Crystal Structures of LRH-1 with Synthetic Agonists Reveal a Novel Mechanism of Receptor Activation

Suzanne G. Mays[1], C. Denise Okafor[1], Richard J. Whitby[3], Devrishi Goswami[4], Józef Stec[3], Autumn R. Flynn[2], Michael C. Dugan[2], Nathan T. Jui[2], Patrick R. Griffin[4], and Eric A. Ortlund[1*]

From the [1]Department of Biochemistry, Emory University School of Medicine, and the [2]Department of Chemistry, Emory University, Atlanta, GA 30322; [3]School of Chemistry, University of Southampton, Southampton, United Kingdom; and the [4]Scripps Research Institute, Jupiter, FL.

To whom correspondence should be addressed: Eric A. Ortlund, Department of Biochemistry, Emory University School of Medicine, Atlanta, GA 30322, USA, Tel.: (404) 727-5014 Fax: (404) 727- 2738 Email: eortlun@emory.edu

Liver receptor homolog 1 (NR5A2, LRH-1) is an orphan nuclear hormone receptor that regulates diverse biological processes, including metabolism, proliferation, and the resolution of endoplasmic reticulum stress. While preclinical and cellular studies demonstrate that LRH-1 has great potential as a therapeutic target for metabolic diseases and cancer, development of LRH-1 modulators has been difficult. Recently, systematic modifications to one of the few known chemical scaffolds capable of activating LRH-1 failed to improve efficacy substantially. Moreover, mechanisms through which LRH-1 is activated by synthetic ligands are entirely unknown. Here, we use x-ray crystallography and other structural methods to explore conformational changes and receptor-ligand interactions associated with LRH-1 activation by a set of related agonists. Unlike phospholipid (PL) LRH-1 ligands, these agonists bind deep in the pocket and do not interact with residues near the mouth, nor do they expand the pocket like PLs. Unexpectedly, two closely related agonists with similar efficacies (GSK8470 and RJW100) exhibit completely different binding modes. The dramatic repositioning is influenced by a differential ability to establish stable, face-to-face π-π-stacking with LRH-1 residue H390, as well as by a novel polar interaction mediated by the RJW100 hydroxyl group. The differing binding modes result in distinct mechanisms of action for the two agonists. Finally, we identify a network of conserved water molecules near the ligand-binding site that are important for activation by both agonists. This work reveals a previously unappreciated complexity associated with LRH-1 agonist development and offers insights into rational design strategies.

Liver Receptor Homolog 1 (LRH-1; NR5A2) is a nuclear hormone receptor (NR) that controls expression of a diverse set of genes important both in normal physiology and disease. In addition to a vital role during development (1,2), LRH-1 regulates many genes related to metabolism, proliferation, and cell survival. In the liver, LRH-1 regulates bile acid biosynthesis (3) and reverse cholesterol transport (4,5), affecting hepatic and circulating cholesterol levels. Glucose metabolism is also regulated by LRH-1 at several points, including GLUT-4-mediated transport (6) and glucose phosphorylation, the latter of which is essential for proper postprandial glucose sensing, flux through glycolysis and glycogenesis pathways, and *de novo* lipogenesis (7). LRH-1 is a key mediator of the cell stress response through control of genes involved in the hepatic acute phase response (8), and in the cytoprotective resolution of endoplasmic reticulum stress (9). Additionally, LRH-1 can be aberrantly overexpressed in certain cancers and can promote tumor growth through estrogen receptor and β-catenin signaling (10-16).

Considering the breadth and significance of these physiological effects, LRH-1 modulators are highly desired as potential therapeutic agents. Chemical modulators would also be extremely useful as tools to dissect complex or temporal aspects of LRH-1 biology. However, development of LRH-1-targeted compounds has been challenging, in part due to a lipophilic binding pocket that becomes occupied with bacterial phospholipids (PL) in recombinant protein. Very

few small molecules are able to displace these PL in library screens. Moreover, ligand-mediated regulation of LRH-1 is poorly understood. Endogenous ligands for LRH-1 are unknown, but exogenous administration of dilauroyl-phosphatidylcholine (DLPC, PC 12:0/12:0) activates LRH-1 and has profound anti-diabetic effects *in vivo*, which are absent in a liver-specific LRH-1 knockout mouse (17). In addition to PCs, the signaling PL phosphatidylinositol 3,4,5-triphosphate (PIP3) binds LRH-1 (18,19), although downstream effects of this interaction have yet to be determined.

Typically, NRs are activated by a ligand-induced conformational change, which promotes recruitment of coactivator proteins to the activation function surface (AFS) in the ligand binding domain (LBD) to drive transcription. Our structural studies with DLPC have shown that, contrary to the canonical model of NR activation, LRH-1 relies on small conformational fluctuations to recruit co-activator or co-repressor proteins. These occur mainly in the AFS (comprised of portions of H3, H4, and the AF-H in the LBD), as well as in the H6/β-sheet region at a distal portion of the LBD (20,21). Flexibility in in the H6/β-sheet region is required for activation by PLs (20). Mechanisms through which LRH-1 is activated by synthetic ligands have not been explored but are likely quite different, given the differing structural composition of synthetic *versus* PL ligands.

There are very few known chemical scaffolds capable of activating LRH-1 above basal levels, the best studied of which are the *cis*-bicyclo[3.3.0]-octenes discovered by Whitby (22,23). The first compound described with this scaffold, named GSK8470 (Figure 1), was somewhat effective but was acid labile (23). Substitution of the aniline group improved compound stability, and the GSK8470-LRH-1 crystal structure provided the basis for an extensive structure-activity relationship study (23). One of the major objectives of this study was to introduce functional groups near select polar residues within the predominantly hydrophobic pocket. The new lead compound produced from this study, named RJW100, contains an *exo* hydroxyl group at the 1-position of the pentalene scaffold (indicated in red in Figure 1), intended to interact with LRH-1 residues H390 or R393. On the other hand, a diastereomer with *endo* stereochemistry (previously known as 24-endo, Figure 1) was not predicted to be able to make these interactions due to the alternative conformation of the hydroxyl oxygen. The *endo* derivative was less active in biochemical assays, seeming to support this hypothesis. However, RJW100 was not much more potent or effective than GSK8470, and the study did not illuminate strategies for further improvement.

