **18,** 4608–4618 (1999).

308. Roe, D. R., Cheatham III, T. E. & Cheatham, T. E. PTRAJ and CPPTRAJ: software for processing and analysis of molecular synamics trajectory data. *J Chem Theory Com* **9,** 3084–3095 (2013).

309. Ye, L., Larda, S. T., Frank Li, Y. F., Manglik, A. & Prosser, R. S. A comparison of chemical shift sensitivity of trifluoromethyl tags: optimizing resolution in 19F NMR studies of proteins. *J. Biomol. NMR* **62,** 97–103 (2015).

310. Anthis, N. J. & Clore, G. M. Visualizing transient dark states by NMR spectroscopy. *Q. Rev. Biophys.* **48,** 35–116 (2015).

311. Lowe, D. M., Winter, G. & Fersht, A. R. Structure-Activity Relationships in Engineered Proteins: Characterization of Disruptive Deletions in the a-Ammonium Group Binding Site of Tyrosyl-tRNA Synthetase. *Biochemistry* **26,** 6038–6043 (1987).

312. Berger, J. *et al.* A PPARγ mutant serves as a dominant negative inhibitor of PPAR signaling and is localized in the nucleus. *Mol. Cell. Endocrinol.* **162,** 57–67 (2000).

313. Masugi, J., Tamori, Y. & Kasuga, M. Inhibition of adipogenesis by a COOH-terminally truncated mutant of PPARγ2 in 3T3-L1 cells. *Biochem. Biophys. Res. Commun.* **264,** 93–99 (1999).

314. Zhang, J., Hu, X. & Lazar, M. a. A novel role for helix 12 of retinoid X receptor in regulating repression. *Mol. Cell. Biol.* **19,** 6448–57 (1999).

315. Fersht, A. R. Relationships between Apparent Binding Energies Measured in Site-Directed Mutagenesis Experiments and Energetics of Binding and Catalysis. *Biochemistry* **27,** 1577–1580 (1988).

316. Creighton, T. E. *Proteins: Structures and Molecular Properties*. (1992).

317. Mottis, A., Mouchiroud, L. & Auwerx, J. Emerging roles of the corepressors NCoR1 and SMRT in homeostasis. *Genes Dev.* **27,** 819–835 (2013).

318. Pochetti, G. *et al.* Structural insight into peroxisome proliferator-activated

receptor γ binding of two ureidofibrate-like enantiomers by molecular dynamics, cofactor interaction analysis, and site-directed mutagenesis. *J. Med. Chem.* **53,** 4354–4366 (2010).

319.   Maier, J. A. *et al.* ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* **11,** 3696–3713 (2015).

320.   Joung, S. & Cheatham, T. E. Molecular dynamics simulations of the dynamic and energetic properties of alkali and halide ions using water-model-specific ion parameters. *J. Phys. Chem. B* **113,** 13279–13290 (2009).

321.   Miao, Y. *et al.* Improved reweighting of accelerated molecular dynamics simulations for free energy calculation. *J. Chem. Theory Comput.* **10,** 2677–2689 (2014).

322.   Liang, H. *et al.* Whole body overexpression of PGC-1 has opposite effects on hepatic and muscle insulin sensitivity. *AJP Endocrinol. Metab.* **296,** E945–E954 (2009).

323.   Acton, J. J. *et al.* Benzoyl 2-methyl indoles as selective PPARγ modulators. *Bioorganic Med. Chem. Lett.* **15,** 357–362 (2005).

Chapter 7: Appendices

7.1 Supplementary material for chapter 2

Supplementary Information for

Defining a conformational ensemble that directs activation of PPARγ

Ian M. Chrisman, Michelle D. Nemetchek, Ian Mitchelle S. de Vera, Jinsai Shang, Zahra Heidari, Yanan Long, Hermes Reyes-Caballero, Rodrigo Galindo-Murillo, Thomas E. Cheatham III, Anne-Laure Blayo, Youseung Shin, Jakob Fuhrmann, Patrick R. Griffin[4], Theodore M. Kamenecka, Douglas J. Kojetin, and Travis S. Hughes
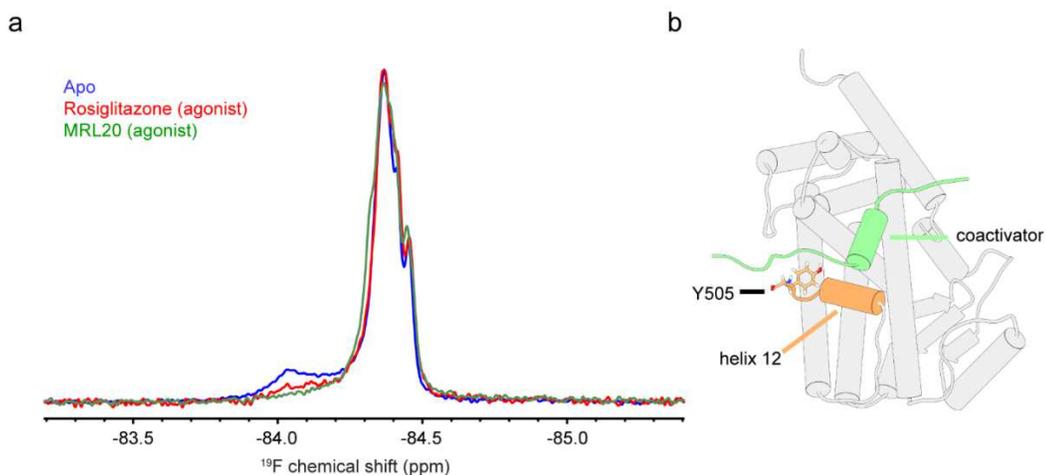
Supplementary Figure 1 | The fluorine NMR spectrum of Y505C C313S PPARγ LBD is largely ligand independent. a) Y505C C313S PPARγ LBD was labeled with BTFA and fluorine NMR was performed at ~376 MHz at room temperature either in apo form (blue spectrum) or in the presence of two agonists (rosiglitazone; red and MRL20; green). b) Y505 is located at the unstructured C-terminus of the PPARγ LBD, which likely contributes to the relative lack of change upon ligand binding. This protein was not delipidated.
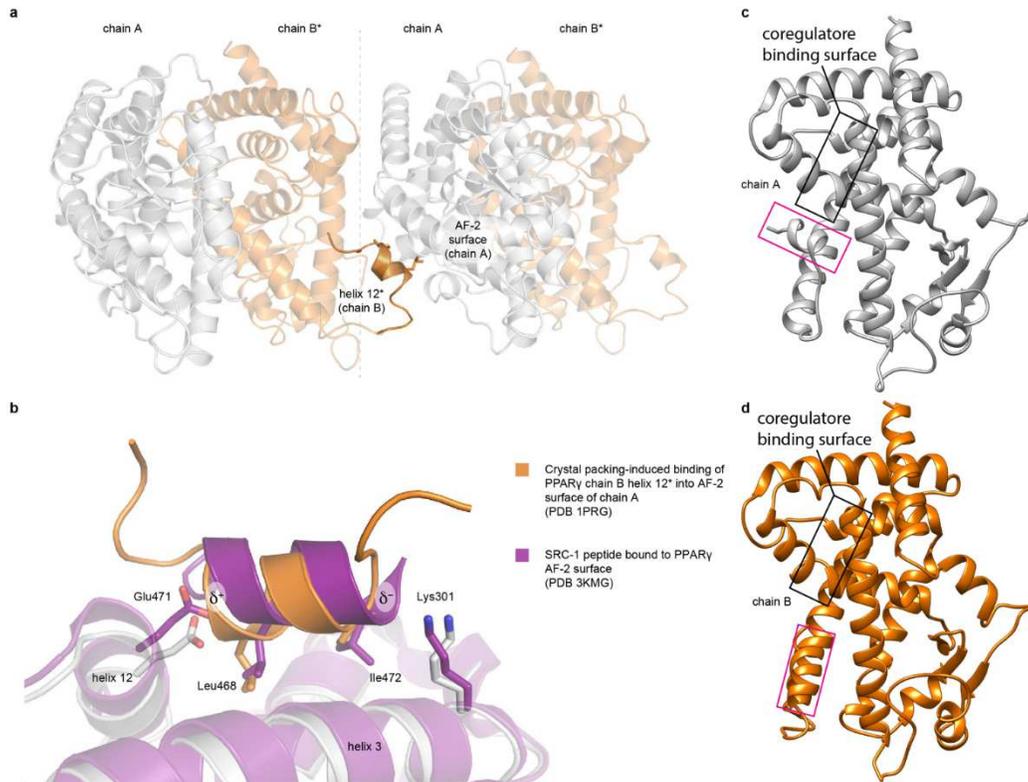
Supplementary Figure 2 | Influence of crystal packing on PPARγ LBD helix 12 conformation. The PPARγ LBD often crystalizes as a homodimer where helix 12 of the symmetry-related B chain (helix 12*) associates with the AF-2 surface of the A chain (a), adopting a conformation similar to a coactivator peptide bound to the AF-2 surface (b). This crystallization artifact distorts the conformation of the B chain helix 12* and influences the conformation of the A chain helix 12 through the formation of "charge clamp" hydrogen bonds with a positively charged (Lys301) of A chain helix 3 and negatively charged (Glu471) residues of A chain helix 12. This creates a dipolar coactivator mimicking protein-protein interaction (δ+/δ-) between monomeric subunits that is further stabilized by hydrophobic interactions originating from Leu468 and Ile472 of B chain helix 12*, which biases the observed structure into an "active" (A chain) or "pseudo-inactive" (B chain) helix 12 conformation. (c) Chain A and (d) chain B forms of PPARγ (PDB code: 1PRG) with helix12 and the coregulator binding surface highlighted.
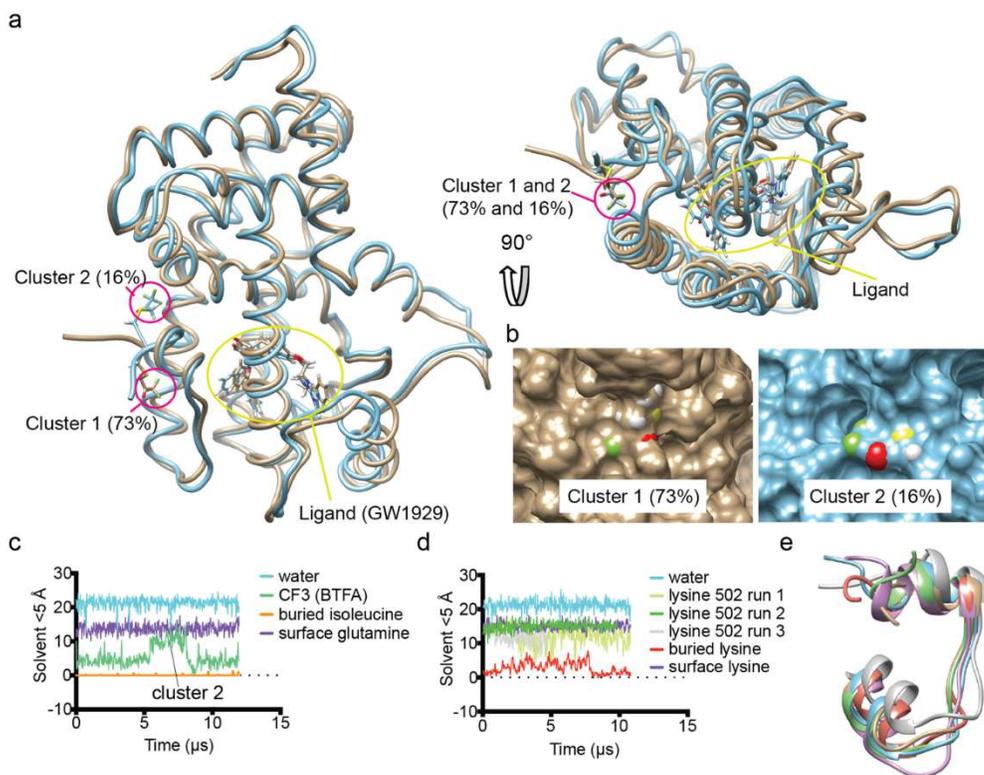
Supplementary Figure 3 | Molecular dynamics simulations indicate that the "agonist" helix 12 crystal conformation results in a solvent exposed side chain zeta nitrogen of residue K502 and terminal carbon of K502C-BTFA. (a) PPARγ^{K502C}-BTFA bound to GW1929 was simulated in explicit water and salts for ~12 μs and clustered into five clusters by label (cysteine-BTFA) position. Representative structures from the two most populated clusters (1 and 2 89% of total frames) are shown. The percent of the total simulation time spent in each cluster is highlighted. The trifluoromethyl of the BTFA label and the ligand are highlighted. Fluorine atoms are green. The remaining 11% of the frames are similar to those in cluster 1 and 2 except for a cluster of 3% of the frames where the BTFA is near the ligand and has very low solvent exposure. (b) The CF$_3$ group of K502C-BTFA is on the surface of the protein in these two clusters (c) Solvent exposure analysis of the PPARγ^{K502C}-BTFA simulation. The number of solvent molecules (water and ions) within 5 angstroms of the trifluoromethyl carbon of BTFA, and terminal carbons on the indicated amino acids and a randomly chosen water molecule are shown. (d) Three independent simulations were performed of GW1929 bound to PPARγ LBD totaling ~25 μs. The number of solvent molecules (water and ions) within 5 angstroms of the zeta Nitrogen of K502, other control lysines, including one pointing into the ligand binding pocket (buried Lysine), and a randomly chosen water molecule are also shown. (e) Representative structures from clustering of the PPARγ^{K502C}-BTFA simulation according to helix 12 RMSD into five clusters along with a crystal structure of an "active" helix 12 conformation (1PRG chain A; grey).

177
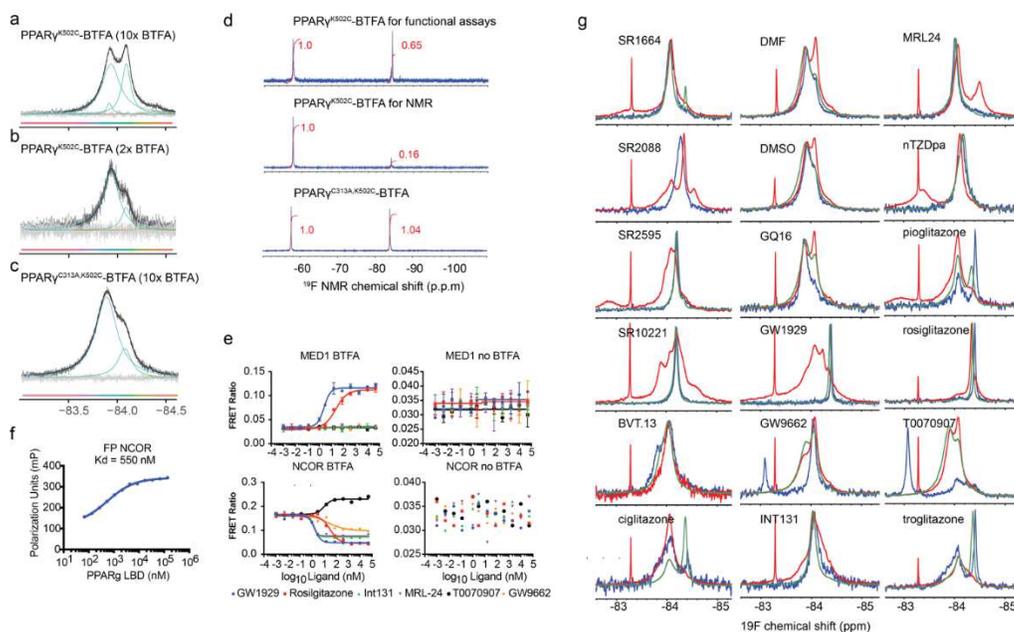
Supplementary Figure 4 | ^{19}F NMR indicates that K502C is preferentially labeled over C313 by BTFA and PPARγ^{K502C} without BTFA label is not active in TR-FRET. PPARγ^{K502C} which contains a native cysteine (C313), and an introduced cysteine (K502C) was incubated with either (a) 10x or (b) 2x molar ratio of BTFA. (c) PPARγ^{C313A,K502C} which contains a single cysteine (K502C) was incubated with 10x BTFA. The PPARγ^{C313A,K502C}-BTFA spectrum is very similar to the 2x BTFA treated PPARγ^{K502C} in panel b indicating that K502C is labeled preferentially. The spectra shown in panel b and c are shown in other figures in this manuscript (d) PPARγ^{K502C} was prepared for use in functional assays (TR-FRET and FP) or NMR and loaded with either 1.1 to 1.25 molar equivalents of MRL24 (which contains a CF$_3$ group; left peak). The signal from MRL24 (left peak) and BTFA (right peak) were integrated (red numbers in figure). The integral of the left peak indicates the amount of protein, while the right peak indicates the fraction labeled. (e) Ligand dependent recruitment of NCOR or MED1 peptide to PPARγ^{K502C} (no BTFA) or PPARγ^{K502C}-BTFA was measured using TR-FRET for select ligands. The failure of non-BTFA labeled PPARγ^{K502C} to produce a TR-FRET signal was also seen in two other separate TR-FRET experiments using GW1929. Mean of two technical replicates and standard deviation are shown. (f) The non-BTFA labeled PPARγ^{K502C} used in panel e has the expected affinity for NCOR as measured by fluorescence polarization (also see Supplementary Figure 7b) (g) Comparison of PPARγ^{C313A,K502C}-BTFA (green; only labeled on C502) with PPARγ^{K502C}-BTFA2x that was loaded *first* with indicated ligands and then exposed to a 2x molar ratio of BTFA (blue; presumably labeled only on C502) and PPARγ^{K502C}-BTFA10x that was exposed to a 10x molar ratio and then loaded with the indicated ligands (red; presumably labeled on both C502 and C313). PPARγ^{K502C}-BTFA2x is more similar to PPARγ^{C313A,K502C}-BTFA than PPARγ^{K502C}-BTFA10x in all cases except for the ligands which covalently attach to C313. (Dimethylformamide; DMF and Dimethyl sulfoxide; DMSO). The narrow peak at ~-83.3 ppm is free BTFA. All spectra displayed in panel g except for the red spectra are displayed in other figures in this manuscript.
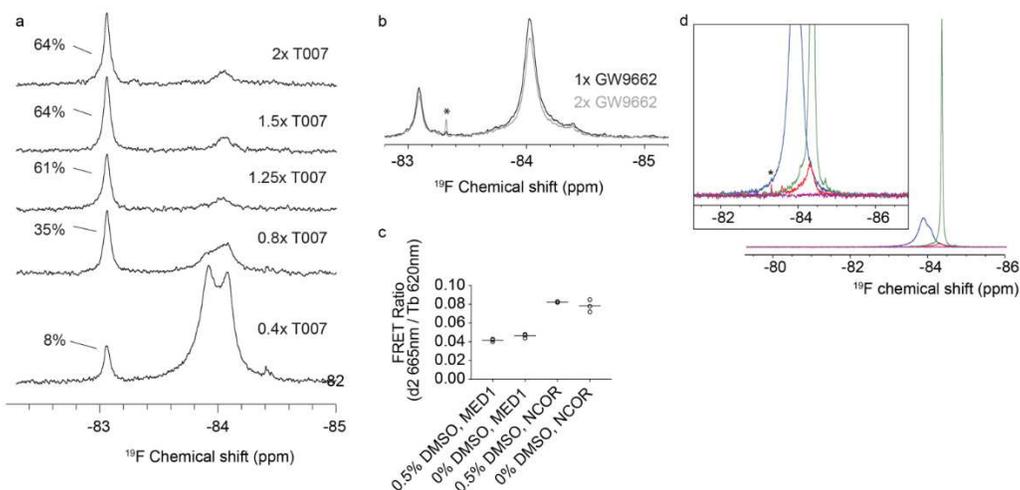
Supplementary Figure 5 | Covalent ligands do not bind to C502, DMSO has a minimal effect on TR-FRET, and signal from any BTFA labeled contaminating protein is undetectable. (a-b) The indicated molar ratios of a) T0070907 or b) GW9662 were added to PPARγ^{K502C} followed by labeling with 10x BTFA. The percent of the total signal area found in the left peak is indicated in panel a. (c) TR-FRET demonstrates that ligand vehicle (DMSO) has negligible impact on coregulator recruitment to apo PPARγ LBD. Three technical replicates are shown along with the mean. (d) PPARγ^{C313A} LBD (no cysteines) without (purple) or with (red) 6x histidine tag was treated with 10x molar ratio of BTFA and measured with fluorine NMR using the same processing and acquisition parameters as PPARγ^{C313A,K502C}-BTFA not bound to ligand (blue) or bound to GW1929 (green). Cleavage of the 6x histidine tag increases the purity of the protein because after cleavage the prep is re-run over the nickel column and contaminating proteins stick to the column while PPAR flows through. The difference in purity that histidine cleavage makes is large, with virtually no contaminating protein detected after cleavage. All spectra shown in this work except supplementary figure 2 used proteins with the histidine tag cleaved off, thus, NMR signal from impurities would not be expected to be observed, however any signal from protein impurities would be expected to be broad and centered at ~-84.35ppm as shown here. A small broad upfield shifted peak around ~-84.35 ppm appears inconsistently in deconvolutions of some spectra. As demonstrated by this control, it is possible, but not likely, that this signal near -84.35 ppm originates from a small amount of contaminating protein that is labeled with BTFA. Inset displays zoomed view of main figure. The * denotes signal from free BTFA. The blue and green spectra in panel d and the spectra in panel b are displayed in other figures in this manuscript.
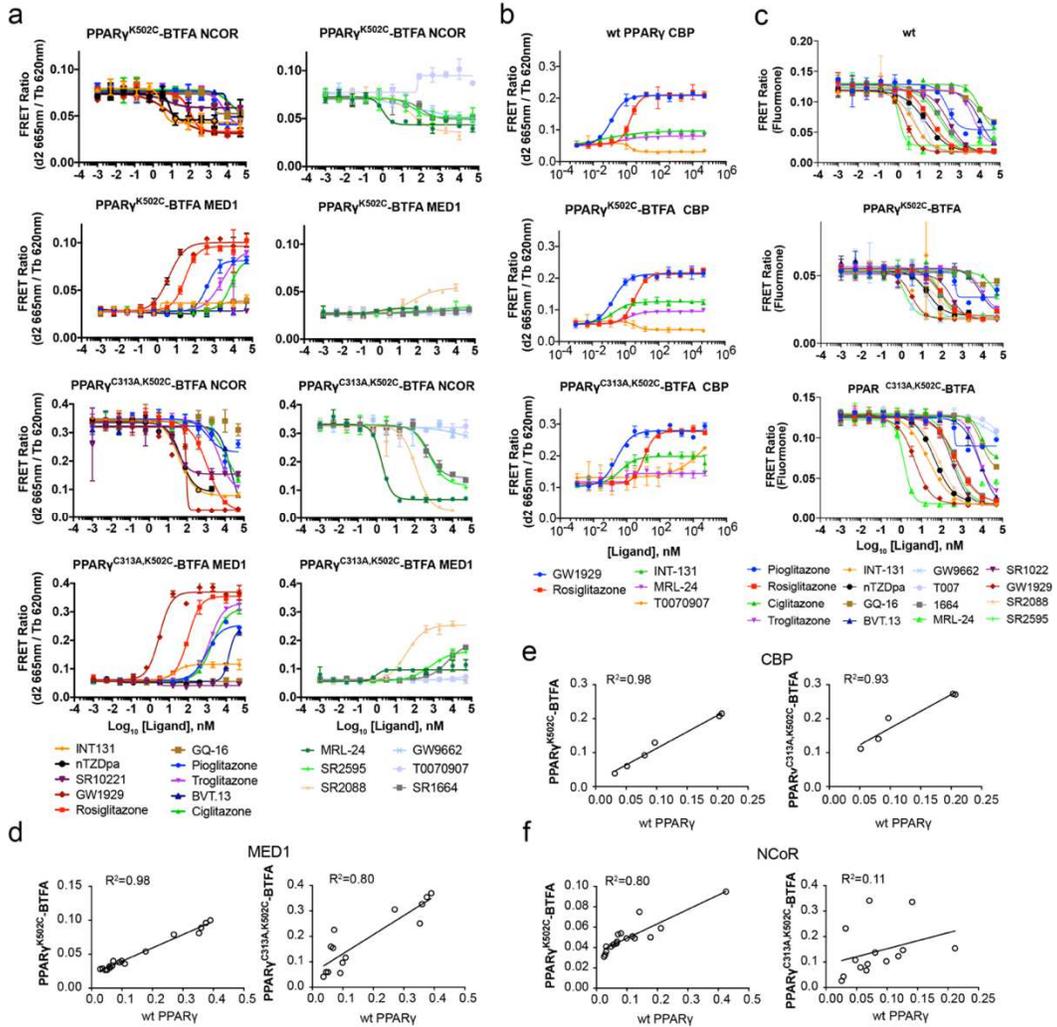
Supplementary Figure 6 | TR-FRET indicates that labeled proteins are functional, that PPARγ^{K502C}-BTFA is most functionally similar to PPARγ. (a) Titration of PPARγ^{C313A,K502C}-BTFA and PPARγ^{K502C}-BTFA with a diverse set of ligands in the TR-FRET assay using NCoR or MED1 biotinylated peptides and His-tagged PPARγ LBD (EC$_{50}$ values are shown in Supplementary Table 3). (b) Titration of PPARγ, PPARγ^{C313A,K502C}-BTFA and PPARγ^{K502C}-BTFA with a subset of ligands in the TR-FRET assay using CBP biotinylated peptides and His-tagged PPARγ LBD (EC$_{50}$ values are shown in Supplementary Table 4). (c) Titration of the ligands into protein (as labeled) preloaded with fluorescently labeled ligand (Fluormone Pan-PPAR Green) was used to determine Ki values for ligands to the proteins. Fitted Ki values are displayed in Supplementary Table 1. Mean and standard deviation (n=2) are shown along with fitted curve (see methods). Error bars represent the standard deviation between two technical replicates in a single experiment. The competition assay to determine K$_i$ values was performed a single time. The TR-FRET with PPARγ^{K502C}-BTFA and PPARγ^{C313A,K502C}-BTFA was performed three times with similar results. TR-FRET with CBP was performed two times with similar results. (d-f) Comparison of PPARγ, PPARγ^{K502C}-BTFA and PPARγ^{C313A,K502C}-BTFA TR-FRET ratios at saturating ligand concentrations for (d) MED1 (e) CBP and (f) NCoR. Correlation coefficient is indicated (R^2).