In this work, we present crystal structures of RJW100 and its *endo* diastereomer bound to LRH-1. We demonstrate that these compounds bind quite differently than PLs and have distinct effects on protein dynamics compared to DLPC. Unexpectedly, these agonists also bind quite differently than the very closely related compound, GSK8470. We identify receptor-ligand interactions driving the repositioning and show that particular interactions are important for LRH-1 activation. These findings provide the first description of mechanisms involved in LRH-1 activation by synthetic molecules.

Results

The Crystal Structure of RJW100 Bound to LRH-1. To understand how RJW100 interacts with LRH-1 and affects receptor conformation, we determined the X-ray crystal structure of LRH-1 LBD bound to the agonist and to a fragment of the coactivator, Tif2, to a resolution of 1.85 Å (Table 1 and Figure 2A). Although the RJW100 used for crystallization was a racemic mixture of two *exo* stereoisomers (Figure 1), the electron density in the structure unambiguously indicates that a single enantiomer is bound (Figure 2B). The bound isomer has R stereochemistry at both the 1- (hydroxyl-substituted) and 3a- (styrene-substituted) positions (hereafter RR-RJW100, Figure 1). The ligand is bound at a single site deep in the binding pocket and is fully engulfed within it. This binding mode is markedly different than that of the PL ligands, DLPC and PIP3, which extend lower in the pocket with the headgroups protruding into the solvent (seen by superposition with PDBs 4DOS and 4RWV, respectively, Figure 2C-D). PL ligands also increase the pocket volume and width compared to RJW100. For example, the mouth of the pocket is ~3 Å wider and nearly 40% larger in volume when DLPC is bound *versus* RJW100 (Fig 2E and Table 2). This effect appears to be mainly due to a shift of H6, which swings

away from the mouth of the pocket in the DLPC structure by approximately 3 Å (Fig. 2E). The direction and magnitude of the H6 movement is similar in published LRH-1-PL structures: comparison of four published human LRH-1-PL structures shows an average H6 shift of 3.0 +/- 0.2 Å relative to LRH-1 in the apo state or when synthetic ligands are bound (mean +/- SEM, Figure 2F). Although these structures exhibit diverse types of crystal packing, the movement of H6 appears to be related to whether the ligand is a PL or small molecule and not to crystal form or packing contacts. It likely occurs to avoid stearic clashes with the PL headgroup. Notably, the H6/β sheet region has been recently identified as a site through which PL ligands allosterically communicate with the activation function surface (AFS) to modulate LRH-1 activity (20,21). The fact that the synthetic agonists do not displace H6 relative to apo receptor suggests that they utilize a different mechanism for receptor activation.

RJW100 Selectively Destabilizes Components of the AFS Relative to DLPC. The overall LRH1-RJW100 structure depicts the AF-H in the active conformation and the Tif2 coactivator peptide bound at the AFS, as expected. However, there are a few indications that the active state may not be fully stabilized. There is substantial disorder in the loop connecting Helix 10 to the AF-H, and three residues within this loop cannot be modeled (dotted line in Fig. 2A). This loop is not disordered in published structures of LRH-1 bound to DLPC or GSK8470, and this does not appear to be related to favorable crystal packing in those structures (not shown). A second site of disorder in our structure occurs in the sidechain of residue E534 (not shown), which plays a critical role in securing coactivators via charge clamp (24).

The disorder in the vicinity of the AFS in our structure suggests that this region is mobile when RJW100 is bound. To test this explicitly, we utilized solution-based hydrogen-deuterium exchange mass spectrometry (HDX-MS) to compare dynamics of purified LRH-1 LBD bound to enantiomerically pure RR-RJW100 or to DLPC. As shown in Figure 3A, RR-RJW100 destabilizes the bottom of H10 and the pre-AF-H loop *versus* DLPC, although the AF-H itself is slightly stabilized. Additional destabilization by RR-RJW100 occurs at the top part of H3 (part of the AFS), and in β2 (part of the alternate activation function region responsible for allosteric communication with the AFS, (20,21)). Interestingly, this localized destabilization occurs even though RJW100 increases overall thermostability of LRH-1 compared to DLPC (Figure 3B). This finding may indicate room for improvement in agonist design: compounds that provide a stable surface for coactivator recruitment would likely be more potent activators.

Dramatic Repositioning of RJW100 Compared to a Closely Related Synthetic Agonist. Perhaps the most striking observation from our structure comes from comparison with GSK8470-bound LRH-1 (PBD 3PLZ). Overall protein conformation is highly similar: the largest movement occurs in the bottom of H3, which moves in the direction of H6 (by 2 Å in the RR-RJW100 structure and by 4 Å in the GSK8470 structure relative to Apo-LRH-1, Figure 4A). However, there is a substantial difference in the positioning of these agonists within the binding pocket. Although GSK8470 and RR-RJW100 bind in the same vicinity, they are rotated nearly 180° from one another. The bicyclic rings at the cores of each molecule are perpendicular to each other, causing the tails to be pointed in opposite directions (Figure 4B). Notably, the rationale for adding a hydroxyl group in the 1 position on this scaffold was to promote an interaction with a "polar patch," consisting of residues R393 and H390 in an otherwise hydrophobic pocket (23). This interaction was predicted based on the position of the ligand in the LRH-1-GSK8470 structure; however, the actual position of RR-RJW100 in the pocket places the hydroxyl group over 6 Å away from these residues (Figure 4C). Such a radically different binding mode for closely related molecules was unexpected, and a propensity to rotate within the pocket may contribute to difficulties improving agonist activity by modification of the GSK8470 scaffold.

The crystal structure of LRH-1 bound to a RJW100 diastereomer. To further explore the effect of agonist structure on ligand binding mode, we determined the 1.93 Å crystal structure of LRH-1 bound to *endo*-RJW100, also in complex with the Tif2 coactivator peptide (Table 1 and Figure 5A). Overall protein conformation is highly

similar to that of the RR-RJW100 (*exo)* structure, and there is disorder at the same portion of the protein backbone within the pre-AFH loop (Figure 5A).

Like *exo*-RJW100, *endo*-RJW100 consists of a mixture of two enantiomers, only one of which is bound in the crystal structure. The presence of a single enantiomer is quite clear from the electron density surrounding the ligand, although the density for the ligand tail is much weaker than in the RR-RJW100 structure (Figure 5B). The bound enantiomer has R stereochemistry at the styrene moiety, as in the *exo*-RJW100 structure, but the hydroxyl group is pointed in the opposite direction (designated SR-RJW100, Figure 1). Superposition of the ligand coordinates from the RR-RJW100 and SR-RJW100 structures reveals nearly identical positioning, with the exception of the hydroxyl group (Figure 5C). However, a key difference is seen in the ligand B-factors. For RR-RJW100, the average ligand B-factor is 34.8, which is slightly less than the average protein B (38.1, Table 1). In contrast, the ligand B-factor for SR-RJW100 is much higher compared to the protein B-factor (59.2 *versus* 46.9, Table 1). The higher average ligand B-factors do not arise solely from the disordered atoms in the ligand tail: for example, the B-factor for the hydroxyl oxygen of SR-RJW100 is 59. This is an indication of more atomic motion for SR-RJW100 compared to RR-RJW100, suggesting that the ligand has a reduced ability to engage in stabilizing intermolecular interactions in the pocket.

Discovery of a Novel LRH-1 Interaction Mediated by the RJW100 Hydroxyl Group. Protein-ligand interactions made by GSK8470 and the RJW100 stereoisomers were examined to gain insight into factors influencing ligand-binding mode. A close view of the LRH-1 binding pocket reveals that RR-RJW100 makes several hydrophobic contacts, many of which are also made by GSK8470 (shown in cyan, Figure 6A). Additionally, RR-RJW100 makes several unique contacts (shown in grey, Figure 6A). Many of these unique contacts are also hydrophobic; however, the RR-RJW100 hydroxyl group forms an indirect polar contact with residue T352 via a water molecule. A portion of the electron density map is shown in Figure 6A to emphasize the strong evidence for this interaction. SR-RJW100 also interacts with T352 through the same water molecule, despite the differing conformation of the hydroxyl group (Figure 6B). The position of the SR-RJW100 hydroxyl group also permits a second water-mediated hydrogen-bonding interaction with the backbone nitrogen of residue V406 (Figure 6B).

Although the interaction with T352 is indirect, we observed that the water molecule involved is part of a network of waters found in every LRH-1 crystal structure in the same location (for examples see PBDs 4DOS, 1YUC, 3PLZ, and 4DOR, as well as Figure 7 in this paper). Thus, this water network appears to be a conserved feature of the binding pocket and may play a role in receptor function or stability. To test the hypothesis that the OH-water-T352 interaction was influencing ligand positioning, we analyzed the stability of this bond using molecular dynamics simulations (MDS). Throughout each simulation (200 ns), the four conserved networked water molecules remained in the same positions (if a particular water molecule occasionally left, it was immediately replaced with another in the same location). Residue T352 maintained a hydrogen bond with the water molecule for 100% of each simulation, regardless of which ligand was bound. Additionally, both RR-RJW100 and SR-RJW100 maintained hydrogen bonding with the water molecule for the majority of the simulations (53.7% of the time for RR-RJW100 and 64.4% of the time for SR-RJW100). When residue T352 was mutated to valine in MDS, the time spent interacting with the T352-coordinated water molecule was drastically reduced (22.9% and 0.5% when RR-RJW100 and SR-RJW100 were bound, respectively), demonstrating that this mutation likely disrupts this water-mediated interaction made by these ligands.

Differences in π-π-Stacking with Residue H390 among LRH-1 Agonists. π-π-Stacking of GSK8470 with residue H390 has been previously described and is hypothesized to be critical for activation of LRH-1 by synthetic compounds (23,25). The RJW100 diastereomers also engage in π-π-stacking with H390, but with some key differences. The π-π-stacking is face-to-face for GSK8470 and edge-to-face for the RJW100 isomers (Figure 7). Additionally, by virtue of the very different orientations in the binding pocket, the agonists do not use analogous phenyl rings for

π-π-stacking: GSK8470 uses the aniline group, whereas the RJW100 isomers use the adjacent phenyl substituent. Moreover, MDS demonstrate ligand-dependent differences in the stability of this interaction. For this latter analysis, time spend in π-π-stacking was quantified over the course of the 200 ns MDS. Face-to-face π-π-stacking was defined as a distance between ring centroids of < 6 A, an angle between ring planes < 45°, and an angle between the centroid-centroid vector and one plane < 60° (26). Edge-to-face stacking was defined as a distance of < 5.5 Å between ring centroids and an angle between the ring planes between 60° and 120° (26,27). Applying these criteria as appropriate for the type of π-π-stacking made by each ligand revealed that GSK8470 maintained π-π-stacking with H390 for most of the simulation (89.5% of the time, Figure 7A). The edge-to-face π-π-stacking made by RR-RJW100 was also fairly stable, maintained for 59.6% of the MDS time (Figure 7B). In contrast, SR-RJW100 made this interaction much less frequently (22.4% of the time) and exhibited much more motion than RR-RJW100 (Figure 7C). These observations are consistent with the relatively high B-factors seen for SR-RJW100 in the crystal structure.

Role of T352 and H390 in LRH-1 Activation by Synthetic Agonists. The importance of the T352 and H390 interactions for binding and activation of LRH-1 by the agonists was investigated using mutagenesis. Binding and stabilization of LRH-1 were detected using DSF. While a T352V mutation (designed to remove the water-mediated hydrogen bond with bound ligands) had little effect on the overall thermostability of DLPC-bound LRH-1, it completely abrogated the stabilizing effect of RR-RJW100 and SR-RJW100 (Figure 8A). Likewise, disrupting this interaction by using a RJW100 analog lacking the hydroxyl group (named 18a) prevented the positive Tm shift in wild-type (WT) LRH-1 (Figure 8B). GSK8470 did not affect the melting profile of LRH-1 in WT or T352V protein, supporting the notion that the hydroxyl group is important for stabilizing the protein-ligand complex.