180
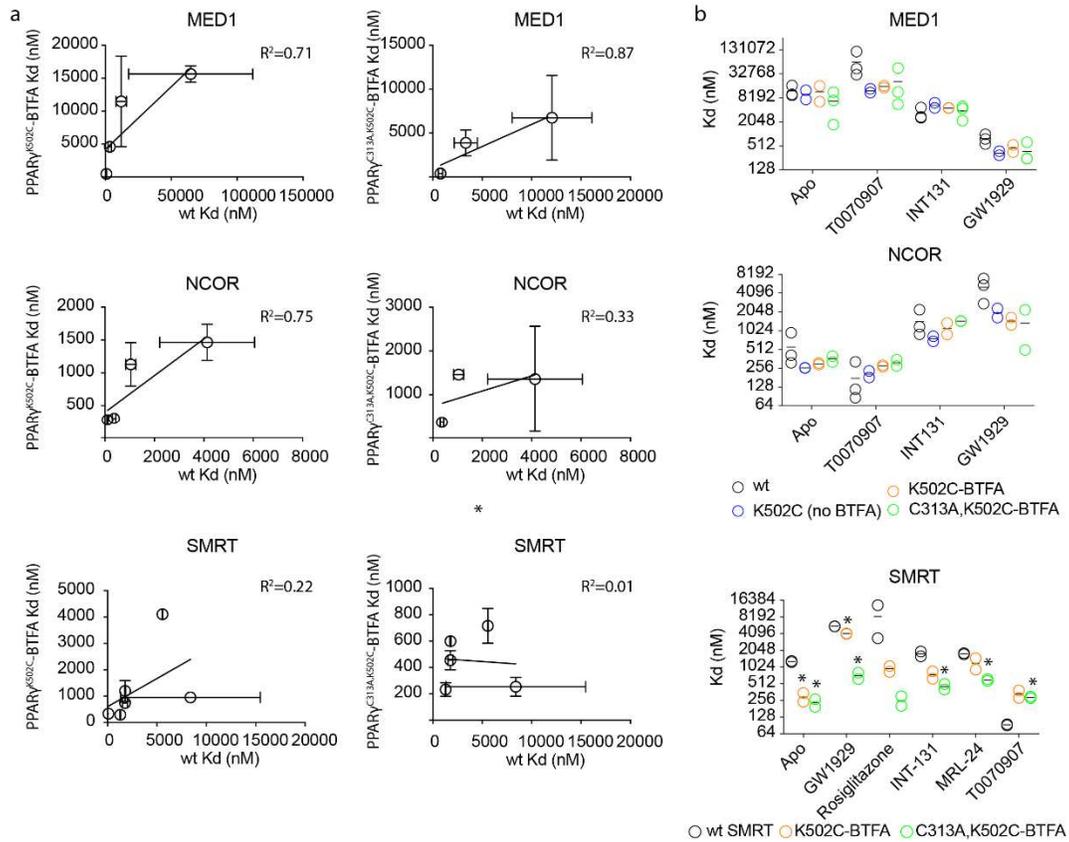
Supplementary Figure 7 | Fluorescence polarization indicates that labeled proteins are functional, that PPARγ^{K502C}-BTFA is most similar to PPARγ and that SMRT binding is perturbed by labeling. (a) Comparison of PPARγ with PPARγ^{K502C}-BTFA and PPARγ^{C313A,K502C}-BTFA affinity for MED1, NCoR and SMRT peptides either without ligand or saturated with T0070907, GW1929 or INT-131 as measured by fluorescence polarization. Assays involving SMRT peptide recruitment also included rosiglitazone and MRL24. Error bars represent standard deviation of 2 (SMRT and wt NCoR) or 3 (PPARγ^{C313A,K502C}-BTFA NCoR and MED1) independently run assays. Correlation coefficient is indicated (R^2). (b) Comparison of peptide affinities for wt and labeled/mutant versions of PPARγ either ligand free or saturated with the indicated ligands. Significant differences between wt and the labeled/mutant PPARγ are indicated by asterisks ($p < 0.05$; unpaired t test with Holm-Sidak correction for multiple comparisons).
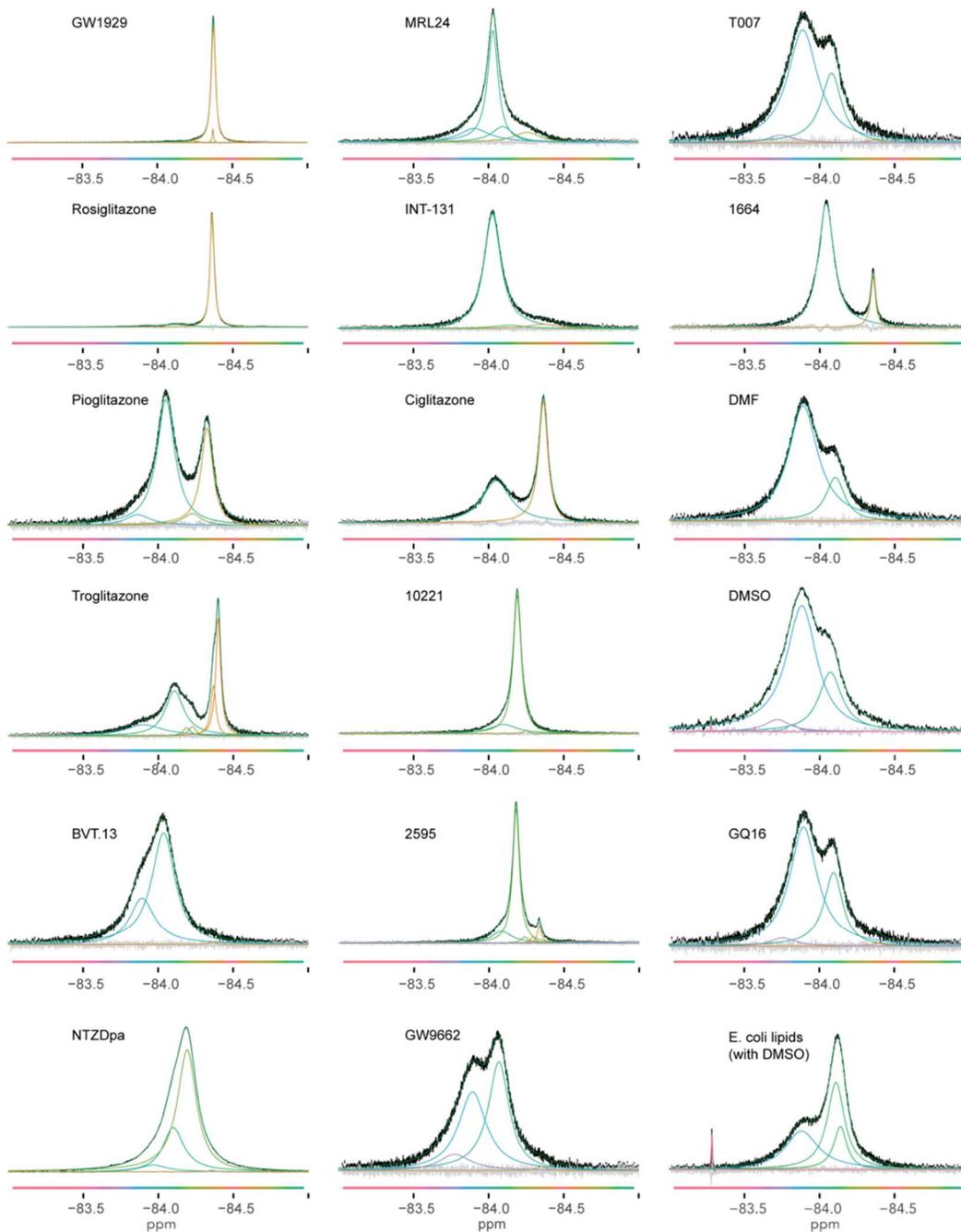
Supplementary Figure 8 | ^{19}F NMR spectra of PPARγ^{C313A,K502C}-BTFA bound to 16 pharmacologically distinct ligands is similar to PPARγ^{K502C}-BTFA. The PPARγ^{C313A,K502C}-BTFA spectra are higher signal to noise than the PPARγ^{K502C}-BTFA spectra because the C313A mutation allows complete labeling with higher BTFA concentrations without spurious labeling of C313 (Supplementary Fig. 4). GQ16 used DMF for a vehicle. All other ligands used DMSO. DMSO and DMF concentrations are the same in all spectra. This DMSO spectrum is also displayed in other figures in this manuscript. The estimated degree of saturation with ligand in these spectra is shown in Supplementary Table 2. Some of these spectra were replicated (Supplementary Fig. 10)
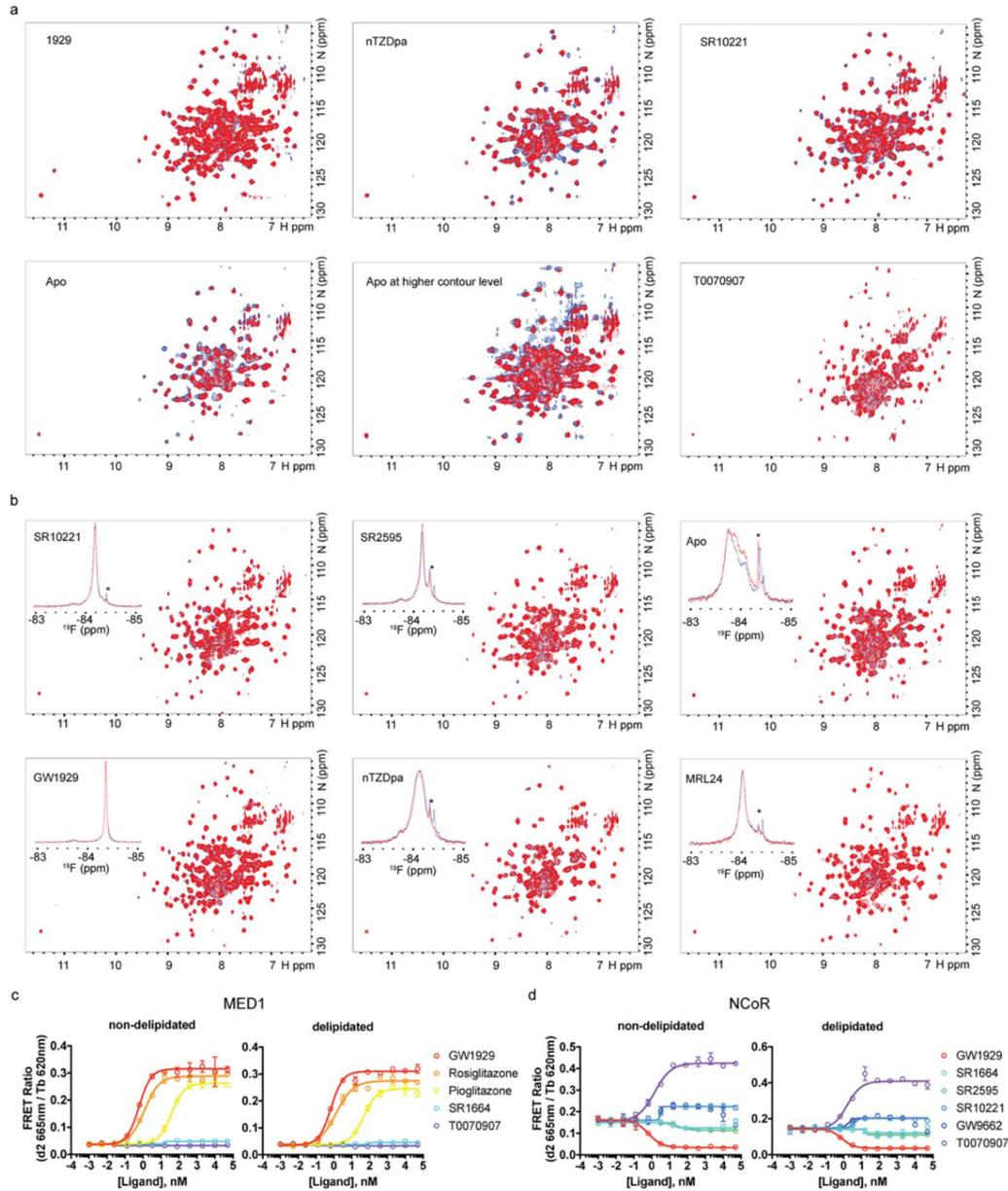
Supplementary Figure 9 | NMR indicates that delipidation produces a small difference in backbone structure and in coregulator recruitment. (a-b) Protein was delipidated (red spectra) or not (blue spectra) and then loaded with the indicated ligand or not loaded with ligand (Apo). 2D [^{1}H,^{15}N]-TROSY-HSQC NMR spectra (Bruker pulse program trosyf3gpphsi19.2) of (a) ^{15}N-labeled PPARγ LBD or (b) ^{15}N-labeled PPARγ^{C313A}-BTFA. Apo spectra are shown at two different contour levels in panel a for clarity of changes vs. liganded states. Insets in panel b show ^{19}F NMR spectra of the same sample. Sharp upfield shifted peaks marked with an asterisk in panel b are likely misfolded protein or another artifact. (c) PPARγ LBD was delipidated or not delipidated prior to measuring coregulator recruitment in response to ligand titration via TR-FRET. No significant difference is observed. Mean ± SE shown. Error bars represent the standard deviation of two technical replicates. TR-FRET was repeated in two independent experiments with similar results. NMR was done once.
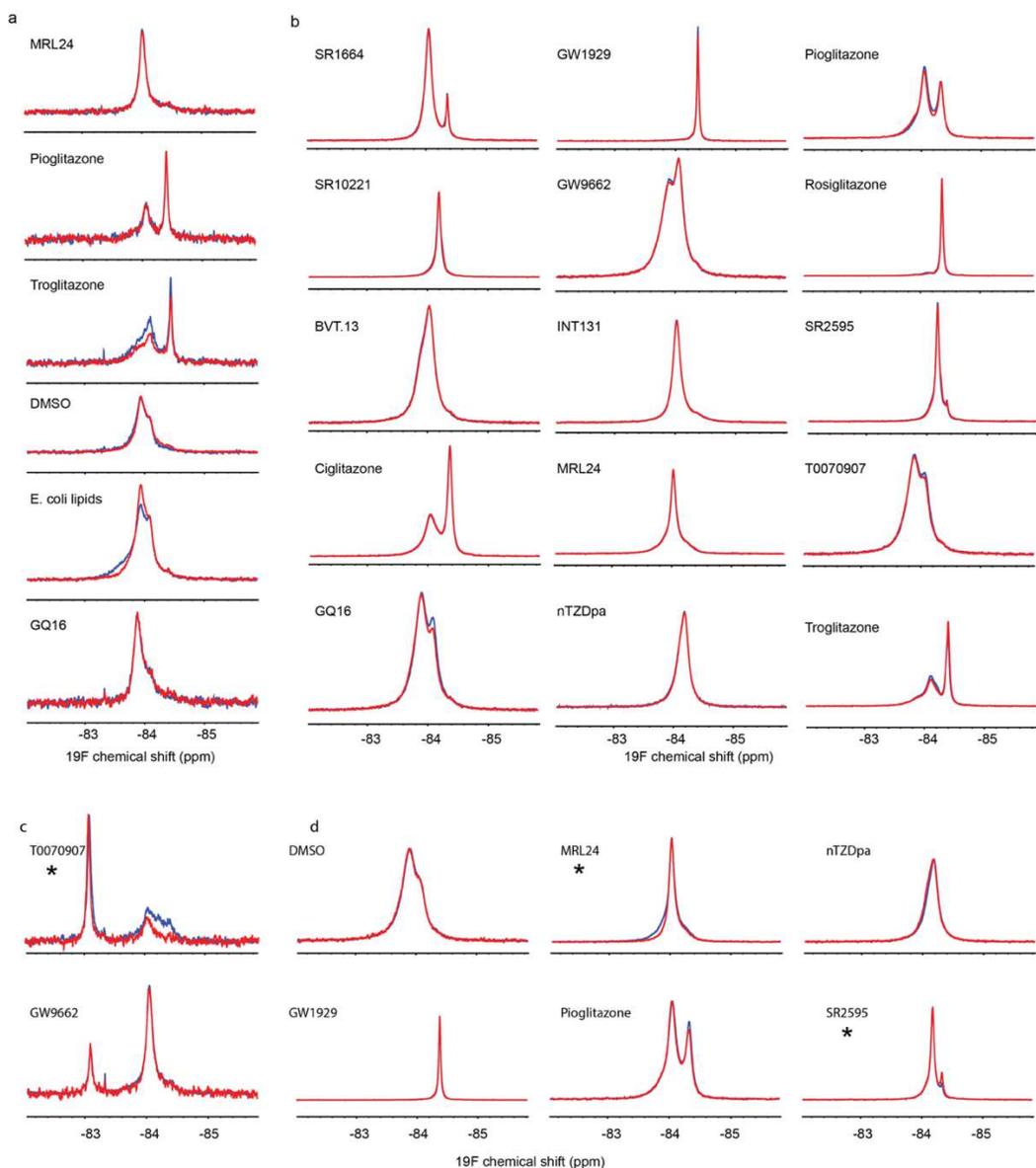
Supplementary Figure 10 | NMR samples are stable over time and replicate samples from different protein preparations produce very similar spectra. Fluorine NMR was collected on (a) PPARγ^{K502C}-BTFA or (b) PPARγ^{C313A,K502C}-BTFA loaded with the indicated ligands initially (blue) and then after two to nine days of storage at room temperature (red). The changes in the *E. coli* lipid panel are likely due to lipid precipitation after unbinding from PPARγ. The blue spectra are displayed in other figures in this manuscript. (c-d) Replicate samples from the same or different protein preparations. (c) replicate samples using PPARγ^{K502C}-BTFA and (d) PPARγ^{C313A,K502C}-BTFA (* denote samples from different protein preparations). Spectra are scaled relative to each other for ease of comparison. All of the spectra in panels c and d are shown elsewhere in this manuscript except for one of the spectra each of pioglitazone and nTZDpa bound to PPARγ^{C313A}-BTFA.
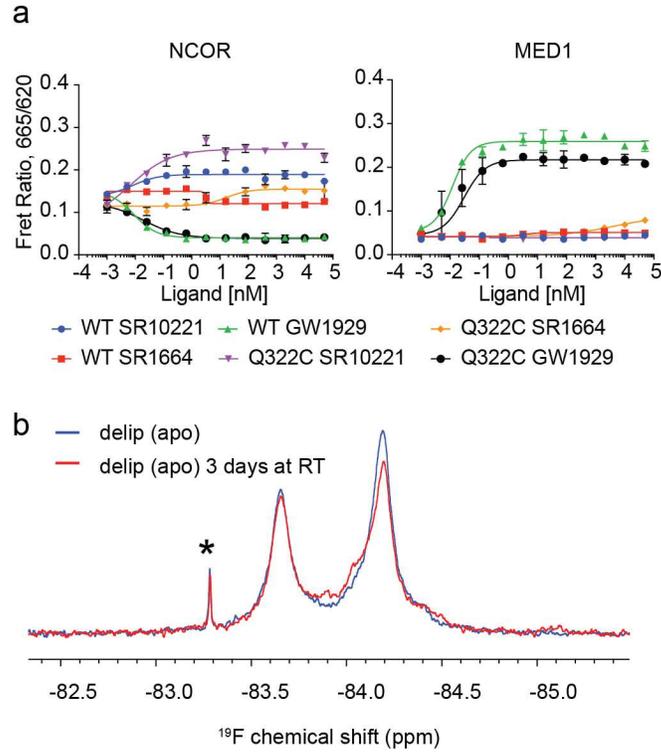
Supplementary Figure 11 | PPARγ^{C313A,Q322C}-BTFA recruits coregulators with similar efficacy to that of wild type PPARγ and the apo PPARγ^{C313A,Q322C}-BTFA spectrum is stable over time. (a) TR-FRET ratio change with increased ligand concentration indicates a change in affinity between the indicated peptide and delipidated PPARγ^{C313A,Q322C}-BTFA induced by ligand binding. (b) ^{19}F NMR spectra of delipidated apo PPARγ^{C313A,Q322C}-BTFA initially and then after 3 days at ambient temperature. The * indicates free BTFA signal. Error bars represent the standard deviation of two technical replicates in a single experiment. TR-FRET was performed twice in two independent experiments with similar results.
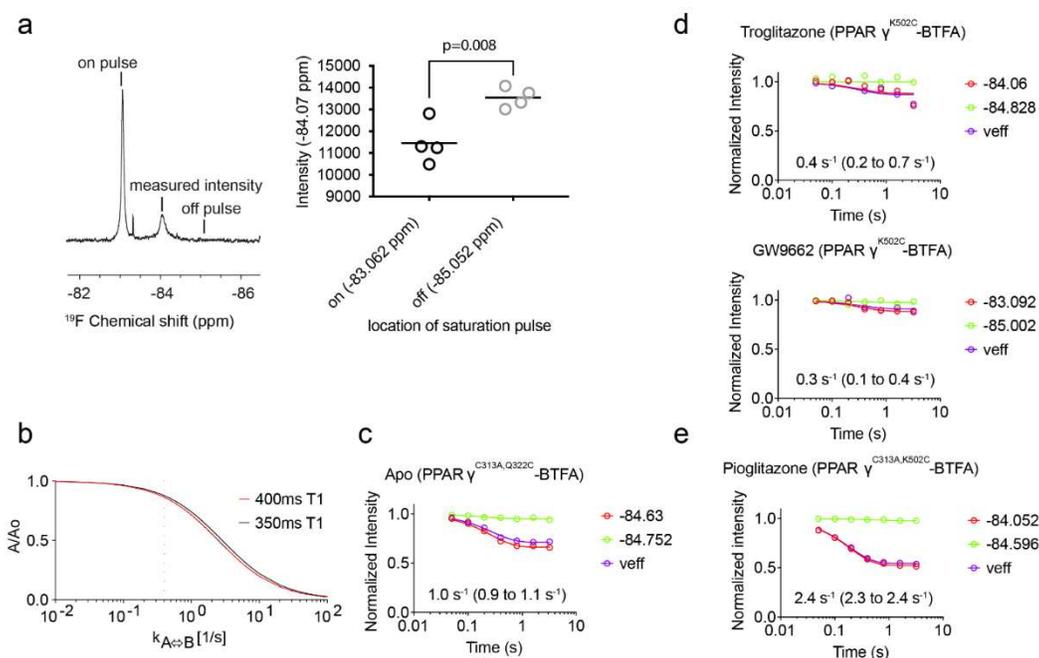
Supplementary Figure 12 | Chemical exchange occurs between ^{19}F NMR resolved peaks.
Chemical Exchange Saturation Transfer (CEST) was performed on probes on helix 12 (a,d,e) and on helix 3 (c). (a) Both on (-83.062 ppm) and off (-85.052 ppm) resonance selective saturating pulses were performed and the height of the smaller peak (-84.07 ppm) in the ^{19}F spectrum of T0070907-bound PPARγK502C-BTFA was monitored (left panel). The p value of a two-tailed t-test comparing the mean intensity (shown) of the smaller peak with on and off resonance selective pulses is shown. While exchange is detected, the exchange is likely very slow, resulting in very small changes in intensity of the small peak. (b) Calculation to determine the detection limits of CEST in our system. The intensity of a peak (peak A) as a function of the rate of exchange with a peak that is selectively saturated (peak B) normalized by the intensity of peak A in the absence of exchange (Ao). Calculations are shown using T1 values of 400 ms (red) and 350 ms (black) and a 1.6 second selective saturation pulse (which is what was used in the CEST figures). Calculation was done using Equation 50 from (Journal of Biomolecular NMR 2000 18:49-63). Given experimental noise and peak A and B sharpness in our experiments exchange may be difficult to reliably detect with exchange rates less than ~0.4 s^{-1} (I/Io>0.86; indicated by dotted vertical line in graph). In addition, exchange rates faster than 10 s^{-1} can be detected but may be difficult to quantify given that we used 50 ms selective pulses. (c-e) Exchange rates are indicated with 95% confidence interval in parentheses. An off resonance soft pulse equidistant from the peak of interest (green) was subtracted from an on resonance soft pulse (red) to yield Veff (purple). Fits were carried out fixing T$_1$ at the experimentally verified rate (Supplementary Table 5) and experimentally obtained non-exchange intensity. Thus, only one parameter was fit leaving 6 degrees of freedom in each fit. These experiments were performed once.

186
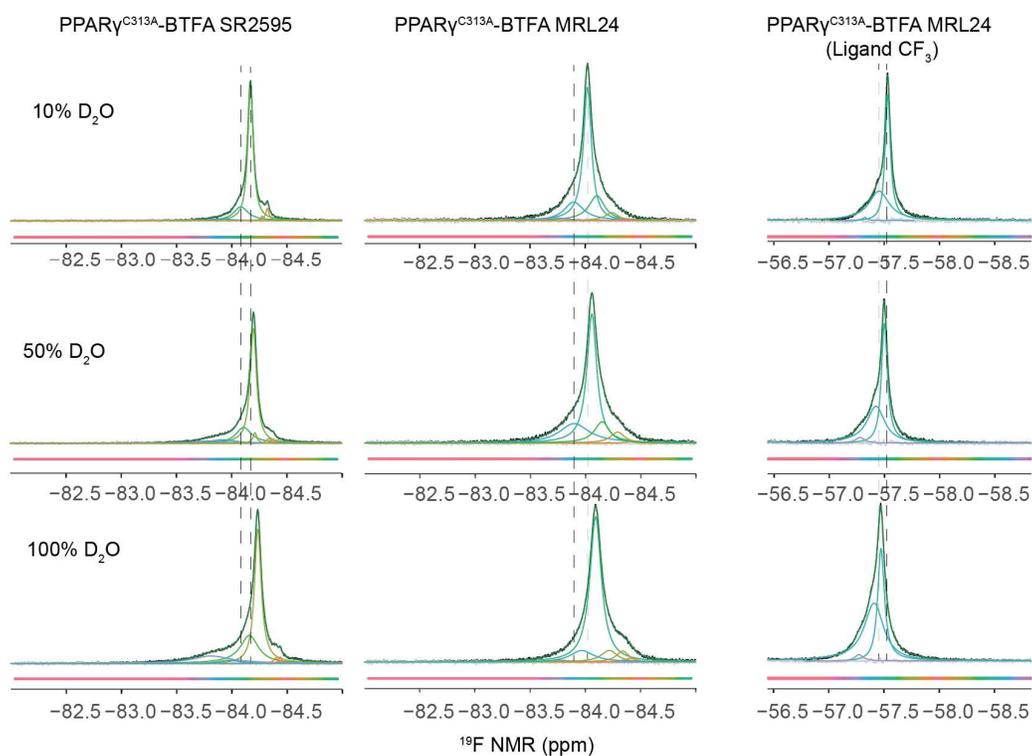
Supplementary Figure 13 | ¹⁹F NMR was performed on samples of the indicated protein and ligand complexes in the presence of 10% (upper panel), 50% (middle panel), or 100% (lower panel) deuterium oxide (D₂O). Vertical lines indicate the chemical shift of select peaks at 10% D₂O concentration and are included to aid in comparison of peak position between different D₂O concentrations. These experiments were performed once.
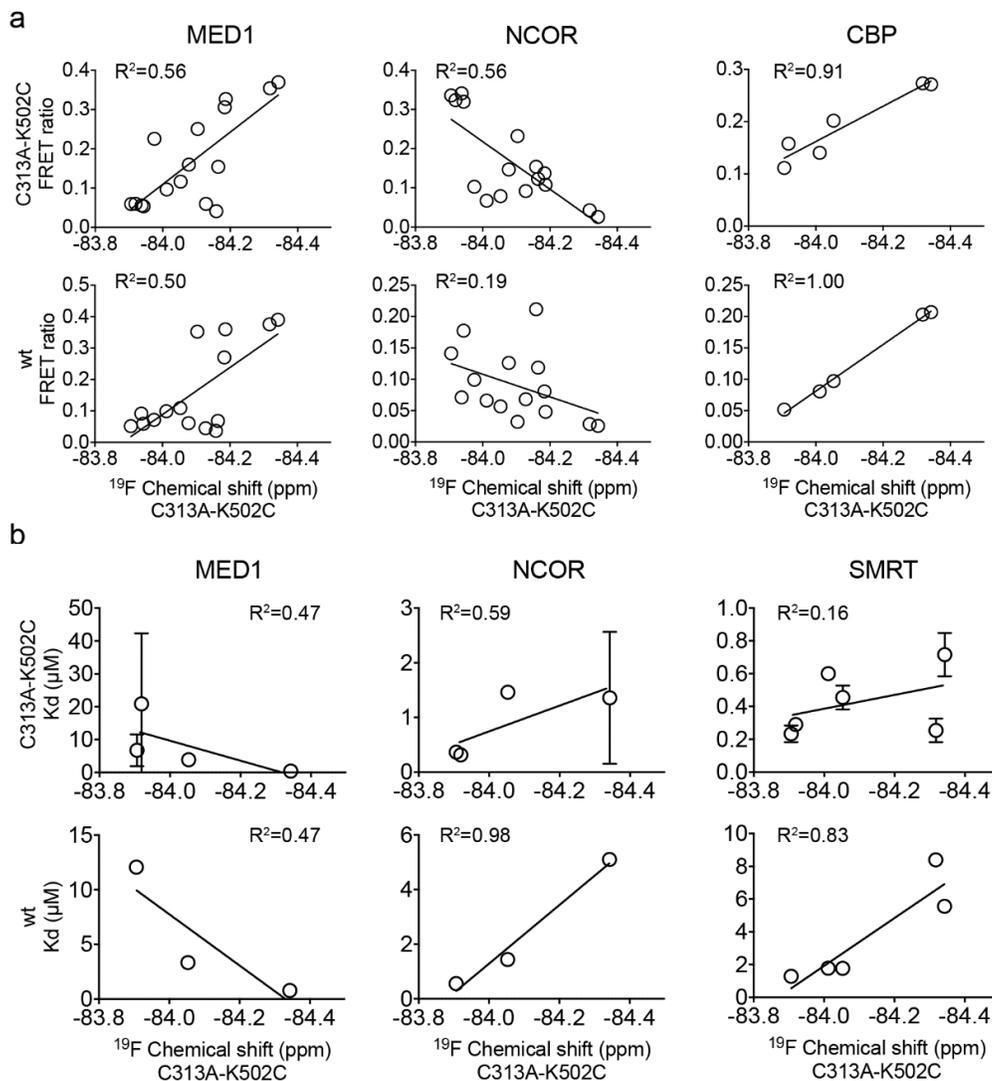
Supplementary Figure 14 | Ligand-directed helix 12 ensemble dictates PPARγ[C313A,K502C]-BTFA-coregulator interaction. (a) Plot of mean ^{19}F NMR chemical shift values (PPARγ[C313A,K502C]-BTFA) versus TR-FRET endpoint data for the recruitment of MED1, NCoR and CBP peptides to PPARγ[C313A,K502C]-BTFA (top panels) and wt PPARγ LBD (bottom panels) for the set of 16 pharmacologically distinct synthetic PPARγ ligands and apo-protein (select ligands for CBP). TR-FRET was performed in three separate experiments for MED1 and NCoR recruitment and two separate experiments for CBP with similar results. b) Plot of mean ^{19}F NMR chemical shift values (PPARγ[C313A,K502C]-BTFA) versus MED1, NCoR and SMRT peptide dissociation constant (Kd) for PPARγ[K502C]-BTFA (top panels) and wt PPARγ LBD (bottom panels) as measured by fluorescence polarization (FP) for a subset of the ligands in panel a. Mean and standard deviation for FP represent 2 (NCOR binding to PPARγ[C313A,K502C]-BTFA and SMRT) or 3 (NCOR binding to wt and MED1) independent replications. Linear regression fit is shown as a solid line and the correlation coefficient (R^2) for the fitted line is indicated. Covalent ligands, which require C313 for binding are not included in the wt graphs as the NMR chemical shift would not be expected to correlate with the FP or TR-FRET values.

188
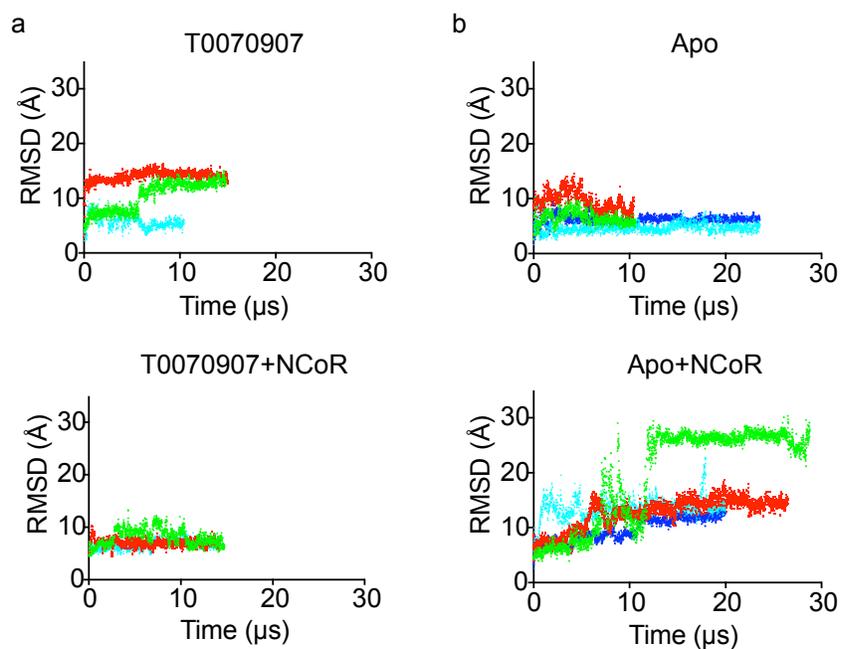
Supplementary Figure 15 | Molecular simulations of PPARγ LBD bound to (a) T0070907 (covalent inverse agonist) alone or co-bound to the corepressor peptide NCoR (lower panel) or (b) apo or apo bound to NCoR (lower panel). The root mean square deviation of helix 12 compared to the conformation of helix 12 in the crystal structure that was used to build these molecules (chain B of 3B0R and 1PRG). Each color represents an independent simulation started from the same coordinates but with different initial atom velocities. Apo without NCoR stays the closest to the starting crystal structure.
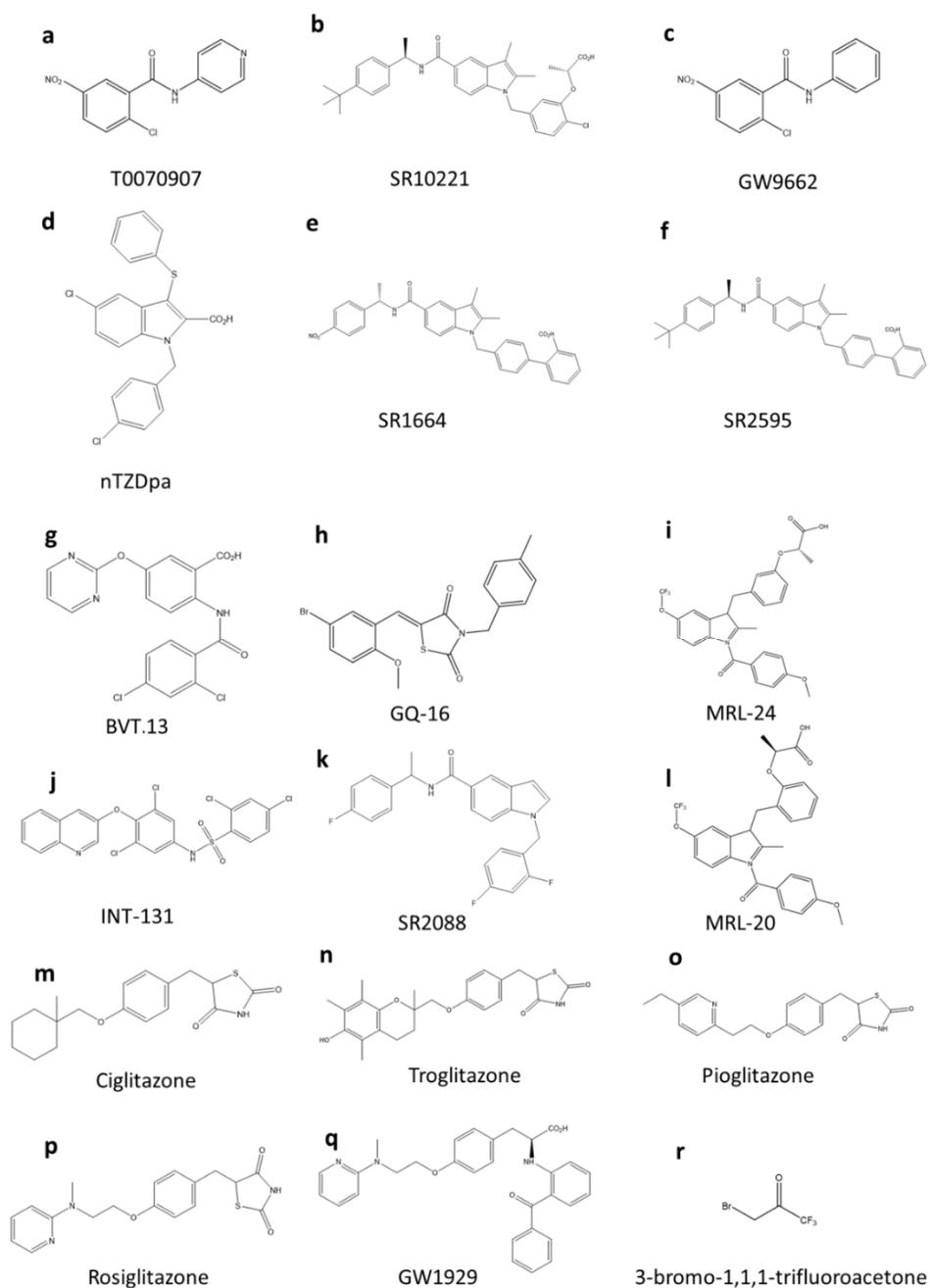
Supplementary Figure 16 | Synthetic PPARγ ligands and ¹⁹F label used in this study. a-q) Chemical structures of the ligands used in this study. r) Chemical structure of the 3-bromo-1,1,1-trifluoroacetone label used in this study

Supplementary Table 1 | Effect of labeling on ligand affinity

Ligand	#EC$_{50}$ (nM)			*K$_i$ (nM)			Fold increase in K$_i$ compared to wild-type	
	wt	PPARγ^{K502C}-BTFA	PPARγ^{C313A,K502C}-BTFA	wt	PPARγ^{K502C}-BTFA	PPARγ^{C313A,K502C}-BTFA	PPARγ^{K502C}-BTFA	PPARγ^{C313A,K502C}-BTFA
Pioglitazone	247	428	412	121	696	1137	6	9
Rosiglitazone	32	81	516	16	132	1422	8	92
Ciglitazone	9930	>50uM	9034	4875	>50uM	24921	36	5
Troglitazone	4699	8723	6387	2306	14177	17619	6	8
INT-131	5	66	22	3	107	60	41	23
nTZDpa	21	19	44	10	31	123	3	12
GQ-16	8481	2171	8873	4163	3528	24477	1	6
BVT.13	10105	7087	7641	4960	11518	21079	2	4
SR10221	292	282	314	143	458	866	3	6
GW1929	2	3	5	1	5	13	5	12
SR2088	33	45	68	16	74	189	5	12
SR2595	110	80	352	54	130	972	2	18
^GW9662	6	14	33204	NA	NA	NA	similar EC$_{50}$	large change in EC$_{50}$
^T0070907	15	25	>50uM	NA	NA	NA	similar EC$_{50}$	large change in EC$_{50}$
SR1664	131	74	551	64	120	1519	2	24
MRL24	0.9	1.3	1.4	0.5	2	4	5	8
					median		5	11
					mean		9	17

NA = Not applicable

#Given that 8nM protein was used in determining EC$_{50}$ values in this table any EC$_{50}$ less than 4 nM indicates that the true EC$_{50}$ is ≤ 4 nM.