The T352 interaction was also found to be important for LRH-1 activation by small molecule agonists. Compound 18a, lacking the hydroxyl group and unable to make this interaction, was an extremely poor LRH-1 activator in luciferase reporter assays (Figure 8C). *Endo*-RJW100 was also a weak agonist, as previously reported (23), although statistically significant activation was achieved at the highest dose with WT LRH-1 (~1.4-fold over DMSO, Figure 8F). RR-RJW100 and GSK8470 were equally effective toward WT LRH-1: both increased activity by approximately 2.5-fold compared to DMSO at the highest dose, and both had EC_{50}s of around 4 μM (Figure 8D-E). Notably, the T352V mutation greatly reduced the ability of RR-RJW100 to activate LRH-1 compared to with WT protein, while not significantly affecting baseline activity (Figure 8D). Unexpectedly, this mutation similarly attenuated activation by GSK8470, perhaps suggesting a broader role for this residue (or perhaps for the water network it coordinates) in ligand-mediated activation. Indeed, introduction of a T352V mutation to GSK8470-bound LRH-1 in MDS disrupts the water network, causing complete displacement of the water molecule typically coordinated by T352 (Figure 8G). The T352V mutation also significantly reduces the amount of time GSK8470 spends π-π-stacking with H390 (25.7% *versus* 89.5% of the simulation, Figures 8H and 7A). Both the destabilization of the water network and the disruption of stable H390 π-π-stacking by the T352V mutation could contribute to the observed loss of activity for GSK8470 in the context of this mutation.

While the T352V mutation resulted in a loss of activity for both RR-RJW100 and GSK8470, mutating H390 to alanine had a different effect on LRH-1 activation depending on the agonist involved. GSK8470 was completely unable to activate H390A-LRH-1, but this mutation had little to no effect on RR-RJW100-mediated activation (Figure 8D-E). This differential reliance on H390 for activation is consistent with the observation that GSK8470 interacts with H390 more stably than RR-RJW100 in MDS. This also provides evidence that RR-RJW100 must utilize a different mechanism of action than GSK8470 for LRH-1 activation.

Capacity for productive π-π-stacking with residue H390 influences agonist positioning. Considering the stable nature of the π-π-stacking GSK8470 with residue H390, reasons that RJW100 would abandon this strong interaction for an entirely different binding mode were unclear. To answer

this question, we artificially re-oriented RR-RJW100 in our structure to be aligned with GSK8470, inspected the fit, and looked at predicted interactions using MDS. While this repositioning did not produce any obvious clashes, the planarity of the styrene moiety restricted its mobility. This constraint prevents the rotation necessary for ideal face-to-face π-π-stacking with H390 (Figure 9A). Although MDS using this re-oriented ligand demonstrated fairly stable face-to-face π-π-stacking for re-oriented RR-RJW100, the interaction was less stable than GSK8470 (present 71.1% of the time *versus* 89.5% for GSK8470, $p < 0.0001$, Figure 9B). There was increased variability in both ring centroid distances and angle between the rings involved in this interaction for repositioned RR-RJW100 (mean centroid distances +/- standard deviations were 5.1 +/- 0.6 Å *vs* 4.2 +/- 0.3 Å for GSK8470, and mean angles were 40 +/- 20 *vs* 30 +/- 10 degrees for GSK8470). This variability indicates more motion of the rings involved in the *π-π*-stacking for the repositioned ligand. Another interesting observation with this MDS pertained to the water network in the ligand-binding pocket. The hydroxyl group of the repositioned ligand was initially within hydrogen-bonding distance of the T352-coordinated water molecule. Unlike native RR-RJW100, however, the hydroxyl group in the repositioned ligand was also within hydrogen-bonding distance of the three other water molecules in the immediate vicinity (Figure 9A, right panel). During the MDS, the water network was very quickly disrupted, with a complete loss one of the four conserved water molecules near the ligand hydroxyl group. One of the remaining three waters moved to coordinate T352, while the ligand hydroxyl group alternated between interactions with each of the three waters without stably interacting with a single water molecule (data not shown). Thus, the water-mediated contact of the R-RJW100 hydroxyl group with T352 is not maintained for the reoriented ligand. Moreover, the loss of a typically conserved water molecule supports the idea that the ligand-binding pocket is more dynamic with the artificially repositioned ligand. Therefore, it appears that an impaired ability of the styrene phenyl ring to interact with H390, combined with a favorable interaction mediated by the hydroxyl group with the T352-coordinated water, is responsible for the dramatically different position adopted by RJW100 compared to a structurally very similar agonist.

Discussion

Although LRH-1 synthetic modulators are highly sought as pharmacological tools and as potential therapeutic agents, a limited understanding of ligand characteristics important for binding and activating LRH-1 has impeded agonist development. This work represents the first detailed exploration of structural mechanisms governing regulation of LRH-1 by synthetic ligands. Relative to the PL LRH-1 agonist, DLPC, the current best agonist (RR-RJW100) constricts the binding pocket and destabilizes portions of the AFS (Figure 1-2). In future studies, it will be interesting to investigate the causes of this latter effect, since stabilization of the AFS may facilitate co-activator binding, leading to greater potency or efficacy. Alternatively, analogs designed to enhance the AFS destabilization may be effective antagonists or inverse agonists.