^these are covalent ligands therefore EC50 doesn't have the same meaning as the non-covalent ligands

Supplementary Table 2 | Fraction of protein bound by ligand for NMR samples

	Assumes all added ligand is soluble.		Assumes 10 μM solubility of ligand.		Fraction bound according to deconvolution of NMR signal	
	PPARγ-BTFA	PPARγ^{C313A}-BTFA	PPARγ-BTFA	PPARγ^{C313A}-BTFA	PPARγ-BTFA	PPARγ^{C313A}-BTFA
Pioglitazone	1.00	0.99	0.96	0.94	1.00	0.89
Rosiglitazone	1.00	0.94	0.99	0.93	1.00	0.95
Ciglitazone	0.55	0.81	0.37	0.69	*1.00	*0.96
Troglitazone	0.92	0.85	0.76	0.74	0.74	0.81
INT-131	1.00	1.00	0.99	0.99	1.00	1.00
nTZDpa	1.00	0.99	1.00	0.99	1.00	1.00
GQ-16	0.93	0.71	0.89	0.69	#0.57	#0.29
BVT.13 Pentahydrate	0.84	0.73	0.79	0.71	0.65	0.69
SR10221	0.99	0.96	0.97	0.95	1.00	1.00
GW1929	1.00	1.00	1.00	1.00	1.00	1.00
SR2088	1.00	0.99	0.99	0.98	1.00	ND
SR2595	1.00	0.96	0.99	0.95	1.00	1.00
^GW9662	1.00	NA	1.00	NA	0.90	NA
^T0070907	1.00	NA	1.00	NA	1.00	NA
SR1664	1.00	0.94	0.99	0.93	1.00	1.00
MRL24	1.00	1.00	1.00	1.00	1.00	0.86
Vehicle (DMSO)					##0.21	##0.25

Fraction bound is calculated using Ki values for ligands.

*Discrepancy between NMR and calculation is likely a combination of inaccurate deconvolution and good solubility of ciglitazone (Cayman Chemical states that ciglitazone is soluble to 1.2mM in 25% DMSO).

^Covalent ligands were added at 2x and so are assumed to be fully bound based on EC50 values, this is supported by LC-ESI MS data.

GQ16 has very poor solubility, which would explain the low bound fraction observed via NMR.

This is either residual bound E. coli lipid or an Apo conformation that is shifted to the right.

Supplementary Table 3 | EC$_{50}$ values for MED1 and NCOR recruitment

Coregulator	MED1 EC$_{50}$ (nM)**					NCOR EC$_{50}$ (nM)**				
PPARγ variant	wt	PPARγ^{K502C}-BTFA	FC	PPARγ^{C313A,K502C}_BTFA	FC	wt	PPARγ^{K502C}-BTFA	FC	PPARγ^{C313A,K502C}_BTFA	FC
Pioglitazone	190	420	2	970	5	350	2200	6	1,200	3
Rosiglitazone	13	29	2	99	8	21	96	5	1,200	57
Ciglitazone	4200	7400	2	1800	<1	2,700	15,000	6	14,000	5
Troglitazone	580	2600	4	1400	2	1,100	6,600	6	4,600	4
INT-131	7	1	<1	18	2	2	3	1	38	19
nTZDpa*	8	<1	<1	<1	<1	4	8	2	47	12
GQ-16	730	1600	2	NC	NA	1,900	1,900	1	NC	NA
BVT.13 Pentahydrate	>50,000	>50,000	NA	14,000	<1	44,000	9,300	<1	19,000	<1
SR10221	8	NC	NA	4	<1	2	28	14	28	14
GW1929	3	4	1	3	1	2	10	5	91	46
SR2088	23	50	2	37	1	13	34	3	120	9
SR2595	100	NC	NA	680	7	75	34	<1	590	8
GW9662#	4	NC	NA	NC	NA	3	37	12	NC	NA
T0070907#	4	NC	NA	>50,000	NA	13	66	5	NC	NA
SR1664	180	NC	NA	5600	31	NC	99	NA	535	NA
MRL24*	2	1	<1	1	<1	<1	1	1	2	2

Median change§	2	1	5	9
Mean change§	2	5	5	15

NA=not applicable

NC=no significant change in TR-FRET ratio was observed, which does not allow curve fitting, which is expected for non-agonists and for non-agonists and inverse agonists in the MED1 assay.

*Data indicates two recruitment events. The EC50 is shown for the initial peptide recruitment event only. The second recruitment event likely is caused by binding of a second ligand to the same PPARγ molecule[208].

these ligands form a covalent adduct to C313 and should not bind to PPARγ^{C313A,K502C}-BTFA.

**Given that 8nM protein was used in determining EC_{50} values in this table any EC_{50} less than 4 nM indicates that the true EC_{50} is ≤ 4 nM.

§not including covalent ligands; values less than 1 (<1) are treated as 1

Supplementary Table 4 | EC$_{50}$ values for CBP recruitment

Coregulator CBP EC$_{50}$ (nM)**

PPARγ variant	wt	PPARγ^{K502C}-BTFA	Fold change	PPARγ^{C313A,K502C}-BTFA	Fold change
Rosiglitazone	2	4	2	11	8
INT-131	<1	<1	1	<1	1
GW1929	<1	<1	1	<1	1
T0070907#	2	4	2	41,000	21,000
MRL24*	<1	1	1	1	1
Median change			1		§1
Mean change			1.4		§2.8

*Data indicates two recruitment events. The EC50 is shown for the initial peptide recruitment event only. The second recruitment event likely is caused by binding of a second ligand to the same PPARγ molecule[208].

\# this ligand forms a covalent adduct to C313 and should not bind to PPARγ^{C313A,K502C}-BTFA.

**Given that 8nM protein was used in determining EC$_{50}$ values in this table any EC$_{50}$ less than 4 nM indicates that the true EC$_{50}$ is ≤ 4 nM.

§not including covalent ligands; values less than 1 (<1) are treated as 1

Supplementary Table 5 | NMR derived spin relaxation values are consistent with deconvolution

Ligand, protein and peak	T_2 value (mean and 95% confidence interval: ms)	Predicted FWHM (Hz) from T_2 value 95% CI $(1/\pi*T_2)$	Deconvolution FWHM (Hz)*	T_1 value (mean and 95% confidence interval: ms)
Pioglitazone PPARγ^{C313A,K502C}-BTFA -84.06 ppm	4.9 (4.6 to 5.3)	47 to 54	95	365 (348 to 382)
Pioglitazone PPARγ^{C313,K502CA}-BTFA -84.33 ppm	6.2 (5.7 to 6.7)	38 to 46	62	367 (356 to 379)
GW1929 PPARγ^{C313A,K502C}-BTFA -84.38 ppm	17.6 (16.9 to 18.4)	16 to 18	22	342 (339 to 344)
Apo PPARγ^{C313A,K502C}-BTFA -84.06 ppm	3.1 (2.9 to 3.4)	94 to 110	112	404 (360 to 451)
Apo PPARγ^{C313A,K502C}-BTFA -83.9 ppm	2.2 (2.0 to 2.3)	135 to 152	174	415 (396 to 434)

*Decovoluted peak widths are all slightly wider than the T_2 predicted linewidth — if they were narrower, this would indicate over-fitting of the data. Thus, the deconvolution provides a fit with the least number of possible peaks (i.e. the number of chemical environments, or conformations) sampled by the BTFA probe attached to helix 12 for these spectra.

Supplementary Methods

Parameters used in simulations

&cntrl

 imin = 0, nstlim = 250000000, dt=0.004,

 ntx = 5, irest = 1,

 ntwx = 25000, ioutfm = 1, ntxo =2, ntpr = 25000, ntwr = 25000,

 iwrap = 1, nscm = 1000,

 ntc = 2, ntf = 2, ntb = 2, cut = 8.0,

 ntt = 3, ig = -1, gamma_ln = 3, temp0 = 310, tempi = 310,

 ntp = 1, taup = 2.0, barostat = 2,

 igb = 0, saltcon = 0.0,

 ntr = 0,

Prepin file

   0   0   2

This is a remark line

molecule.res

CYB INT 0

CORRECT   OMIT DU BEG

 0.0000

 1 DUMM DU M   0 -1 -2   0.000   .0   .0   .00000

 2 DUMM DU M   1 0 -1   1.449   .0   .0   .00000

 3 DUMM DU M   2 1 0   1.523 111.21   .0   .00000

```
 4 N    N    M   3  2  1    1.540  111.208 -180.000 -0.415700
 5 H    H    E   4  3  2    1.001  128.253   75.116  0.271900
 6 CA   CT   M   4  3  2    1.453   46.684   -6.789  0.021300
 7 HA   H1   E   6  4  3    1.085  106.975  -51.989  0.112400
 8 CB   CT   3   6  4  3    1.532  114.777   70.403 -0.052400
 9 HB2  H1   E   8  6  4    1.083  111.004  -48.445  0.125000
10 HB3  H1   E   8  6  4    1.080  109.153   70.321  0.125000
11 SG   S    S   8  6  4    1.820  113.917 -171.908 -0.275000
12 CD   CT   3  11  8  6    1.812   99.158   84.434 -0.125700
13 HD2  H1   E  12 11  8    1.085  111.557  -58.090  0.067200
14 HD3  H1   E  12 11  8    1.085  110.619   60.963  0.067200
15 CE   C    B  12 11  8    1.510  111.238 -178.762  0.509200
16 O1   O    E  15 12 11    1.181  125.861   -0.134 -0.397100
17 CZ   CT   3  15 12 11    1.537  114.766  179.713  0.511800
18 F3   F    E  17 15 12    1.323  109.977   58.455 -0.191500
19 F1   F    E  17 15 12    1.305  111.986  179.382 -0.191500
20 F2   F    E  17 15 12    1.324  109.860  -59.791 -0.191500
21 C    C    M   6  4  3    1.534  104.761 -167.409  0.597300
22 O    O    E  21  6  4    1.192  120.251  -90.053 -0.567900
```

LOOP

IMPROPER

 CD CZ CE O1

 CA +M   C   O

DONE

STOP

AMBER parameter database derived frcmod file

Remark line goes here

MASS

N 14.010   0.530

H 1.008   0.161

CT 12.010   0.878

H1 1.008   0.135

S 32.060   2.900

C 12.010   0.616

O 16.000   0.434

F 19.000   0.320

BOND

H -N   434.00 1.010

CT-N   337.00 1.449

CT-H1 340.00 1.090

CT-CT 310.00 1.526

C -CT 317.00 1.522

CT-S   227.00 1.810

C -O   570.00 1.229

CT-F   367.00 1.380

ANGLE

H1-CT-N   50.000   109.500

CT-CT-N   80.000   109.700

C -CT-N   63.000   110.100

CT-N -H   50.000   118.040

CT-CT-H1 50.000   109.500

CT-CT-S   50.000   114.700

CT-C -O   80.000   120.400

C -CT-H1 50.000   109.500

C -CT-CT 63.000   111.100

CT-S -CT 62.000   98.900

H1-CT-H1 35.000   109.500

H1-CT-S   50.000   109.500

CT-C -CT 63.000   117.000

F -CT-F   77.000   109.100

DIHE

H1-CT-CT-N   9   1.400   0.000   3.000

N -CT-CT-S   9   1.400   0.000   3.000

O -C -CT-N   6   0.000   0.000   2.000

H1-CT-N -H   6   0.000   0.000   2.000

CT-CT-N -H   6   0.000   0.000   2.000

C -CT-N -H   6   0.000   0.000   2.000

CT-CT-S -CT 3   1.000   0.000   3.000

H1-CT-CT-H1 9   1.400   0.000   3.000

H1-CT-CT-S   9   1.400   0.000   3.000

O -C -CT-H1 1   0.800   0.000   -1.000

O -C -CT-H1 1   0.000   0.000   -2.000

O -C -CT-H1   1   0.080   180.000   3.000

O -C -CT-CT   6   0.000   0.000   2.000

H1-CT-S -CT   3   1.000   0.000   3.000

C -CT-S -CT   3   1.000   0.000   3.000

C -CT-CT-H1   9   1.400   0.000   3.000

C -CT-CT-S   9   1.400   0.000   3.000

O -C -CT-S   6   0.000   0.000   2.000

CT-C -CT-S   6   0.000   0.000   2.000

CT-C -CT-F   6   0.000   0.000   2.000

CT-C -CT-H1   6   0.000   0.000   2.000

O -C -CT-F   6   0.000   0.000   2.000

IMPROPER

CT-CT-C -O   10.5   180.0   2.0   Using general improper torsional
angle X- X- C- O, penalty score= 6.0)

NONBON

 N   1.8240 0.1700

 H   0.6000 0.0157

 CT   1.9080 0.1094

 H1   1.3870 0.0157

 S   2.0000 0.2500

 C   1.9080 0.0860

 O   1.6612 0.2100

 F   1.7500 0.0610

GAFF parameter database derived frcmod file

Remark line goes here

MASS

N 14.010   0.530   same as n

H 1.008   0.161   same as hn

CT 12.010   0.878   same as c3

H1 1.008   0.135   same as h1

S 32.060   2.900   same as ss

C 12.010   0.616   same as c

O 16.000   0.434   same as o

F 19.000   0.320   same as f

BOND

H -N   403.20   1.013   same as hn- n, penalty score= 0.0

CT-N   328.70   1.462   same as c3- n, penalty score= 0.0

CT-H1 330.60   1.097   same as c3-h1, penalty score= 0.0

CT-CT 300.90   1.538   same as c3-c3, penalty score= 0.0

C -CT 313.00   1.524   same as c-c3, penalty score= 0.0

CT-S   215.90   1.839   same as c3-ss, penalty score= 0.0

C -O   637.70   1.218   same as c- o, penalty score= 0.0

CT- F 356.90   1.350   same as c3- f, penalty score= 0.0

ANGLE

H1-CT-N   49.840   108.880   same as h1-c3-n , penalty score= 0.0

CT-CT-N   65.910   111.610   same as c3-c3-n , penalty score= 0.0

C -CT-N   67.000   109.060   same as c -c3-n , penalty score= 0.0

CT-N -H   45.800   117.680   same as c3-n -hn, penalty score= 0.0

CT-CT-H1   46.390   109.560   same as c3-c3-h1, penalty score= 0.0

CT-CT-S   61.300   110.270   same as c3-c3-ss, penalty score= 0.0

CT-C -O   67.400   123.200   same as c3-c -o , penalty score= 0.0

C -CT-H1   47.040   108.220   same as c -c3-h1, penalty score= 0.0

C -CT-CT   63.270   111.040   same as c -c3-c3, penalty score= 0.0

CT-S -CT   60.240   99.240   same as c3-ss-c3, penalty score= 0.0

H1-CT-H1   39.240   108.460   same as h1-c3-h1, penalty score= 0.0

H1-CT-S   42.060   108.760   same as h1-c3-ss, penalty score= 0.0

C -CT-S   61.850   108.840   same as c -c3-ss, penalty score= 0.0

CT-C -CT   62.040   116.500   same as c3-c -c3, penalty score= 0.0

C -CT-F   66.260   110.000   same as c -c3-f , penalty score= 0.0

F -CT-F   70.890   107.360   same as f -c3-f , penalty score= 0.0

DIHE

H1-CT-CT-N   9   1.400   0.000   3.000   same as X -c3-c3-X , penalty
score= 0.0

N -CT-CT-S   9   1.400   0.000   3.000   same as X -c3-c3-X , penalty
score= 0.0

O -C -CT-N   6   0.000   180.000   2.000   same as X -c -c3-X , penalty
score= 0.0

H1-CT-N -H   6   0.000   0.000   2.000   same as X -c3-n -X , penalty
score= 0.0

CT-CT-N -H   6   0.000   0.000   2.000   same as X -c3-n -X , penalty
score= 0.0

C -CT-N -H   6   0.000   0.000   2.000   same as X -c3-n -X , penalty
score= 0.0

CT-CT-S -CT   3   1.000   0.000   3.000   same as X -c3-ss-X , penalty
score= 0.0

H1-CT-CT-H1 9   1.400   0.000   3.000   same as X -c3-c3-X , penalty
score= 0.0

H1-CT-CT-S   9   1.400   0.000   3.000   same as X -c3-c3-X , penalty score= 0.0

O -C -CT-H1   1   0.800   0.000   -1.000   same as h1-c3-c -o

O -C -CT-H1   1   0.000   0.000   -2.000   same as h1-c3-c -o

O -C -CT-H1   1   0.080   180.000   3.000   same as h1-c3-c -o , penalty score= 0.0

O -C -CT-CT   6   0.000   180.000   2.000   same as X -c -c3-X , penalty score= 0.0

H1-CT-S -CT   3   1.000   0.000   3.000   same as X -c3-ss-X , penalty score= 0.0

C -CT-S -CT   3   1.000   0.000   3.000   same as X -c3-ss-X , penalty score= 0.0

C -CT-CT-H1   9   1.400   0.000   3.000   same as X -c3-c3-X , penalty score= 0.0

C -CT-CT-S   9   1.400   0.000   3.000   same as X -c3-c3-X , penalty score= 0.0

O -C -CT-S   6   0.000   180.000   2.000   same as X -c -c3-X , penalty score= 0.0

CT-C -CT-S   6   0.000   180.000   2.000   same as X -c -c3-X , penalty score= 0.0

CT-C -CT-F   6   0.000   180.000   2.000   same as X -c -c3-X , penalty score= 0.0

CT-C -CT-H1   6   0.000   180.000   2.000   same as X -c -c3-X , penalty score= 0.0

O -C -CT-F   6   0.000   180.000   2.000   same as X -c -c3-X , penalty score= 0.0

IMPROPER

CT-CT-C -O   1.1   180.0   2.0   Using the default value

NONBON

N	1.8240	0.1700	same as n
H	0.6000	0.0157	same as hn
CT	1.9080	0.1094	same as c3
H1	1.3870	0.0157	same as h1
S	2.0000	0.2500	same as ss
C	1.9080	0.0860	same as c
O	1.6612	0.2100	same as o
F	1.7500	0.0610	same as f

7.2 Supplementary material for chapter 3

Data in this section is in final preparation for publication.

Supplementary information for

Helix 4 determines the specificity of coactivator interactions in nuclear receptors

Ian M Chrisman, Zahra Heidari, Trey Patton, Tung-Chung Mou, Travis S Hughes
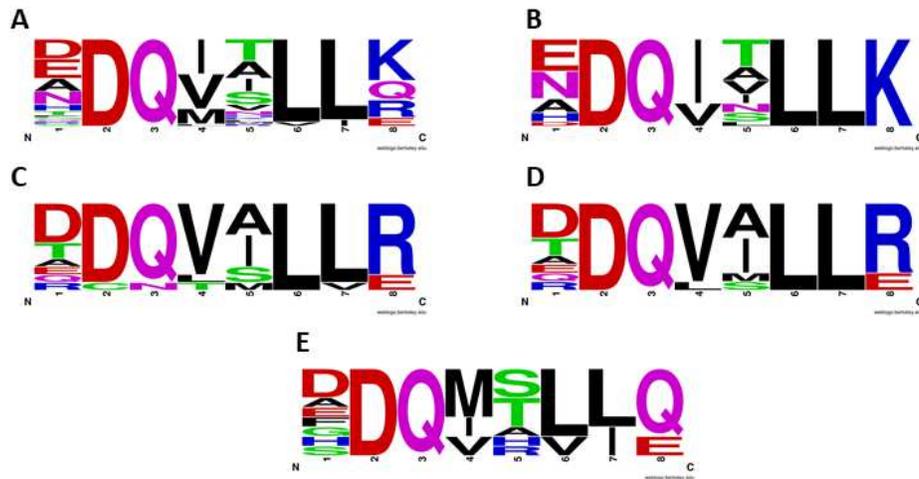
Figure S1. Consensus logos of the helix 4 motif in nuclear receptors. A: The consensus logo of the 8-residue motif on helix 4 for all 40 NR which are thought to bind coregulators on the AF-2 surface. In one NR, DAX1, this motif has expanded to be 9-residues and cannot be aligned. Of the remaining seven the motif cannot be located in two and is degenerate in five others. All NR where the motif is degenerate are not thought to bind coregulators on the AF-2 surface. B: Consensus logo of the helix 4 motif for the NR1 family of NR. This shows a much higher degree of conservation in this family relative to other NR families. C: Consensus logo of the helix 4 motif for the NR2 family. D: Consensus logo of the helix 4 motif for the NR2 family with TR4 omitted. This protein has a highly degenerate sequence in this region and has been shown not to bind coregulators on the AF-2 surface. E: Consensus logo of the helix 4 motif for the NR3 family. This motif is the least conserved in this family which may indicate the least homology in coregulator interactions within an NR family. This is merely hypothetical however and would require further research to demonstrate.

Protein	Sequence	Location	Co-peptide	PDB accession
Androgen receptor (NR3C4)	**D**D**Q**MAVI**Q** also **R**	731-738/726	NCoA2, NCoA4, NCoA3	2Q7J, 3L3X, 4OLM, 1T63
Constitutive androstane receptor (NR1I3)	EDQISLL**K** also **R**	188-195/183	NCoA1	1XV9
Coup TF1 (NR2F1)	TDQVSLLR	246-253	No structures	
Coup TF2 (NR2F2)	TDQVSLLR	239-246	No coregulator	
ERBAL2 (NR2F6)	ADQVALLR	229-236	No structures	
DAX1 (NR0B1)	LDQQLVLVR	280-288	No coregulators	
Estrogen receptor α (NR3A1)	HD**Q**VHLL**E**	373-380	NCoA1, NCoA2	4TUZ, 5KR9
Estrogen receptor β (NR3A2)	FD**Q**VRLL**E**	327-332	NCoA1	3OMO
Estrogen related receptor α (NR3B1)	SD**Q**MSVLQ	255-262	PGC1α	3D24
Estrogen related receptor β (NR3B2)	GDQMSLLQ	270-277	No structures	
Estrogen related receptor γ (NR3B3)	ADQMSLLQ	295-302	NCoA1, NRIP1, SMRT- **No interactions**	
Farnesoid x Receptor (NR1H4)	**HE**D**Q**IALL**K**	318-325	NCoA2, NCoA1	5WZX/5Q0A/ 3OOF
GCNF (NR6A1)			No structures	
Glucocorticoid Receptor (NR1C1)	**D**D**Q**MTLL**Q** also **R**	590-597/585	NCoA2, NCoR	3MNR, 3H52, 4CSJ, 3K22, 3E7C
HNFα (NR2A1)	**D**D**Q**VALL**R**	205-212	NCoA1	1PZL
HNFγ (NR2A2)	DDQVALLR	166-173	No coregulators	
LRH1 (NR5A2)	**D**D**Q**MKLL**Q**	372-379	Dax1, SRC2	4RWV, 3PLZ
LXRα (NR1H3)	**E**D**Q**IALL**K**	284-291	NCoA1	3IPQ
LXRβ (NR1H2)	**RE**D**Q**IALL**K**	298-305	NCoA1	4DK7
Mineralocorticoid receptor (NR3C2)	ED**Q**ITLIQ	796-803	NCoA1	5L7E
NOR1 (NR4A3)	EDQTLLIE	464-471	No structures	
NURR1 (NR4A1)	ADQDLLFE	433-440	No coregulators	
NUR77 (NR4A2)	ADQDLLLE	433-440	No coregulators	
PGR (NR3C3)	**D**DQITLI**Q**	745-752	SMRT	4OAR
PNR (NR2E3)	RDQVILLE	247-254	No coregulators	
PPARα (NR1C1)	**N**DQVTLL**K**	303-310	NCoA1	3FEI, 3G8I
PPARγ (NR1C3)	**N**DQVTLL**K**	312-319	MED1	
PPARδ (NR1C2)	NDQVTLLK	276-283	No coregulators	
PXR (NR1I2)	**E**DQISLL**K**	270-277	NCoA1	5X0R, 2O9I
RARα (NR1B1)	ADQITLLK	255-262	SMRT, NCoA1- **No interactions**	
RARβ (NR1B2)	ADQITLL**K** also **T**	258-264/253	NCoA1	4DM6, 4JYI
RARγ (NR1B3)	ADQITLLK	257-264	No coregulators	

REV-ERBα (NR1D1)	HDQVNLLK	466-473	NCoR- **No interactions**	
REV-ERBβ (NR1D2)	HDQVNLLK	432-439	No coregulators	
RORα (NR1F1)	**Q**NDQIVLLK	349-357	NRIP1	4S15
RORβ (NR1F2)	**Q**NDQILLLK	288-296	NCoA1	1NQ7
RORγ (NR1F2)	**Q**NDQIVLL**K**	346-354	NCoA2, NRIP1, SMRT	5YP5, 3L0L, 6FZU, 5NTI, 5X8Q
RxRα (NR2B1)	DDQVILL**R**	295-302	NCoA2, TIF2	4RMC, 5MK4
RXRβ (NR2B2)	DDQVILLR	366-373	NCoA2- **No interactions**	
RXRγ (NR2B3)	EDQVILLR	296-303	No coregulators	
SF1 (NR5A1)	ADQMTLL**Q**	293-299	NCoA2, SHP, PGC1α	1ZDT
SHP (NR0B2)	QDQRRLLQ	82-89	No coregulators	
THRα (NR1A1)	EDQIILLK	245-252	No coregulators	
THRβ (NR1A2)	EDQIILLK	299-306	No coregulators	
TLX (NR2C2)	QDQLMLLE	202-209	No coregulators	
TR2 (NR2C1)			No structures	
TR4 (NR2C2)	DCNTSLVR	413-420	No coregulators	
VDR (NR1I1)	**D**DQIVLL**K**	253-260	MED1	5AWK, 5B41

Figure S2. Conserved helix 4 motif in all nuclear receptors. For all human nuclear receptors, listed alphabetically in left column, the eight-residue helix 4 sequence is shown along with the residue numbers this motif covers. Bolded residues are those which are resolved to interact with coregulator peptides in crystal structures. The coregulators listed are those which show interaction and the PDB accession numbers for the crystal structures these interactions were identified in are provided. For those proteins which have been crystallized with coregulators but show no helix 4 interactions (ERRγ, RARα, REV-ERBα and RXRβ) all coregulators that have been crystallized with this protein are listed. In one case, DAX1, this motif has expanded to include one extra residue. For the proteins GCNF and TR2 this motif is highly degenerate and cannot be located due to the lack of crystal structures.

Species	PPARγ	Estrogen Receptor α	Glucocorticoid Receptor	Testicular Receptor 4	Nurr1
Homo Sapiens (Human)	NDQVTLLK	HDQVHLLE	DDQMTLLQ	DCNTSLVR	ADQDLLFE
Sus Scrofa (Pig)	NDQVTLLK (0.98)	HDQVHLLE (0.92)	DDQMTLLQ (0.89)	DCNTSLVR (0.99)	ADQDLLFE (0.99)
Manis Javanica (Pangolin)	NDQVTLLK (0.98)	HDQVHLLE (0.91)	DDQMTLLQ (0.95)	DCNTSLVR (0.98)	ADQDLLFE (0.99)
Physeter Catodon (Sperm Whale)	NDQVTLLK (0.98)	HDQVHLLE (0.89)	DDQMTLLQ (0.93)	DCNTSLVR (0.99)	ADQDLLFE (0.99)
Gallus Gallus (Chicken)	NDQVTLLK (0.87)	HDQVHLLE (0.79)	DDQMTLLQ (0.75)	ECNTSLVR (0.93)	TDQDLLFE (0.95)
Danio Rerio (Zebrafish)	NDQVTLLK (0.64)	HDQVQLLE (0.58)	DDQMTLLQ (0.49)	ECNTALVR (0.80)	CDQELLFE (0.92)
Xenopus Laevis (Frog)	NDQVTLLK (0.82)	HDQVHLLE (0.69)	DDQMTLLQ (0.63)	DCNTNLVR (0.88)	QDQDLLFE (0.87)

Figure S3. Sequence conservation of the helix 4 motif. The sequence conservation of the 8-residue helix 4 motif for several nuclear receptors was examined across several species which range from close relatives of humans to distantly related species. Total protein identity is shown in parentheses for each non-human sequence. For all nuclear receptors shown this motif is fully conserved among mammals. The only sequences found to not be fully conserved are from proteins which do not meet the standard motif. Fully conserved residues are shown in bold while residues which deviate from the human sequence are in red.
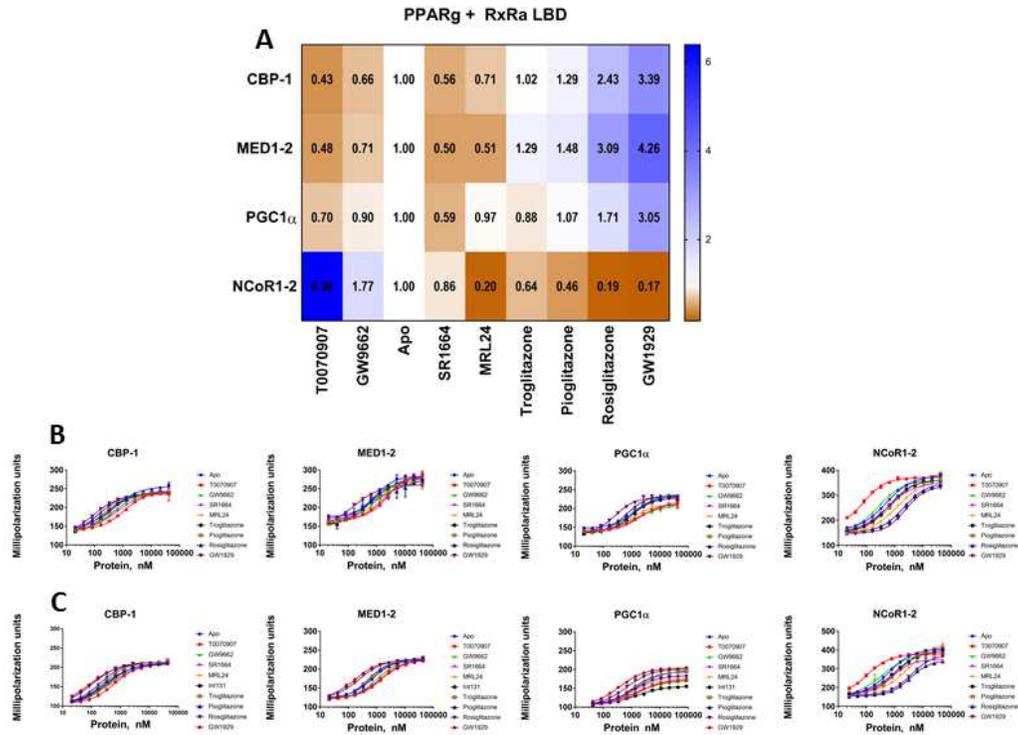
Figure S4. Experimental replicate of the LBD PPARγ-RxRα FP. A: Heatmap from an experimental replicate of the LBD PPARγ-RxRα. All values shown reflect the average value from two technical replicates within the single experimental replicate. B: Curve fits used to generate the heatmap shown in figure 2. C: Curve fits used to generate the heatmap shown in panel A. Error bars reflect the difference in two technical replicates within the same experimental replicate.

Figure S5. Experimental replicate of the FL PPARγ-RxRα-SULT2A1 FP. A: Heatmap from an experimental replicate of the FL PPARγ-RxRα-SULT2A1 PPRE. All values shown reflect the average value from two technical replicates within the single experimental replicate. B: Curve fits used to generate the heatmap shown in figure 2. C: Curve fits used to generate the heatmap shown in panel A. Error bars reflect the difference in two technical replicates within the same experimental replicate.
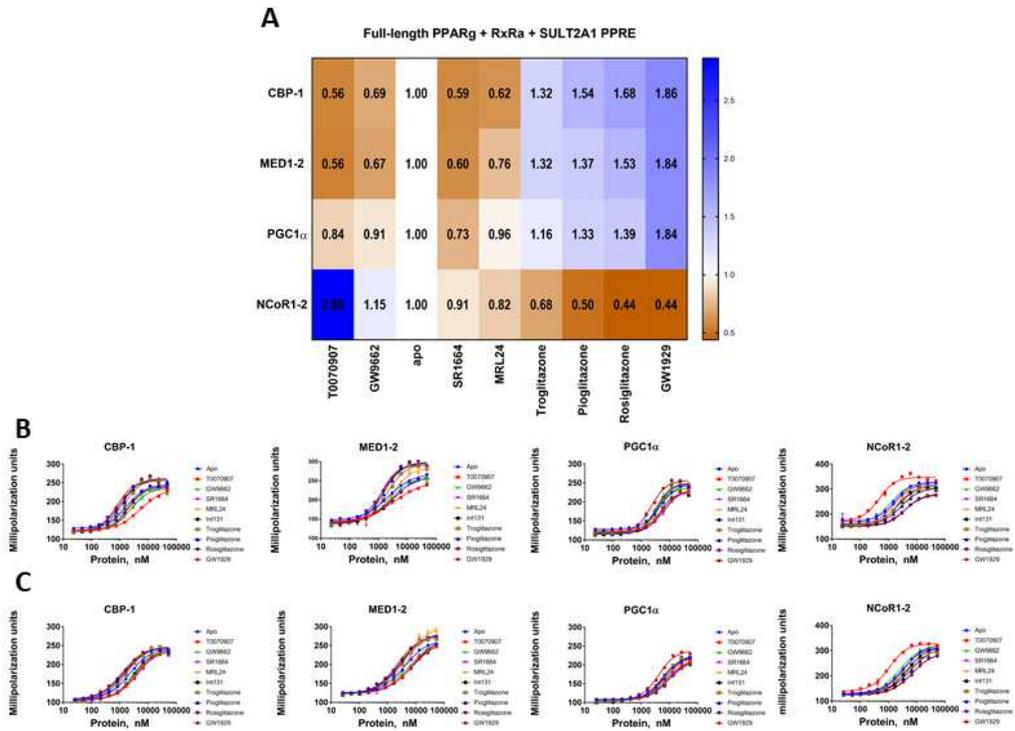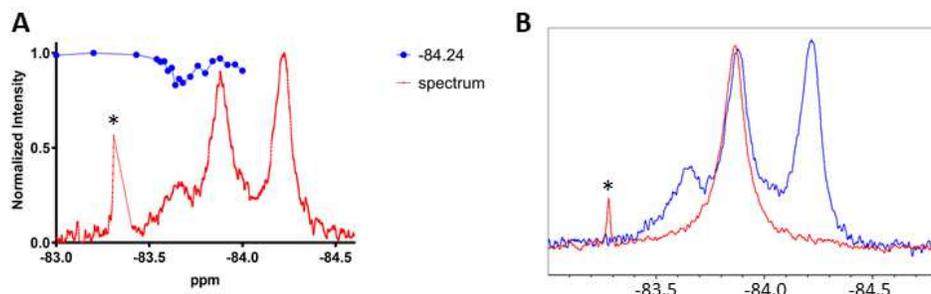
Figure S6. CEST NMR indicates that the central peak observed in ^{19}F NMR of PPARγ^{Q294C}-BTFA arises from labeling of the native cysteine. A: Variable location CEST NMR of PPARγ^{Q294C}-BTFA reveals that in this variant that the rightmost peak is not in exchange with the central peak but does exchange with the leftmost. In this spectrum the peak for free BTFA, denoted with an asterisk, was truncated so as to not drown out the peaks from the protein. This truncation is why the observed peak is much wider than expected for a small molecule. B: Overlay of PPARγ^{Q294C}-BTFA (Blue) and a sample WT PPARγ-LBD labeled with a 40x molar excess of BTFA (Red). The WT PPARγ peak nearly perfectly overlays with the central peak of PPAγ^{Q294C}-BTFA supporting that this peak arises from labeling of C285. In all spectra an * denotes residual free BTFA.
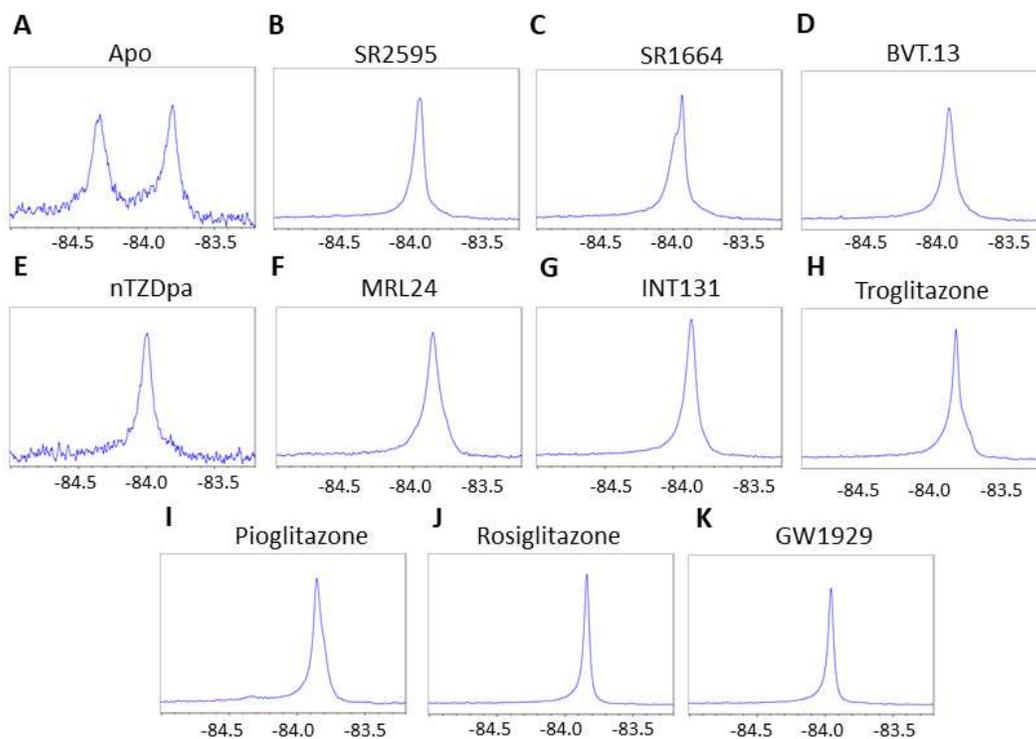
Figure S7. ^{19}F NMR of PPARγ$^{C285A, Q294C}$-BTFA. A-L: The ^{19}F NMR spectra of PPARγ$^{C285A, Q294C}$-BTFA in 12 distinct ligand bound states. The covalent ligands T0070907 and GW9662 were not run with this mutant as they require C285 for function. All NMR spectra were collected once.