In a previous study, RJW100 was the most effective of a large series of GSK8470 derivatives but still only modestly increased LRH-1 activation (23). Indeed, we find that these two agonists are statistically indistinguishable in luciferase reporter assays measuring LRH-1 activity (Figure 8D). Given the similarities in structures and efficacies for these ligands, we expected them to utilize similar mechanisms of action; however, this is not the case. Our crystal structure reveals a dramatically different binding mode for RR-RJW100 compared to GSK8470 (Figure 4). While this was surprising, it is not unreasonable, considering that LRH-1 has a very large hydrophobic binding pocket and that these agonists are also quite hydrophobic, filling only 37% of the available space (excluding waters). It is possible that many of the GSK8470 analogs investigated in the previous SAR study adopt a variety of different conformations. This seems to be the case in our docking studies with these ligands: multiple, very different binding modes with similar energies are predicted (data not shown). Importantly, however, the repositioning of RR-RJW100 in our structure appears to be driven by particular interactions, since SR-RJW100 assumes a very similar pose (Figure 5). This occurs despite the fact that the SR derivative

exhibits signs of motion in our crystal structure, with significant disorder in the tail of the ligand and higher relative B-factors than RR-RJW100.

A major factor driving repositioning of the RJW100 enantiomers was the hydrogen bonding interaction made by the hydroxyl group. Although the contact with residue T352 is indirect, it is mediated by a water molecule that is part of a network of waters found in every published LRH-1 crystal structure (with the exception of PDB 4DOR, in which a major portion of the ligand binding pocket is disordered (20)). The existence of conserved water molecules, as well as their participation in ligand binding has been described (28-30). Thus, this interaction could serve as an anchor point to secure the compound in a predictable orientation, enabling the targeting of desired parts of the binding pocket via strategic addition of substituents to the ligand's scaffold. Moreover, replacing the RJW100 hydroxyl group with a larger polar moiety may allow direct contact with T352, leading to a stronger interaction. This strategy is being actively explored in our laboratory.

The role of the T352 interaction in LRH-1 activation by RR-RJW100 was demonstrated through the marked loss of activation by this compound when this residue was mutated (Figure 8). In addition, a RJW100 analog lacking a hydroxyl group and thus unable to make this was a poor activator. Unexpectedly, the T352V mutation also resulted in a loss of activity for GSK8470, although this compound does not interact with the T352-coordinated water molecule. However, we show that the T352V mutation weakens GSK8470's interaction with H390, perhaps via destabilization of the conserved water network (Figure 8G-H). This could be responsible for the loss of activity of GSK8470 when T352 is mutated.

It has been hypothesized in *in silico* studies that π-π-stacking with residue H390 is critical for activation of LRH-1 by this ligand class (25); however, this had not been explicitly tested. We find that this is the case for GSK8470, which stably interacts with H390 via face-to-face π-π-stacking. Interestingly, while mutation of H390 to alanine ablated LRH-1 activation by GSK8470, it had no effect on RR-RJW100-mediated activation (Figure 8). We show that the RR-RJW100 interaction with H390 is much less stable than that of GSK8470 and is mediated by a different phenyl ring (Figure 7). Substitution of the GSK8470 aniline group with the styrene appears to have had the unexpected effect of making face-to-face stacking with H390 less favorable (Figure 9). This, combined with the favorable, water-mediated interaction with T352, influences the positioning of RJW100.

Together, these findings reveal that the interaction of small molecule agonists with LRH-1 is more complex than originally supposed. Not only do these agonists affect receptor conformation differently than PL ligands, but they also exhibit an unexpected variability in binding modes. This work has uncovered some of the molecular interactions responsible for both positioning and activation of two very similar agonists, which provide insights into strategies to improve the design of LRH-1-targeted compounds.

Experimental Procedures

Materials and reagents—GSK8470, RJW100, and analogs were synthesized as previously described (22,23). RR-RJW100 was separated from SS-RJW100 by chiral preparative chromatography (Diacel OD-H column, (31)). *Endo*-RJW100 is a diastereoisomer of RJW100, previously referred to as "24-endo" (23). In this paper, we use "SR-RJW100" and "RS-RJW100" to refer to the enantiomers of *endo*-RJW100 (Figure 1). pCI empty vector was purchased from Promega. The SHP-luc and Renilla reporters, as well as pCI LRH-1, have been previously described (20). The vector for His-tagged tobacco etch virus (TEV) was a gift from John Tesmer (University of Texas at Austin). The pMSC7 (LIC_HIS) vector was provided by John Sondek (University of North Carolina at Chapel Hill). The Tif2 NR Box 3 peptide was purchased from RS Synthesis (Louisville, KY). DNA oligonucleotide primers were synthesized by IDT (Coralville, IA).

Protein expression and purification—LRH-1 LBD (residues 299-541) in the pMSC7 vector was expressed and purified as previously described (21). Briefly, protein was expressed in BL21(DE3) pLysS *E. coli* by induction with IPTG (1 mM) for 4 hr at 30°C. Protein was purified by nickel affinity chromatography. Protein used for DSF experiments was incubated with DLPC (five-fold molar excess) for four hours at room temperature,

and then repurified by size-exclusion into an assay buffer of 20 mM Tris-HCl, pH 7.5, 150 mM NaCl, and 5% glycerol. Protein used for crystallization was incubated with TEV protease to cleave the His tag. The cleaved protein was then separated from the His tag and TEV by a second round of nickel affinity chromatography. To make protein-ligand complexes, protein was incubated with ligands overnight (10-fold molar excess) and repurified by size-exclusion, using a final buffer of 100 mM ammonium acetate, pH 7.4, 150 mM sodium chloride, 1 mM DTT, 1 mM EDTA, and 2 mM CHAPS.

Crystallization—Protein-ligand complexes were incubated with a peptide derived from human Tif2 NR Box 3 (+H_3N-KENALLRYLLDKDDT-CO_2-) at four-fold molar excess for two hours at room temperature and then concentrated to 6.5 mg/ml. A crystallant of 0.05 M sodium acetate, pH 4.6, 5-11% PEG 4000 and 0-10% glycerol was used. Crystals were grown by hanging drop vapor diffusion in drops containing 1 µl protein and 1 µl crystallant, at a temperature of 18-20°C.

Structure determination—Crystals were flash frozen in liquid nitrogen, using a cryopreservative consisting of crystallant plus 30% glycerol. Data were collected remotely from the South East Regional Collaborative Access Team at the Advanced Photon Source, 22ID beamline (Argonne National Laboratories, Chicago, IL). Data were processed and scaled using HKL2000 (32) and phased by molecular replacement using Phaser-MR (Phenix (33)). For the RR-RJW100 structure, PDB 3PLZ (23) was used as the search model, with the ligand and a portion of the bottom of the receptor omitted. For the SR-RJW100 structure, the search model was the RR-RJW100 structure with the ligand omitted. Model building and refinement were conducted with Coot (34) and phenix.refine (33), respectively. Figures were constructed using Pymol (Schrödinger, LLC) (35).

Structure analysis—Dimensions of the binding pocket in the presence of various ligands were calculated using CastP software (36). Ligplot+ was used to identify residues interacting with the ligands (37).

Mutagenesis—Mutations were introduced to pMSC7 and pCI LRH-1 constructs using the Quikchange site-directed mutagenesis kit (Strategene). Constructs were sequence-verified prior to use.

Cell culture—HeLa cells were purchased from Atlantic Type Culture Collection and grown in phenol red-free MEMα media (CellGro) supplemented with 10% charcoal-stripped fetal bovine serum (Atlanta Biologicals). Cells were maintained using standard culture conditions.

Reporter gene assays—HeLa cells were seeded at a density of 10,000 cells per well in white-walled, clear-bottomed 96-well culture plates. The next day, cells were transfected with LRH-1 and reporters, using Fugene HD (Roche) at a ratio of 5:2 Fugene:DNA. The transfected plasmids included full-length LRH-1 in a pCI vector (5 ng/ well), and a SHP-luc reporter, encoding the LRH-1 response element and surrounding sequence from the SHP promoter cloned upstream of firefly luciferase in the pGL3 basic vector (50 ng/ well). Cells were also co-transfected with a constitutive *Renilla* luciferase reporter (utilizing the CMV promoter), which was used for normalization of firefly signal (1 ng/ well). Control cells received pCI empty vector at 5 ng/ well in place of LRH-1-pCI. Following an overnight transfection, cells were treated with agonists for 24 hours at the concentrations indicated in the figure legends. Agonists were dissolved in DMSO and then diluted into media, with a final concentration of 0.3% DMSO in all wells. Luciferase signal was quantified using the DualGlo kit (Promega). Experiments were conducted at least three times in triplicate.

Hydrogen-deuterium exchange mass spectrometry—Purified LRH-1 LBD protein (His tag removed) was incubated with of a five-fold molar excess DLPC or synthetic agonist overnight at 4°C. Protein-ligand complexes were then re-purified by size exclusion to remove displaced phospholipids and unbound ligands. An additional bolus of agonist or DLPC (5-fold molar excess) was added to the complexes prior to analysis by HDX. The assay buffer consisted of 20 mM Tris HCl, pH 7.4, 150 mM NaCl, and 5% glycerol. Solution-phase amide HDX was carried out with a fully automated system as described previously (38) Briefly, 5 µl of protein was diluted to 25 µl with D_2O-containing HDX buffer and incubated at 25 °C for 10 s, 30 s, 60 s, 900 s or 3,600 s. Following on exchange, back exchange was minimized and the protein was denatured by dilution to 50 µL in a low pH and low temperature buffer containing 0.1% (v/v) trifluoroacetic acid in

5 M urea (held at 1 °C). Samples were then passed across an immobilized pepsin column (prepared in house) at 50 µl min-1 (0.1% v/v TFA, 15 °C); the resulting peptides were trapped on a C8 trap cartridge (Hypersil Gold, Thermo Fisher). Peptides were then gradient-eluted (4% (w/v) CH_3CN to 40% (w/v) CH_3CN, 0.3% (w/v) formic acid over 5 min, 2 °C) across a 1 mm × 50 mm C18 HPLC column (Hypersil Gold, Thermo Fisher) and electrosprayed directly into an Orbitrap mass spectrometer (Q Exactive, Thermo Fisher). Data were processed with in-house software (39) and visualized with PyMOL (Schrödinger, LLC, (35)). To measure the difference in exchange rates, we calculated the average percent deuterium uptake for hLRH-1 LBD – RR-RJW100 complexes following 10, 30, 60, 900 and 3,600 s of on exchange. From this value, we subtracted the average percent deuterium uptake measured for the DLPC-hLRH-1 LBD complex. Negative perturbation values indicate exchange rates are slower for these regions within the RR-RJW100-LRH-1 complex relative to DLPC-bound LRH-1.

Differential scanning fluorimetry (DSF)—Purified LRH-1 LBD-His protein (0.2 mg/ml) was incubated overnight with 50 µM of each compound at 4°C. The final DMSO concentration in the reactions was 1%. SYPRO orange dye (Invitrogen) was then added at a 1:1000 dilution. Reactions were heated at a rate of 0.5°C per minute, using a StepOne Plus Real Time PCR System (ThermoFisher). Fluorescence was recorded at every degree using the ROX filter (602 nm). Data were analyzed by first subtracting baseline fluorescence (ligands + SYPRO with no protein) and then fitting the curves using the Bolzman equation (GraphPad Prism, v6) to determine the Tm.

Model construction for molecular dynamics simulations—Three crystal structures of LRH-1 LBD in complex with Tif2 were used to construct models for the simulations. These were (1) PDB 3PLZ, chains B and D (GSK8470 ligand), (2) PDB 5L11 (RR-RJW100 ligand), and (3) PDB 5SYZ (SR-RJW100 ligand). For consistency, the structures were modified at the N- and C-termini so that all contained residues 300-540 of LRH-1 and residues 742-751 of the Tif2 peptide. Missing residues within this protein sequence were added, as well as missing protein sidechains (3PLZ). The T352V mutation was introduced by mutating the sequence of WT LRH-1 in each of these structures. Finally, in a separate simulation, RR-RJW100 ligand was artificially reoriented in WT LRH-1 to be aligned with GSK8470, allowing a face-to-face π-π-stacking interaction of LRH-1 residue H390 with the aniline ring of RR-RJW100. Seven complexes in total were used in molecular dynamics simulations.