215
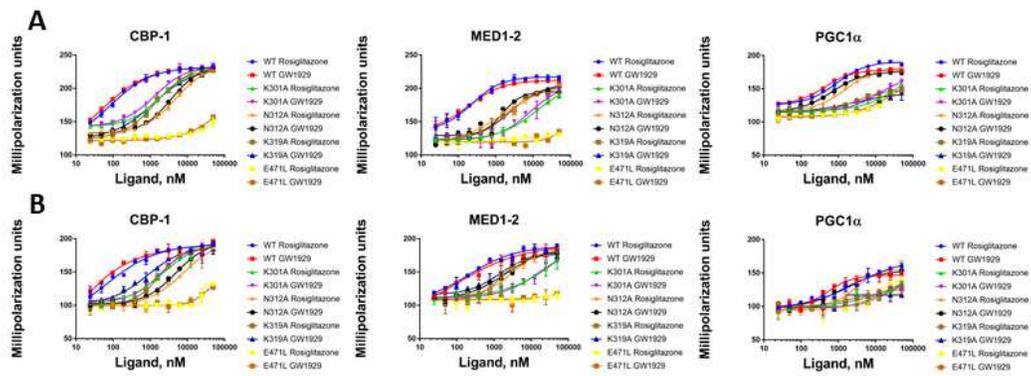
Figure S8. Curve fits of Mutant FP. A: The curve fits for FP of point mutants from a single experimental replicate. B: The curve fits for FP of point mutants from the second experimental replicate. Error bars signify the difference in two technical replicates within an experimental replicate.
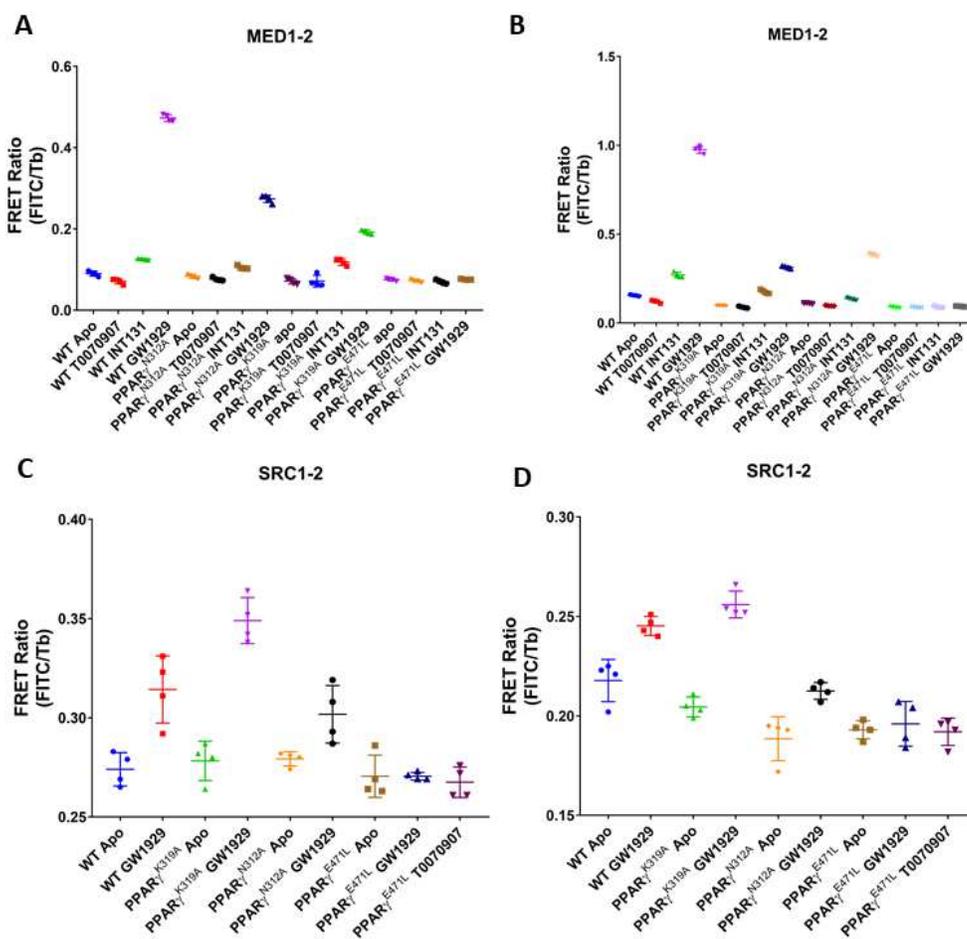
Figure S9. TR-FRET evidences the same ligand behavior as is seen in FP data. A: TR-FRET data for MED1-2 association with more ligands including a partial agonist, INT131, and an inverse agonist T0070907. This data shows the same pattern of ligand induced affinity as is seen in the FP assays with this peptide. Data in this panel is from the same experiment as is shown in figure 4. B: Experimental replicate of the MED1-2 TR-FRET data. Data in this experimental replicate was collected in a manner identical to that in panel A. C: Experimental replicate of the SRC1-2 TR-FRET shown in figure 4. Data in this panel was collected identically to that in figure 4. D: Experimental replicate of the SRC1-2 TR-FRET shown in figure 4. Data in this experimental replicate was collected at a concentration of SRC1-2 half that used in figure 4. The pattern is identical to that seen at the higher concentration but the signal to noise is significantly worse.

217
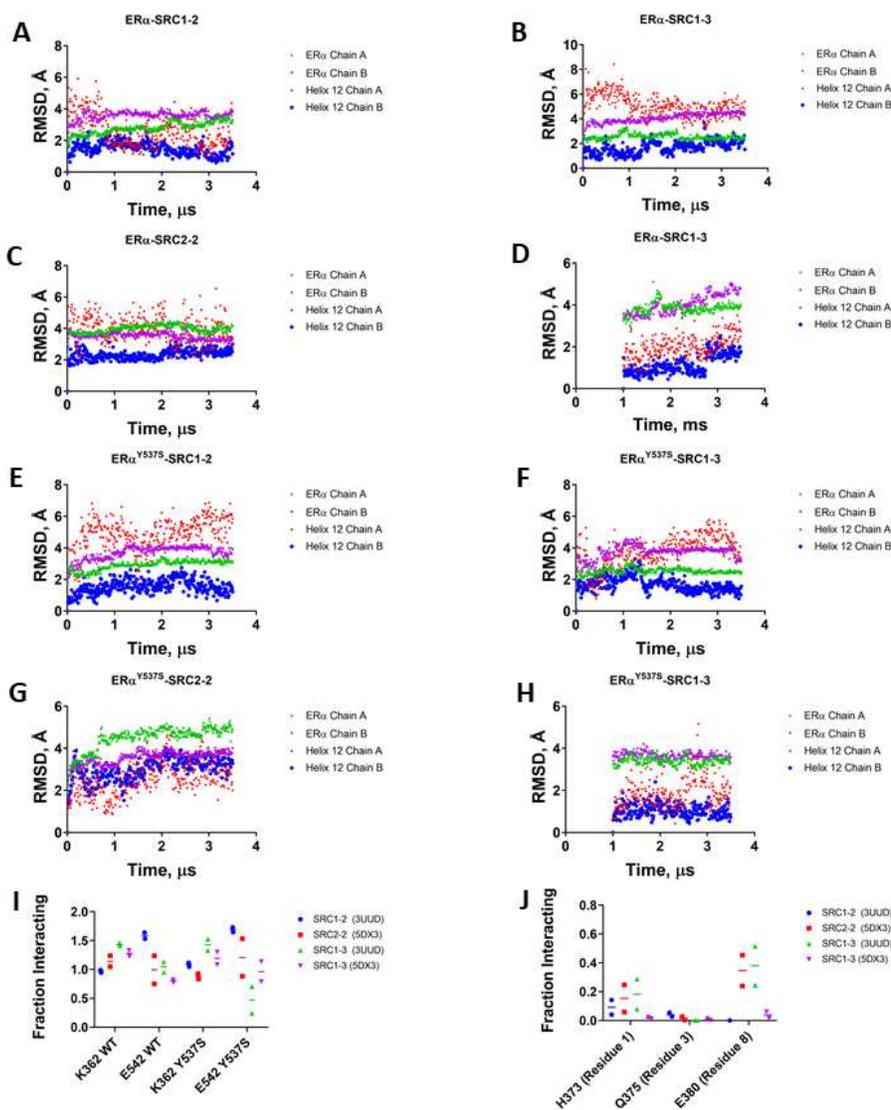
Figure S10. Molecular dynamics of ERα. A: RMSD of ERα bound to SRC1-2 started from 3UUD. B: : RMSD of ERα bound to SRC1-3 introduced to 3UUD. C: : RMSD of ERα bound to SRC202 started from 5DX3. D: : RMSD of ERα bound to SRC1-3 introduced to 5DX3. E: : RMSD of ERα^{Y537S} bound to SRC1-2 started from 3UUD. F: RMSD of ERα^{Y537S} bound to SRC1-2 introduced to 3UUD. G: RMSD of ERα^{Y537S} bound to SRC2-2 started from 3UUD. H: RMSD of ERα^{Y537S} bound to SRC1-3 introduced to 5DX3. In panel H the first μs of data is omitted due to a saving error. To allow protein to reach a stable state only the data from 1.5-3.5μs was considered in this work. I: Fractional charge clamp interaction for all data in this work. In the WT ERα interaction is always between 100-200% indicating robust interaction and proper peptide location. In ERα^{Y537S} the interaction is not always well populated indicating there may be issues in the peptide location. J: Fraction of interaction between the helix 4 residues and peptides in ERα^{Y537S}.

218

A

SR2595

B

SR1664

C

nTZDpa

D

BVT.13

E

MRL-24

F

INT-131

G

Troglitazone

H

Pioglitazone

I

Rosiglitazone

J

GW1929

K

3-bromo-1,1,1-trifluoroacetone
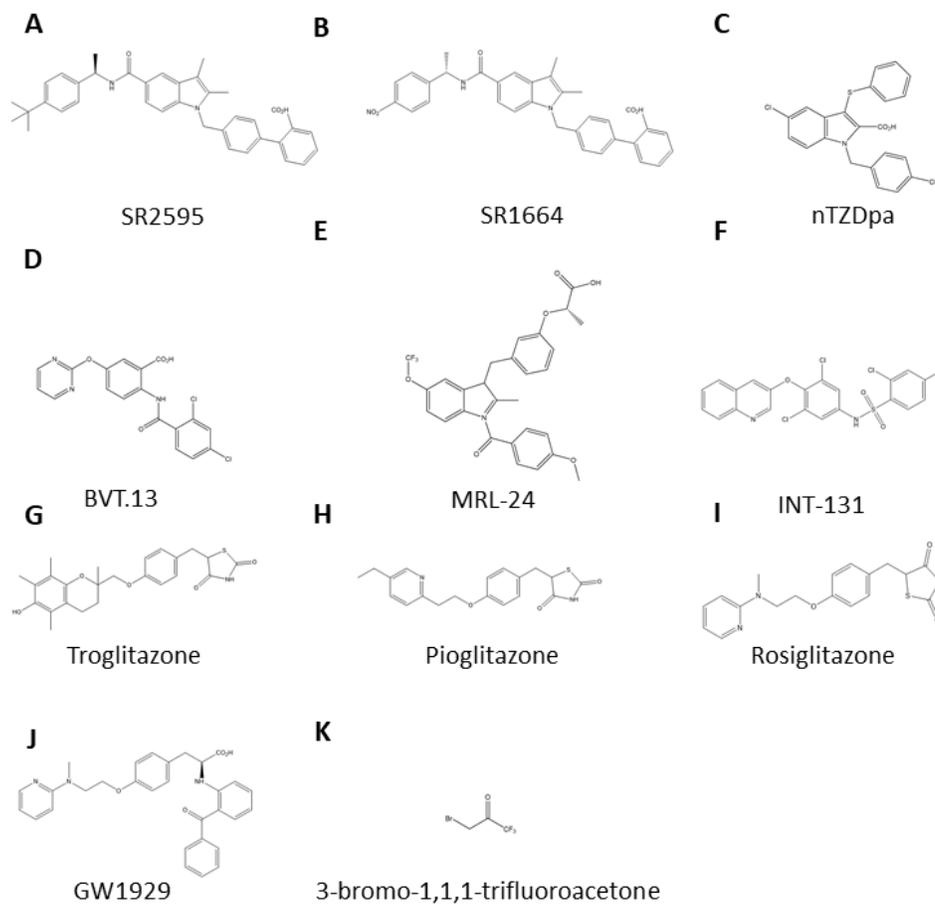
Figure S11. Chemical structures of the synthetic ligands and the ^{19}F label used in this work.
A-M: The chemical structures of all PPARγ ligands used in this work. N: The chemical structure
of the BTFA molecule used to add the ^{19}F label to PPARγ.

	T0070907	GW9662	Apo	SR1664	MRL24	Troglitazone	Pioglitazone	Rosiglitazone	GW1929
CBP-1	1,118	520.7	414.5	601.7	381.2	280.8	182.2	131	81.8
	602.6	387.7	257.7	459.9	364.5	252.9	199.7	105.9	76.1
MED1-2	2,953	1,492	861.9	2,093	1,814	641.4	462	347.5	145.5
	1,533	1,036	732.9	1,465	1,450	567.5	494.2	237.4	172
PGC1α	2,008	1,256	1,235	1,761	1,029	984	893.4	658.4	257.1
	1,586	1,230	1,110	1,868	1,149	1,264	1,042	648.9	363.6
NCoR-1	78.4	273.1	497.5	555.3	1,404	602.6	875.9	2087	2,374
	100	359.9	638.8	738.6	3,194	1,001	1,387	3,445	3,777

Table S1. K$_d$ values for the LBD PPARγ-RxRα complex with coactivator peptides. The first experimental replicate is shown in blue while a second experimental replicate is shown in orange. All K$_d$ values are in nanomolar and were obtained from non-linear curve fits of 12 data points ranging from 20-41600nM for the first experimental replicate or 21-43000nM for the second experimental replicate. Each curve was comprised of two technical replicates.

	T0070907	GW9662	Apo	SR1664	MRL24	Troglitazone	Pioglitazone	Rosiglitazone	GW1929
CBP-1	3,953	3,211	2,227	3,805	3,624	1,693	1,444	1,327	1,198
	5,431	2,385	1,791	2,579	2,020	1,375	1,134	1,161	1,015
MED1-2	7,240	6,032	4,020	6,753	5,271	3,049	2,928	2,623	2,181
	3,390	2,592	2,149	2,795	2,869	1,893	1,507	1,311	1,185
PGC1α	8,844	8,203	7,425	10,144	7,720	6,408	5,571	5,340	4,034
	6,075	3,474	4,601	6,471	3,680	4,061	3,553	3,336	2,205
NCoR-1	1,055	2,628	3,022	3,324	3,692	4,444	6,015	6,877	6,875
	455	1,161	1,552	1,541	2,183	1,946	2,741	3,710	3,785

Table S2. K_d values for the full-length PPARγ-RxRα-DNA complex with coactivator peptides. The first experimental replicate is shown in blue while a second experimental replicate is shown in orange. All K_d values are in nanomolar and were obtained from non-linear curve fits of 12 data points ranging from 26-53000nM for the first experimental replicate or 23-47600nM for the second experimental replicate. Each curve was comprised of two technical replicates.

	PPAR:GW1929	PPAR:MED1-2:GW1929
Data collection		
Beamline	APS 19BM	APS 19BM
Wavelength (Å)	0.979	0.979
Ligand	1929	MED1+1929
Resolution range (Å)*	30.31 - 1.90 (1.97 - 1.90)	43.81 - 1.90 (1.97 - 1.90)
Space group	P 41 21 2	C 2 2 21
Unit cell dimensions		
a, b, c (Å)	61.9 61.9 167.9	55.4 87.6 121.6
α, β, γ (∘)	90 90 90	90 90 90
Unique reflections*	26487 (2520)	23571 (2313)
Redundancy*	8.0 (4.3)	9.8 (6.2)
Completeness (%)*	99.22 (97.30)	99.26 (98.72)
Mean I/σ (I)*	13.5 (2.3)	19.2 (7.6)
Wilson B-factor	23.08	25.50
R_{sym}[†],*	0.11 (0.65)	0.09 (0.56)
Refinement		
R_{work}[§],*	0.2082 (0.3873)	0.2008 (0.3524)
R_{free}[§],*	0.2446 (0.4103)	0.2574 (0.3881)
Number of total atoms		
protein	2217	2295
1929	37	37
Solvent	185	131
Total protein residues	276	286
RMS deviations		
Bonds lengths (Å)	0.004	0.012
Bond angles (∘)	0.56	1.44
Ramachandran favored/allowed/outliers (%)[††]	97.81/ 2.19/0.00	98.23/1.77/0.00
Average B-factor		
Macromolecules	36.01	45.24
GW1929	44.01	30.09
water	38.05	45.26
Number of TLS groups	0	1

Table S3. Crystallography data collection and refinement statistics.

[*] Data for highest resolution shell are given in brackets. [†] $R_{Sym} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$, where $I_i(hkl)$ is the ith observation of the intensity of the reflection hkl. [§] $R_{work} = \sum_{hkl} ||F_{obs}| - |F_{calc}|| / \sum_{hkl} |F_{obs}|$, where F_{obs} and F_{calc} are the observed and calculated structure-factor amplitudes for each reflection hkl. R_{free} was calculated with 5% of the diffraction data that were selected randomly and excluded from refinement. [††] Calculated using MolProbity (Chen et al., 2010).

	CBP-1	MED1-2	PGC1α
WT PPARγ-LBD	82.12	226.80	884.00
Rosiglitazone	125.90	194.80	1966.0
WT PPARγ-LBD	48.71	180.80	413.7
GW1929	70.00	97.61	538.70
PPARγ^{K301A}	2,040	12,953	46,132
Rosiglitazone	2,135	10,894	>50,000
PPARγ^{K301A}	1,310	8,299	12,348
GW1929	1,853	8,177	12,310
PPARγ^{N312A}	4,787	1,674	1,590
Rosiglitazone	7,838	1,130	2,973
PPARγ^{N312A}	3,663	1,495	675
GW1929	4,886	1,127	1,306
PPARγ-LBD^{K319A}	1,625	2,332	2,946
Rosiglitazone	1,247	1649	4,249
PPARγ-LBD^{K319A}	839.5	2,407	1,484
GW1929	980.4	1,379	751.2
PPARγ-LBD^{E499L}	>50,000	NBD	>50,000
Rosiglitazone	>50,000	NBD	>50,000
PPARγ^{E499L}	>50,000	NBD	34,686
GW1929	>50,000	NBD	>50,000

Table S4. K$_d$ values for the PPARγ mutants with coactivator peptides. The first experimental replicate is shown in blue while a second experimental replicate is shown in orange. All K$_d$ values are in nanomolar and were obtained from non-linear curve fits of 12 data points ranging from 24-49,500nM for both experimental replicates. Each curve was comprised of two technical replicates. A value of >50,000 indicates the affinity was lower than the highest concentration used so an accurate affinity cannot be determined. A value of NBD indicates that there was no indication of binding at the highest concentration used.

7.3 Supplementary material for chapter 4

Data in this section is in final preparation for publication

Supplementary material for

Inverse agonists induce drug specific structure and function via distinct molecular mechanisms in the nuclear receptor PPARγ

Zahra Heidari, Ian Chrisman, Scott Novick, Michelle Nemetchek, Theodore Kamenecka, Patrick Griffin, Travis Hughes

Supplementary Table 1. Duration of aMD simulations

PPARγ ligand binding domain complex	Total length of simulations (μs)	number of independent runs
*apo-NCoR	11.9	5
*apo	56.25	7
active apo	12.75	6
*T0070907-NCoR	13.5	6
*T0070907	15.75	7
*GW9662-NCoR	10.8	5
*GW9662	8.1	6
*SR10221-NCoR	11.25	5
*SR10221	9	4
rosiglitazone	9	4

* started with helix 12 in an inactive conformation
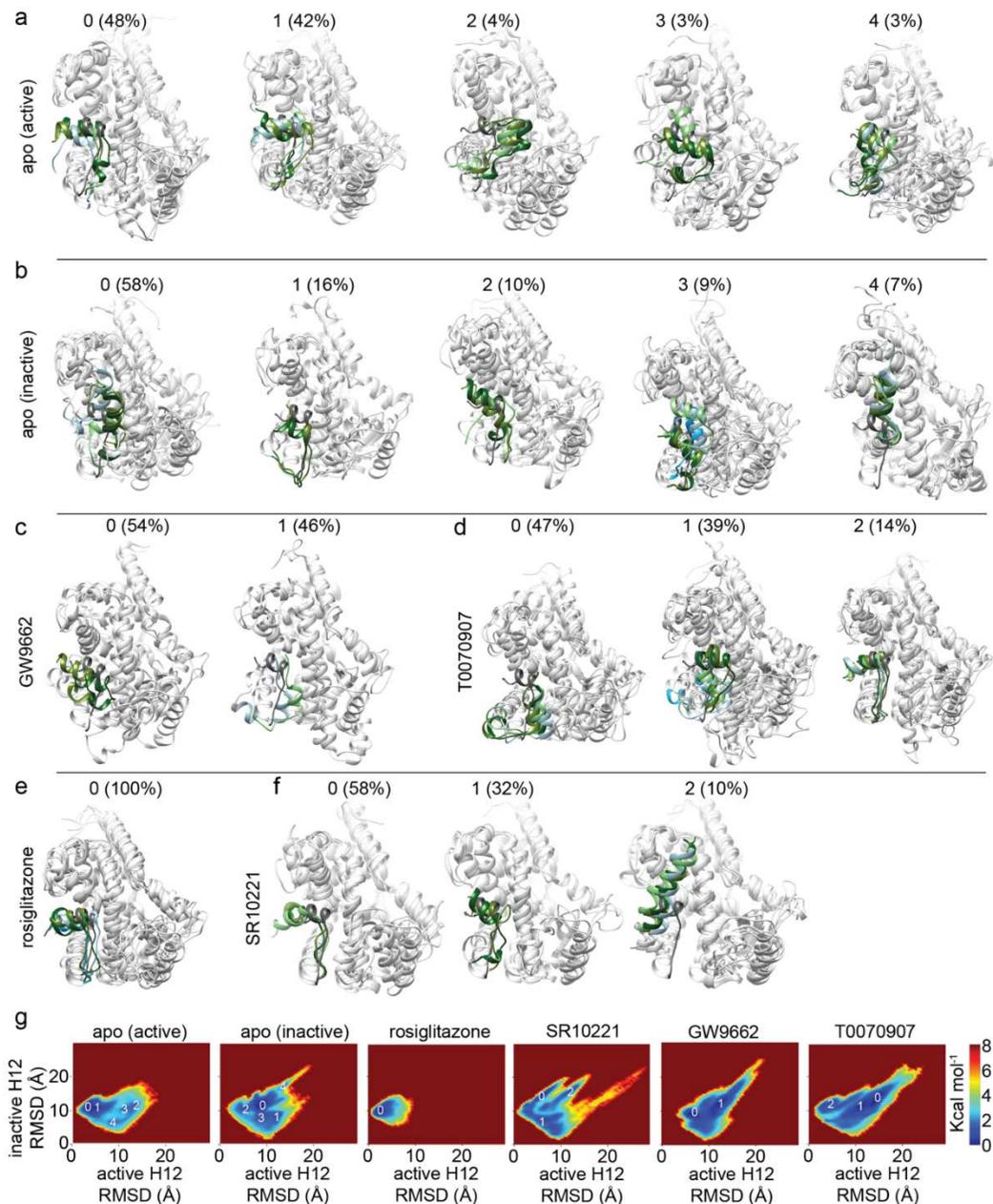
Supplementary Figure 1. Accelerated MD reveals a more diverse structural ensemble for apo PPARγ and PPARγ bound to inverse agonists/antagonists compared to agonist bound PPARγ. a-f) Representative structures from clustering based on helix 12 RMSD of structures in the wells indicated in panel g. The identity of the energy well and the calculated prevalence of each well's structures within the Boltzmann ensemble (see methods) is indicated by the number and percent in parentheses respectively. The prevalence of the structures within a particular well is indicated by the color of helix 12 ranging from dark green for the most prevalent to olive drab, light green, light blue and then deep sky blue for the least prevalent. For reference, an active

226

structure (rosiglitazone bound PPARγ; 2PRG chain A) is shown with helix 12 colored dim grey. g) The locations of the wells that were sampled to produce the representative structures in panels a-f is shown. Supplementary table 9 shows the relative energies of the wells indicated in panel g.