Molecular dynamics simulations—The complexes were solvated in an octahedral box of TIP3P water with a 10 Å buffer around the protein complex. Na^+ and Cl^- ions were added to neutralize the protein and achieve physiological conditions. All systems were set up using xleap in AmberTools (40) with the parm99-bsc0 forcefield (41). Parameters for all ligands (GSK8470, RR-RJW100, and SR-RJW100) were obtained using Antechamber (42) in AmberTools. All minimizations and simulations were performed with Amber14 (43). Systems were minimized with 5000 steps of steepest decent followed by 5000 steps of conjugate gradient minimization with 500 kcal/mol·Å^2 restraints on all atoms. Restraints were removed from all atoms excluding the atoms in both the ligand and the Tif2 peptide, and the previous minimization was repeated. The systems were heated from 0 to 300 K using a 100-ps run with constant volume periodic boundaries and 5 kcal/mol·Å^2 restraints on all protein and ligand atoms. Twelve ns of MD equilibration was performed with 10 kcal/mol·Å^2 restraints on protein and ligand atoms using the NPT ensemble. Restraints were reduced to 1 kcal/mol·Å^2 for an additional 10 ns of MD equilibration. Then restraints were removed and 200 ns production simulations were performed for each system in the NPT ensemble. A 2-fs timestep was used and all bonds between heavy atoms and hydrogens were fixed with the SHAKE algorithm (44). A cut-off distance of 10 Å was used to evaluate long-range electrostatics with Particle Mesh Ewald (PME) and for van der Waals forces. 10,000 evenly spaced frames were taken from each simulation for analysis. Analysis was performed with the CPPTRAJ module (45) of AmberTools. Statistical significance of differences in time π-π-stacking with residue H390 for GSK8470 *versus* the re-oriented RR-RJW100 was assessed using a Chi Squared test with Yate's correction (GraphPad Prism, v6).

Acknowledgements
This work was supported in part by NIH T32-GM008602 (SGM), NIH F31DK111171 (SGM), R01DK095750 (EAO) and an Emory Catalyst Grant to EAO and NTJ. RJW and JS thank GlaxoSmithKline for generous funding. The authors thank Sally Bloodworth (University of Southhampton, UK) for assistance with synthetic chemistry, Bradley Kossmann and Dr. Ivaylo Ivanov (Georgia State University) for helpful discussions about MDS, and Dr. Kay Diederichs (Universität Konstanz) for advice in model refinement.

Conflict of interest
No conflicts

Author contributions. SGM and EAO conceived the studies. SGM purified protein-ligand complexes, determined crystal structures and conducted DSF, mutagenesis, and cellular assays. CDO performed and analyzed molecular dynamics simulations. DG and PRG conducted HDX and analyzed results. RJW, JS, MCD, ARF, and NTJ synthesized GSK8470, RJW100, and analogs. SGM, CDO, RJW, NTJ, DG, PRG, and EAO wrote the manuscript.

References
1. Gu, P., Goodwin, B., Chung, A. C., Xu, X., Wheeler, D. A., Price, R. R., Galardi, C., Peng, L., Latour, A. M., Koller, B. H., Gossen, J., Kliewer, S. A., and Cooney, A. J. (2005) Orphan nuclear receptor LRH-1 is required to maintain Oct4 expression at the epiblast stage of embryonic development. *Molecular and cellular biology* **25**, 3492-3505
2. Wagner, R. T., Xu, X., Yi, F., Merrill, B. J., and Cooney, A. J. (2010) Canonical Wnt/beta-catenin regulation of liver receptor homolog-1 mediates pluripotency gene expression. *Stem cells* **28**, 1794-1804
3. Lu, T., Makishima, M., JJ, R., Schoonjans, K., Kerr, T., Auwerx, J., and DJ, M. (2000) Molecular Basis for Feedback Regulation of Bile Acid Synthesis by Nuclear Receptors. *Molecular Cell* **6**, 507-515
4. Stein, S., Oosterveer, M. H., Mataki, C., Xu, P., Lemos, V., Havinga, R., Dittner, C., Ryu, D., Menzies, K. J., Wang, X., Perino, A., Houten, S. M., Melchior, F., and Schoonjans, K. (2014) SUMOylation-dependent LRH-1/PROX1 interaction promotes atherosclerosis by decreasing hepatic reverse cholesterol transport. *Cell metabolism* **20**, 603-613
5. Schoonjans, K., Annicotte, J.-S., Huby, T., Botrugno, O., Fayard, E., Ueda, Y., Chapman, J., and Auwerx, J. (2002) Liver receptor homolog 1 controls the expression of the scavenger receptor class B type I. *EMBO Reports* **3**, 1181-1187
6. Bolado-Carrancio, A., Riancho, J. A., Sainz, J., and Rodriguez-Rey, J. C. (2014) Activation of nuclear receptor NR5A2 increases Glut4 expression and glucose metabolism in muscle cells. *Biochemical and biophysical research communications* **446**, 614-619
7. Oosterveer, M. H., Mataki, C., Yamamoto, H., Harach, T., Moullan, N., van Dijk, T. H., Ayuso, E., Bosch, F., Postic, C., Groen, A. K., Auwerx, J., and Schoonjans, K. (2012) LRH-1-dependent glucose sensing determines intermediary metabolism in liver. *The Journal of clinical investigation* **122**, 2817-2826
8. Venteclef, N., Smith, J. C., Goodwin, B., and Delerive, P. (2006) Liver receptor homolog 1 is a negative regulator of the hepatic acute-phase response. *Molecular and cellular biology* **26**, 6799-6807
9. Mamrosh, J. L., Lee, J. M., Wagner, M., Stambrook, P. J., Whitby, R. J., Sifers, R. N., Wu, S. P., Tsai, M. J., Demayo, F. J., and Moore, D. D. (2014) Nuclear receptor LRH-1/NR5A2 is required and targetable for liver endoplasmic reticulum stress resolution. *eLife* **3**, e01694