Supplementary Table 2. Comparison of coregulator peptide affinity for the ligand binding domain of wt PPARγ, PPARγ^{Q299C}-BTFA

Table Q299C probe effects

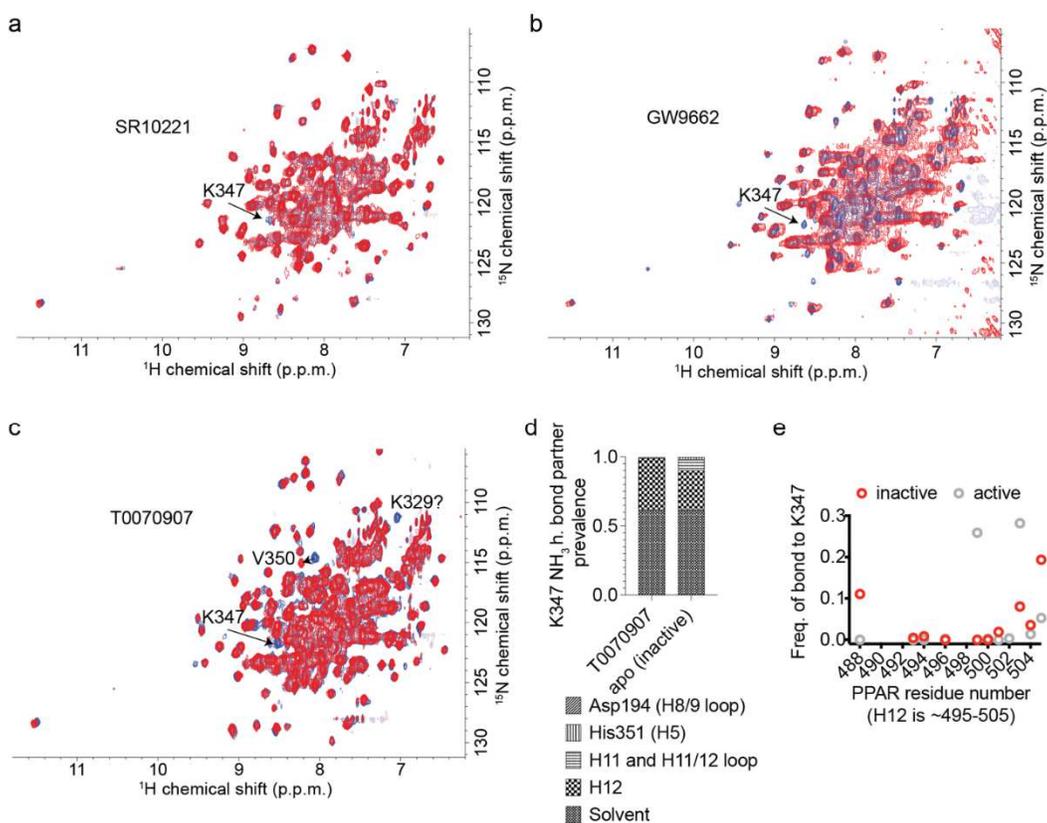| DRUG	WT K_d (nM)		PPARγ^{Q299C}-BTFA K_d (nM)			
	NCoR	MED1	NCoR	Fold Change vs WT	MED1	Fold Change vs WT
T007	105	20740	69	0.7	20440	1.0
GW9662	396	3296	151	0.4	4345	1.3
NO LIGAND (APO)	394	3640	497	1.3	6765	1.9
SR1664	377	2330	322	0.9	3722	1.6
MRL24	892	2664	767	0.9	2426	0.9
INT131	1181	2639	1183	1.0	2548	1.0
ROSIGLITAZONE	4433	542	2766	0.6	388	0.7
GW1929	3412	261	5517	1.6	328	1.3
			median	0.9	median	1.1
			average	0.9	average	1.2

**Supplementary Table 3. Comparison of coregulator peptide affinity for the ligand binding domain of wt PPARγ, PPARγ^{Q322C}-BTFA, and PPARγ^{C313A,Q322C}-BTFA.**

| Coregulator	MED1 K$_d$ (nM)					NCoR K$_d$ (nM)				
PPARγ Variant	Wt	PPARγ^{C313A,Q322C}-BTFA	FC	PPARγ^{Q322C}-BTFA	FC	Wt	PPARγ^{C313A,Q322C}-BTFA	FC	PPARγ^{Q322C}-BTFA	FC
Apo	7437	9999	0.74377438	3465	2.14632035	441.3	302.4	1.5	392.6	1.1
T0070907	19852	6310		6881	2.88504578	66.77	397.6	6	69.84	1
GW9662	4647	8033				240.2	397.6	1.7		
SR1664	3864	5834	0.66232431			378.9	353.7	1.1		
MRL24	2075	8178	0.25372952	1761	1.17830778	956.4	1348	1.4	385.5	2.5
Int131	1925	4047	0.47566098			1150	1322	1.1		
Troglitazone	1378	2401	0.57392753			535.2	313.2	1.7		
Pioglitazone	945.7	2426	0.38981863			625.8	361.2	1.7		
Rosiglitazone	311.4	1084	0.28726937	243.1	1.28095434	1604	639.8	2.5	6293	3.9
GW1929	210.5	819.2	0.25695801	201.7	1.04362915	4423	24392	5.5	7657	1.7
Median Change			2.3		1.7			1.3		1.7
Mean Change			0.45543284		1.70685148			1.7		2.1

Supplementary Figure 2. Protein and ^{19}F NMR of the helix 12 conformational ensemble. a-c) Comparison of K347A mutant, that would disrupt interaction between helix 4 and helix 12, and wt PPARγ LBD bound to the indicated ligands. Probable assignments, based on rosiglitazone bound PPARγ and apo PPARγ assignments, for residues affected by the mutation are indicated. The mutation consistently affects a resonance with shifts between 121 and 123 and 8.5 and 8.7 in all forms. Two rosiglitazone assigned residues are in this area, K347 and L463. L463 is far from the coregulator binding surface so we assume this resonance is K347 (the residue we mutated). d) The Boltzmann average values for interaction of the terminal side chain protons on K347 (i.e. NH$_3$) and all other residues was calculated for PPARγ alone (apo) or bound to T0070907. When K347 is not interacting with helix 12 it interacts almost solely with solvent. c) all 58 representative structures shown in Supplementary Figure 1 were subjectively classified as having active-like or inactive-like helix 12 conformations and prevalence of bonding between helix 12 residues and K347 were calculated from 1μs simulations of each representative structure.
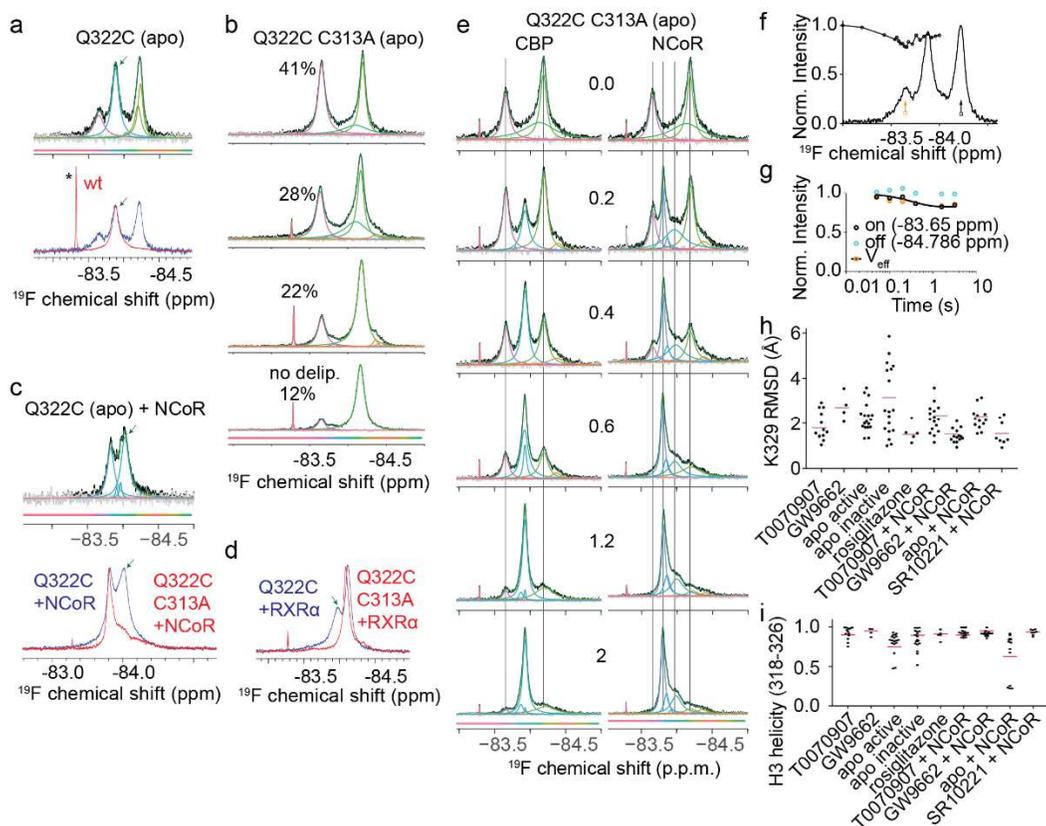
Supplementary Figure 3. The c-terminal portion of helix 3 is composed of a diverse structural ensemble in apo that consolidates upon binding coregulators, RXRα or ligands. a) The spectrum of wt PPARγ-BTFA (red) overlays well with the middle peak of PPARγ^{Q322C}-BTFA, indicating that the middle peak originates from labeling of the native cysteine (C313). The green arrow indicates signal from C313. b) Examples of spectra from delipidated PPARγ^{C313A,Q322C}-BTFA from different purifications and from non-delipidated PPARγ^{C313A,Q322C}-BTFA (bottom panel). The delipidation efficacy is variable and affects the relative population of apo peaks with delipidation increasing the population of the left peak. This variability is also observed with spectra of apo bound to peptides. These data indicate that the left peak in these spectra originates from apo. The right peak may also be representative of apo PPARγ or *E. coli* lipid bound PPARγ or the lipid may simply push the apo helix 3 structure equilibrium towards the right peak. c-d) Overlay of PPARγ^{Q322C}-BTFA (blue), which can be labeled at the native cysteine (C313), and PPARγ^{C313A,Q322C}-BTFA (red) which cannot, reveals the signal originating from C313 when PPARγ is bound to either c) NCoR or d) RXRα (green arrows). NCoR and RXRα binding consolidate the Q322C-BTFA signal into one main peak. e) Either CBP or NCoR was titrated into ligand free PPARγ^{C313A,Q322C}-BTFA (apo) at the indicated molar ratios of peptide to PPARγ starting with just PPARγ (0.0). Titration of both peptides again leads to peak consolidation. f) CEST indicates exchange between the peak denoted by the orange arrow and the black arrow, but not the central peak (which originates from the native cysteine). The black circles indicate the location of the saturating peak, while the height of the circles represent the height of the right peak labeled with the black arrow. g) Saturating NMR pulses were applied at the orange peak (in panel f) and an off-target area equidistant from the peak of interest (the right peak labeled with the black arrow in panel f) and the height of the right peak (panel f) was measured after various durations of the saturating NMR pulse. The data were then fit to a single exponential, which indicates

231

0.5 s^{-1} exchange (0.3-0.7 95% CI), however this is at the limit of detection and the exchange rate may be lower. h) Boltzmann averages of the RMSD of helix 3 charge clamp (K329) indicate higher variance in the position of K329 compared to when bound to inverse agonists/antagonists or agonists. i) Boltzmann averages of the helicity of a section of helix 3 near the charge clamp and containing the Q322C BTFA probe shows higher variance in the helicity in apo PPARγ compared to when bound to inverse agonists/antagonists or agonists.
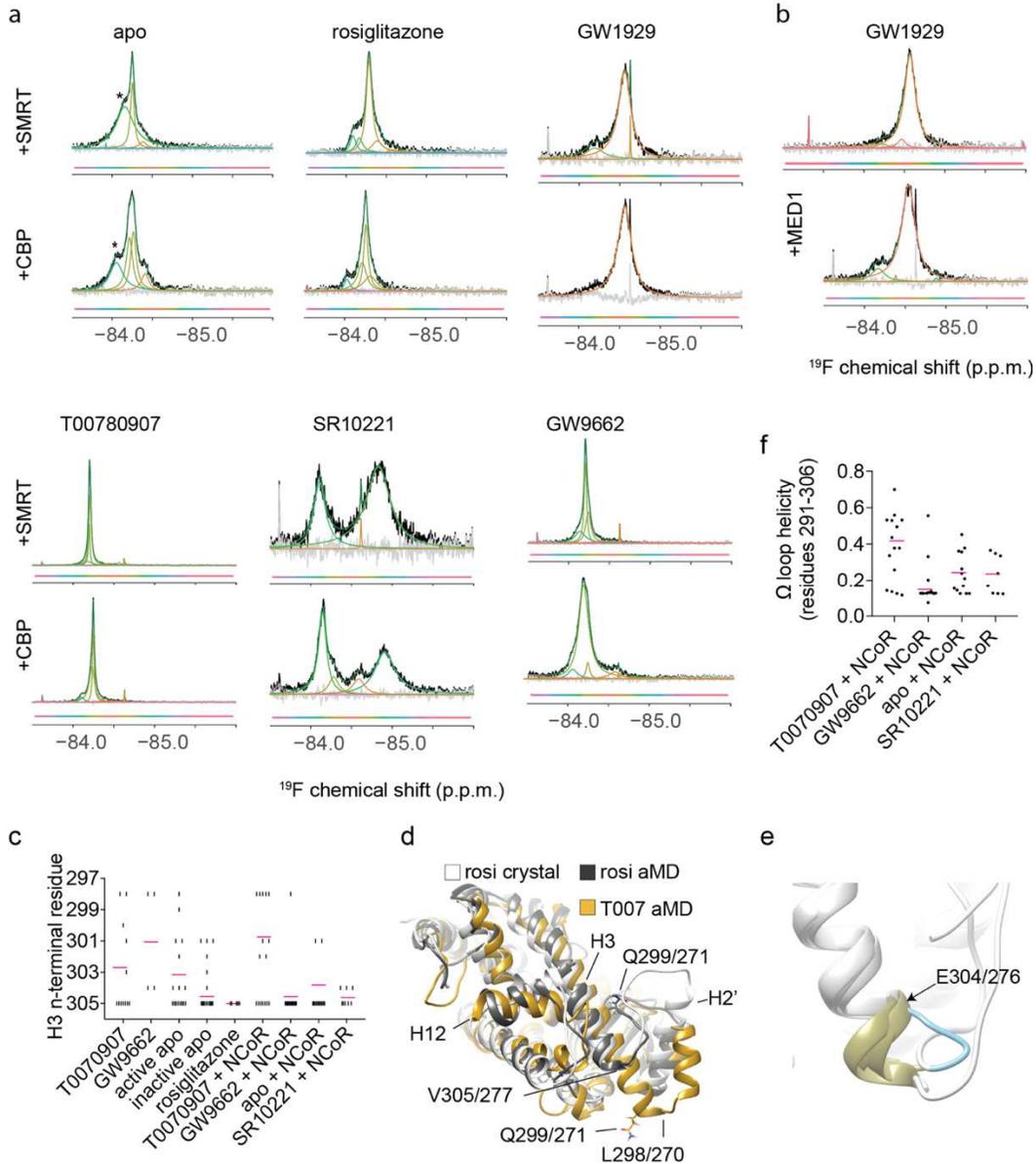
Supplementary Figure 4. The n-terminal portion of helix 3 shows diverse ligand dependent conformations. a-b) PPARγ-Q299C was labeled with BTFA bound to the indicated ligands (or apo) and peptides and fluorine NMR spectra were collected. c) The extent of helix 3 extension beyond the normally observed n-terminus (V305) is indicated for individual simulations along with the Boltzmann average (pink bar). d) Comparison of the crystal structure of rosiglitazone bound PPARγ (white; PDB code 2PRG) with simulation structures of rosiglitazone bound PPARγ (dim grey) and T0070907 bound PPARγ (gold). The location of the probe (Q299) and the usual n-terminus of helix 3 (V305) is shown. e) Three crystal structures show n-terminal helix 3 extension to residue 301/272 (khaki), which is different from the non-helical form usually observed in this region (light blue; PDB code 1PRG). The three structures with n-terminal extension are PPARγ bound to a partial agonist (PDB code 1ZEO) and PPARγ bound to two antagonists (6C5Q, 6C5T). f) Overall Boltzmann average (pink bars) and average helicity (black circles) of a region (291-306) where n-terminal extension is observed in some

simulations. This region was chosen for analysis as it is the same region as observed peptides from the HDX-MS analysis (**Supplementary Figure 8**)
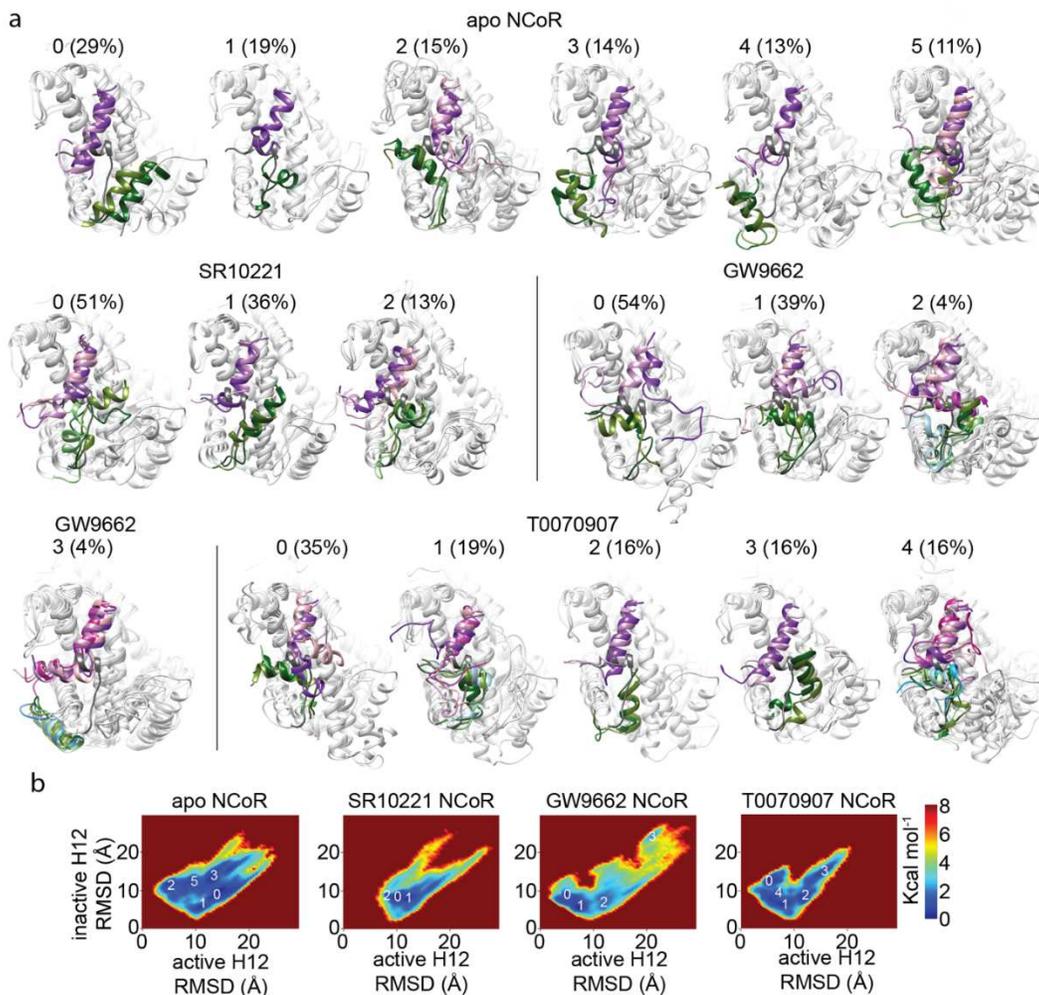
Supplementary Figure 5. Accelerated molecular simulations indicate that two inverse agonists induce different conformations when bound to a peptide from the corepressor peptide NCoR. aMD simulations were run of PPARγ bound to the indicated ligands and a NCoR peptide. a) Representative structures from clustering based on helix 12 RMSD of structures in the wells indicated in panel g. The identity of the energy well and the calculated prevalence of each well's structures within the Boltzmann ensemble (see methods) is indicated by the number and percent in parentheses respectively. The prevalence of the structures within a particular well is indicated by the color of helix 12 ranging from dark green for the most prevalent to olive drab, light green, light blue and then deep sky blue for the least prevalent. For reference, an active structure (rosiglitazone bound PPARγ; 2PRG chain A) is shown with helix 12 colored dim grey. The relative energies of the wells indicated in panel b are shown in **Supplementary Table 9**.

Supplementary Table 4. Solvation of NCoR residues in wt and mutant PPARγ-NCoR complexes.

Complex	carbonyl O of NCoR residue A2270			Arginine side chain of NCoR residue R2268		
	wt	N340A	N304A-wt	wt	K329A	K329A-wt
T0070907	1.2422	2.1903	0.95	7.2963	7.7423	0.45
GW9662	1.1333	2.1412	1.00	7.3156	7.4256	0.11
SR10221	1.1846	1.8381	0.65	7.1038	7.6183	0.51
apo	1.44918	2.2088	0.76	7.3769	7.3099	-0.07

**Supplementary Table 5.** Fold change in corepressor (NCoR) affinity compared to WT PPARγ LBD

Mutation	PPARγ (apo)	PPARγ- T0070907	PPARγ- SR10221	PPARγ- GW9662
Q314A	NS 1.2x ↓	NS 2.6x ↓	NS 1.4x ↑	4.4x ↓
K329A	8.9x ↓	7.2x ↓	7.0x ↓	6.1x ↓
N340A	1.5x ↓	1.5x ↓	NS 1.2x ↓	1.4x ↓
K347A	2.7x ↓	3.2x ↓	NS 1.1x ↓	2.5x ↓
E499L	1.2x ↑	1.6x ↓	4.1x ↑	1.6x ↓
Deletion of Helix 12	5.6x ↑	1.0x	7.9x ↑	2.8x ↑

Fold change for each construct and ligand is shown relative to the same condition in WT PPARγ LBD. Significance is determined by T-test corrected for multiple comparison using the Holm-Sidak method $P \leq 0.05$. A blue arrow indicates a loss of affinity relative to WT while a red arrow indicates a higher affinity. NS indicates that the value is not statistically significant.

**Supplementary Table 6, K_d values for PPARγ constructs and NCoR when bound to each ligand used in this study**

Construct	PPARγ (apo)	PPARγ-T0070907	PPARγ-SR10221	PPARγ-GW9662
WT PPARγ	388.25±38.86	97.33±5.07	424.83±124.86	255.93±23.21
Q314A	480.10±86.27	255.75±107.27	296.50±16.40	1114.20±307.73
K329A	3474.35±18.88	703.05±90.16	2985.00±791.96	1556.50±40.31
N340A	572.70±53.03	142.40±5.80	503.25±29.34	369.95±1.77
K347A	1056.55±64.42	307.45±32.74	473.9±13.01	632.5±55.44
E499L	318.43±5.80	158.93±16.05	104.82±15.24	396.80±84.00
Deletion of Helix 12	70.13±35.64	97.35±61.36	53.65±29.05	90.15±56.35

Values are reported in nanomolar. Error represents the standard deviation of n separate experiments for each condition. For WT PPARγ n = 4, for Q314A n = 2, for K329A n = 2, for N340A n = 2, for K347A n = 2, for E499L n = 3, for Deletion of helix 12 n = 4. K_d values were determined from non-linear regression four-parameter variable slope fit of 12 data points with concentrations ranging from 25 – 50,000nM ligand and protein. All fits were performed in GraphPad Prism 7.02.
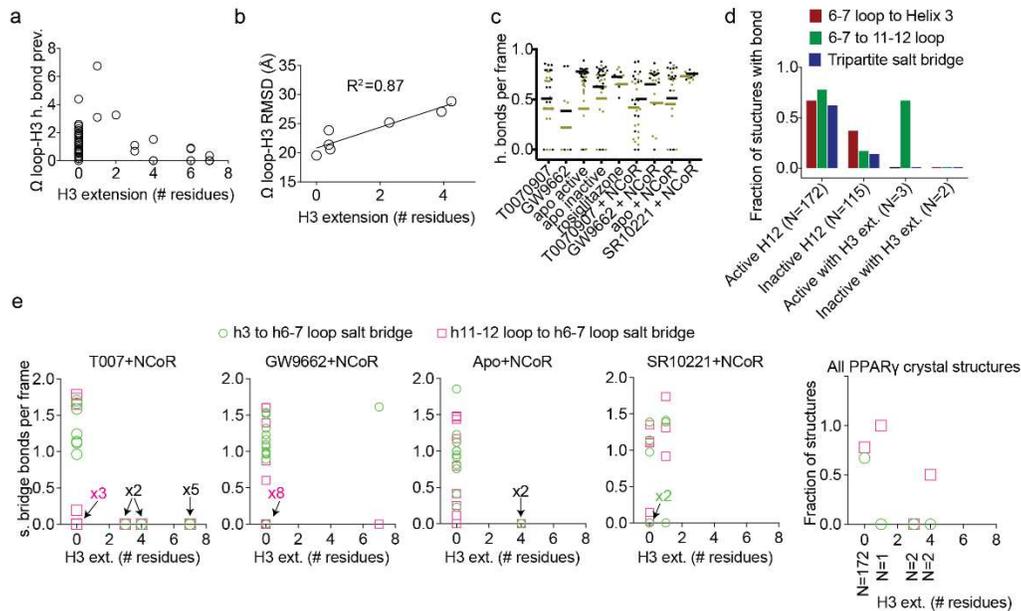
Supplementary Figure 6. Helix 3 extension reduces omega loop-helix 3 interaction and salt-bridge interactions. a) Comparison of the total number of hydrogen bonds between the omega loop and helix 3 for simulations of various complexes and b) comparison between the Boltzmann average of the root mean square deviation between helix 3 and the omega loop and helix 3 extension. c) Comparison of the prevalence of hydrogen bonding between beta sheet residues for the indicated complexes. Bars represent Boltzmann averages, dots represent values of the individual aMD inspired cMD simulations.

d) Analysis of the indicated salt bridge formation for all available crystal structures of PPARγ .e) Comparison of the prevalence of the indicated salt bridges and helical extension for the indicated complexes and for all available crystal structures. The helix 6-7 loop (R385/357) to helix 11-12 loop (E488/460) salt bridge is not observed for the 13 structural ensembles with an n-terminal helix 3 extension of 2 or more residues beyond V305/277 (i.e. the n-terminus of helix 3 normally observed in crystal structures) while it is observed in 22 of 27 structures that lack helix 3 extension. 12 of 13 structural ensembles with an extended helix 3 lack the R385 to E304 salt bridge while it is present in 26 of 27 structural ensembles without an extended helix 3.
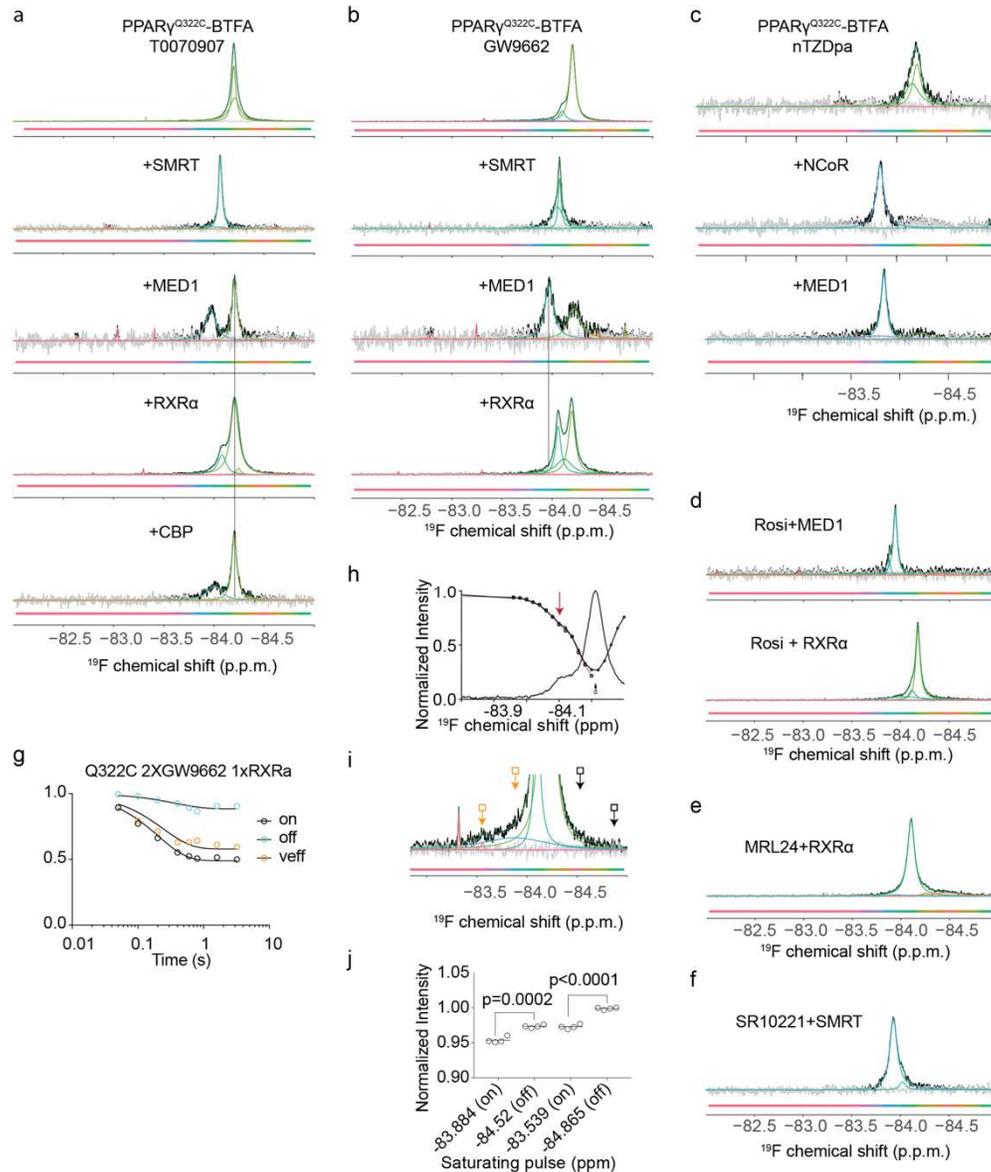
Supplementary Figure 7. ^{19}F NMR reveals structurally and functionally distinct states in the charge clamp region of helix 3.

a-b) PPARγ^{Q322C} covalently bound to T0070907 or GW9662 (which bind to the sole native cysteine, C313) was labeled with high concentration of BTFA (40x the protein concentration of ~150 μM) to yield high signal to noise spectra with no possibility of labeling the native cysteine. The indicated coregulators or RXRα were added at a 2:1 molar ratio. These spectra yield evidence of multiple peaks even without other proteins or peptides bound, most obviously for PPARγ^{Q322C}-BTFA bound to GW9662 (panel b). d-f) PPARγ^{Q322C}-BTFA was loaded with the indicated ligands and 19F NMR spectra were acquired followed by deconvolution. g) Saturating pulses (circles) were applied at the minor peak chemical shift (on; black circles) and at a control location equidistant from the main peak (off; blue circles) for various durations. The normalized intensity of the main peak was monitored and plotted against the duration of the pulses and fit to determine the exchange rate (1.7 s^{-1}; 95% CI=1.5-2.0). h) CEST indicates exchange between the

239

minor left peak and the major peak in PPARγ^{Q322C}-BTFA (top spectrum; panel b). The position of the black circles indicate the region of saturation and their y position indicates the height of the major peak (black arrow) upon saturation. A small deviation is noted when saturating near the minor peak (red arrow) indicating exchange between the minor and major peak. i) deconvolution of the GW9662 spectrum indicated the presence of signal downfield (left) of the readily visible peaks. We tested for exchange of these peaks with the major peak using selective saturating pulses in this region (orange arrows) and control upfield pulses equidistant from the major peak (black arrows). The intensity of the major peak was measured after saturating pulses at these locations and compared. This was repeated 4 times on the same sample at each location. A two-tailed t-test indicates that the downfield peaks exist and exchange with the major peak.

Sequence	Charge	Start	End	Structure	Apo vs T007	SR10221 vs T007
RALAKHLYDS	3	240	249		0 (2)*	0 (2)*
RALAKHLYDSY	3	240	250		0 (1)*	0 (1)*
IKSFPLTKGAKARAIL	3	251	265		23 (5)	39 (3)
TGKTTDKSPFVIYDM	3	266	280		-7 (2)	N/A
TGKTTDKSPFVIYDMNSLM	3	266	284		-10 (3)	N/A
MGEDKIKFKHITPLQEQSKE	3	285	304		-3 (3)*	-1 (2)*
MGEDKIKFKHITPLQEQSKEVA	3	285	306		-4 (2)*	0 (2)*
KIKFKHITPLQEQSKEVA	3	289	306		-3 (2)*	-1 (3)*
IRIFQGCQ	2	307	314		N/A	8 (3)
AVQEITE	1	320	326		-21 (2)	-6 (1)
YAKSIPGF	2	327	334		0 (1)*	N/A
DLNDQVTL	1	338	345		-1 (1)*	-1 (1)*
LKYGVHE	2	346	352		-10 (1)	-8 (1)
LKYGVHEIIY	2	346	355		-4 (0)*	-2 (0)*
LKYGVHEIIYTM	3	346	357		-3 (1)*	-2 (1)*
ASLMNKDGVL	2	359	368		-8 (2)	-1 (1)*
MNKDGVL	2	362	368		-8 (2)	-2 (1)*
ISEGQGFMTRE	2	369	379		-22 (3)	0 (3)*
ISEGQGFMTREFL	2	369	381		-16 (3)	1 (2)*
FLKSLRKPFGD	2	380	390		-1 (1)*	0 (2)*
FLKSLRKPFGDFMEPKFEF	3	380	398		-14 (2)	-4 (2)*
LRKPFGDF	2	384	391		-33 (3)	-23 (3)
LRKPFGDFMEPKFEF	3	384	398		-16 (3)	-8 (2)
AVKFNAL	2	399	405		-9 (2)	-3 (2)*
NALELDDSDL	1	403	412		1 (1)*	1 (1)*
IILSGDRPGLL	2	419	429		0 (2)*	1 (2)*
IILSGDRPGLLNVKPIE	3	419	435		0 (2)*	1 (2)*
IILSGDRPGLLNVKPIED	3	419	436		0 (2)*	-1 (2)*
NVKPIEDIQDNL	2	430	441		0 (1)*	1 (1)*
NVKPIEDIQDNLLQA	2	430	444		0 (1)*	0 (1)*
LELQLKLNHPESSQL	3	445	459		0 (1)*	0 (1)*
ELQLKLNHPESSQL	2	446	459		-1 (2)*	0 (2)*
QLKLNHPESSQL	2	448	459		0 (1)*	0 (1)*
KLNHPESSQL	2	450	459		-1 (2)*	-1 (2)*
FAKLLQKMTDL	2	460	470		-1 (1)*	-1 (2)*
FAKLLQKMTDLRQ	3	460	472		-9 (2)	-4 (2)*
LQKMTDL	2	464	470		-2 (2)*	-2 (1)*
LQKMTDLRQ	3	464	472		-7 (2)	-7 (3)
RQIVTE	2	471	476		-30 (3)	N/A
RQIVTEHVQL	3	471	480		-32 (2)	N/A
LQVIKKTETDM	2	481	491		-6 (2)	-3 (2)*
LQVIKKTETDMSLHPLL	3	481	497		-12 (2)	-10 (2)
LQVIKKTETDMSLHPLLQE	3	481	499		-18 (2)	-17 (2)
IKKTETDMSLHPLL	3	484	497		-11 (2)	-11 (2)
HPLLQEIYKDLY	2	494	505		-72 (4)	-81 (3)
QEIYKDLY	1	498	505		-62 (3)	-69 (3)

Supplementary Figure 8. HDX-MS data of PPARγ bound to NCoR peptide alone or co-bound to T0070907 or SR10221. Detected peptide sequences residue numbers and change in HDX is shown. A decrease in exchange in the T0070907 complex peptide compared to the indicated complexes is shown as a negative number while an increase is shown as a positive number. Standard deviation is shown in parantheses. Peptides with significant differences are shown in color, while insignificant are shown in grey and have an asterisk. NA means matching peptides were not found.

Supplementary Table 7. K_d values for PPARγ constructs and SMRT when bound to each ligand used in this study

Construct	PPARγ (apo)	PPARγ-T0070907	PPARγ-SR10221	PPARγ-GW9662
WT PPARγ	1353.25±300.41	120.57±62.12	481.25±163.72	592.85±276.32
Q314A	1367.00±417.44	248.23±79.76	246.37±36.23	2050.33±489.18
K329A	5772.67±527.37	270.30±4.950	1868.00±134.35	1468.50±324.56
N340A	2173.00±297.86	107.73±41.91	799.48±154.84	861.88±175.17
K347A	1986.00±360.43	147.50±11.49	466.80±109.70	556.17±43.63
E499L	1235.00±157.63	312.77±36.06	309.87±89.00	811.23±82.30
Deletion of Helix 12	270.27±92.22	181.33±27.61	102.46±43.99	303.23±49.77

Values are reported in nanomolar. Error represents the standard deviation of n separate experiments for each condition. For WT PPARγ n = 4, for Q314A n = 3, for K329A n = 2, for N340A n = 4, for K347A n = 3, for E499L n = 3, for Deletion of helix 12 n = 3. K_d values were determined from non-linear regression four-parameter variable slope fit of 12 data points with concentrations ranging from 25 − 50,000nM ligand and protein. All fits were performed in GraphPad Prism 7.02.

Supplementary 8. K_d values for PPARγ constructs and SMRT when bound to each ligand used in this study

Mutation	PPARγ (apo)	PPARγ-T0070907	PPARγ-SR10221	PPARγ-GW9662
Q314A	NS 1x	NS 2.1x ↓	NS 2x ↑	3.5x ↓
K329A	4.3x ↓	NS 2.2x ↓	3.9x ↓	NS 2.5x ↓
N340A	1.6x ↓	NS 1.1x ↑	6x ↑	NS 1.5x ↓
K347A	NS 1.5x ↓	NS 1.2x ↓	NS 1x	NS 1.1x ↑
E499L	NS 1.1x ↑	2.6x ↓	NS 1.6x ↑	NS 1.4x ↓
Deletion of Helix 12	5x ↑	NS 1.5x ↓	4.7x ↑	NS 2x ↑

Fold change for each construct and ligand is shown relative to the same condition in WT PPARγ LBD. Significance is determined by T-test corrected for multiple comparison using the Holm-Sidak method $P \leq 0.05$. A blue arrow indicates a loss of affinity relative to WT while a red arrow indicates a higher affinity. NS indicates that the value is not statistically significant.
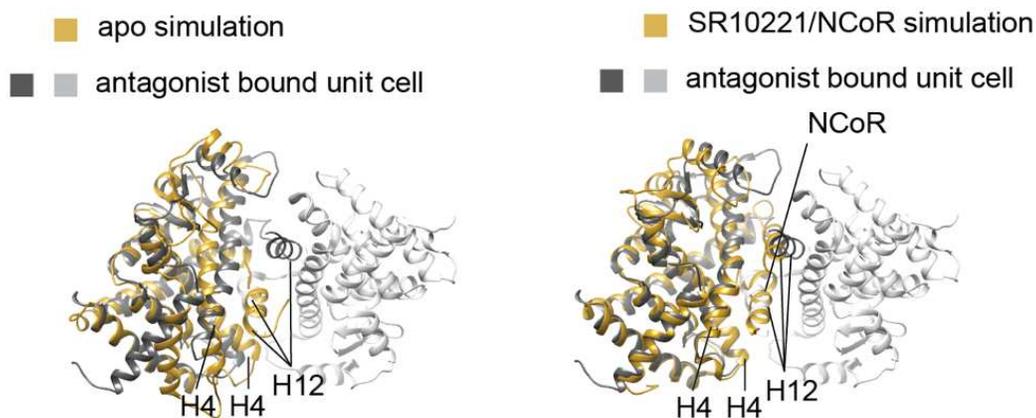
Supplementary Figure 9. Helix 12 binds to the coregulator surface in many crystal structures in a similar manner to NCoR and Helix 12 in simulations. Representative structures from the fifth lowest energy well in apo PPARγ where helix 12 binds to the coregulator binding surface (left; gold) and the second lowest energy well from the PPARγ cobound to SR10221 and NCoR (right; gold) simulations are aligned with the unit cell of an antagonist bound structure (PDB code 6c5q).
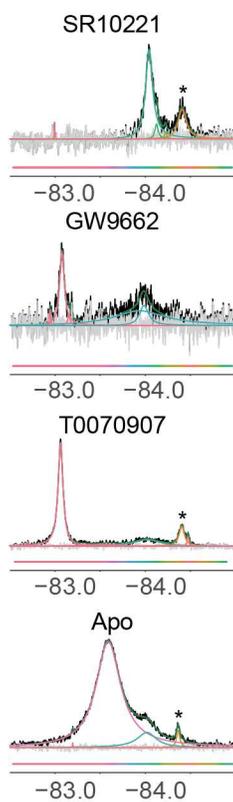
Supplementary Figure 10. **SMRT induces similar changes to helix 12 as NCoR for PPARγ bound to the indicated ligands.** PPARγ^{K502C}-BTFA was saturated with the indicated ligands and a 2:1 molar ratio of SMRT peptide was added. Fluorine NMR was then performed.

Supplementary Table 9. Energy Values of minimums in the potential energy landscape.

energy level	0	1	2	3	4	5
	(kcal mol^{-1} relative to well with lowest energy)					
apo PPARγ-NCoR	0	0.25	0.40	0.42	0.47	0.56
PPARγ-SR10221-NCoR	0	0.20	0.80			
PPARγ-GW9662-NCoR	0	0.20	1.61	1.61		
PPARγ-T0070907-NCoR	0	0.35	0.46	0.47	0.55	
apo PPARγ (chain a)	0	0.07	0.44	0.45	0.46	
apo PPARγ (chain b)	0	0.75	1.06	1.12	1.23	
PPARγ-GW9662	0	0.09				
PPARγ-T0070907	0	0.10	0.69			
PPARγ-rosiglitazone	0					
PPARγ-SR10221	0	0.34	1.04			