10. Bayrer, J. R., Mukkamala, S., Sablin, E. P., Webb, P., and Fletterick, R. J. (2015) Silencing LRH-1 in colon cancer cell lines impairs proliferation and alters gene expression programs. *Proceedings of the National Academy of Sciences of the United States of America* **112**, 2467-2472
11. Bianco, S., Brunelle, M., Jangal, M., Magnani, L., and Gevry, N. (2014) LRH-1 governs vital transcriptional programs in endocrine-sensitive and -resistant breast cancer cells. *Cancer research* **74**, 2015-2025
12. Chand, A. L., Herridge, K. A., Thompson, E. W., and Clyne, C. D. (2010) The orphan nuclear receptor LRH-1 promotes breast cancer motility and invasion. *Endocrine-related cancer* **17**, 965-975
13. Clyne, C. D., Kovacic, A., Speed, C. J., Zhou, J., Pezzi, V., and Simpson, E. R. (2004) Regulation of aromatase expression by the nuclear receptor LRH-1 in adipose tissue. *Molecular and cellular endocrinology* **215**, 39-44
14. Lai, C. F., Flach, K. D., Alexi, X., Fox, S. P., Ottaviani, S., Thiruchelvam, P. T., Kyle, F. J., Thomas, R. S., Launchbury, R., Hua, H., Callaghan, H. B., Carroll, J. S., Charles Coombes, R., Zwart, W., Buluwela, L., and Ali, S. (2013) Co-regulated gene expression by oestrogen receptor alpha and liver receptor homolog-1 is a feature of the oestrogen response in breast cancer cells. *Nucleic acids research* **41**, 10228-10240
15. Thiruchelvam, P. T., Lai, C. F., Hua, H., Thomas, R. S., Hurtado, A., Hudson, W., Bayly, A. R., Kyle, F. J., Periyasamy, M., Photiou, A., Spivey, A. C., Ortlund, E. A., Whitby, R. J., Carroll, J. S., Coombes, R. C., Buluwela, L., and Ali, S. (2011) The liver receptor homolog-1 regulates estrogen receptor expression in breast cancer cells. *Breast cancer research and treatment* **127**, 385-396
16. Lin, Q., Aihara, A., Chung, W., Li, Y., Huang, Z., Chen, X., Weng, S., Carlson, R. I., Wands, J. R., and Dong, X. (2014) LRH1 as a driving factor in pancreatic cancer growth. *Cancer letters* **345**, 85-90
17. Lee, J. M., Lee, Y. K., Mamrosh, J. L., Busby, S. A., Griffin, P. R., Pathak, M. C., Ortlund, E. A., and Moore, D. D. (2011) A nuclear-receptor-dependent phosphatidylcholine pathway with antidiabetic effects. *Nature* **474**, 506-510
18. Sablin, E. P., Blind, R. D., Uthayaruban, R., Chiu, H. J., Deacon, A. M., Das, D., Ingraham, H. A., and Fletterick, R. J. (2015) Structure of Liver Receptor Homolog-1 (NR5A2) with PIP hormone bound in the ligand binding pocket. *Journal of structural biology*
19. Krylova, I. N., Sablin, E. P., Moore, J., Xu, R. X., Waitt, G. M., MacKay, J. A., Juzumiene, D., Bynum, J. M., Madauss, K., Montana, V., Lebedeva, L., Suzawa, M., Williams, J. D., Williams, S. P., Guy, R. K., Thornton, J. W., Fletterick, R. J., Willson, T. M., and Ingraham, H. A. (2005) Structural analyses reveal phosphatidyl inositols as ligands for the NR5 orphan receptors SF-1 and LRH-1. *Cell* **120**, 343-355
20. Musille, P. M., Pathak, M. C., Lauer, J. L., Hudson, W. H., Griffin, P. R., and Ortlund, E. A. (2012) Antidiabetic phospholipid-nuclear receptor complex reveals the mechanism for phospholipid-driven gene regulation. *Nature structural & molecular biology* **19**, 532-537, S531-532
21. Musille, P. M., Kossmann, B. R., Kohn, J. A., Ivanov, I., and Ortlund, E. A. (2015) Unexpected Allosteric Network Contributes to LRH-1 Coregulator Selectivity. *The Journal of biological chemistry*
22. Whitby, R. J., Dixon, S., Maloney, P. R., Delerive, P., Goodwin, B., Parks, D. J., and Wilson, T. M. (2006) Identification of Small Molecule Agonists of the Orphan Nuclear Receptros Liver Receoptor Homolog-1 and Steroidogenic Factor-1. *Journal of medicinal chemistry* **49**, 6652-6655
23. Whitby, R. J., Stec, J., Blind, R. D., Dixon, S., Leesnitzer, L. M., Orband-Miller, L. A., Williams, S. P., Willson, T. M., Xu, R., Zuercher, W. J., Cai, F., and Ingraham, H. A. (2011) Small molecule agonists of the orphan nuclear receptors steroidogenic factor-1 (SF-1, NR5A1) and liver receptor homologue-1 (LRH-1, NR5A2). *Journal of medicinal chemistry* **54**, 2266-2281

24. Li, Y., Lambert, M. H., and Xu, H. E. (2003) Activation of Nuclear Receptors. *Structure* **11**, 741-746
25. Lalit, M., Gangwal, R. P., Dhoke, G. V., Damre, M. V., Khandelwal, K., and Sangamwar, A. T. (2013) A combined pharmacophore modeling, 3D-QSAR and molecular docking study of substituted bicyclo-[3.3.0]oct-2-enes as liver receptor homolog-1 (LRH-1) agonists. *Journal of Molecular Structure* **1049**, 315-325
26. Zhang, Z., Chen, H., Bai, H., and Lai, L. (2007) Molecular dynamics simulations on the oligomer-formation process of the GNNQQNY peptide from yeast prion protein Sup35. *Biophys J* **93**, 1484-1492
27. Schrodinger, L. (2016) In the Ligand Interaction Diagram, what are the maximum lengths of hydrogen bonds, pi-pi stacking interactions, pi-cation interactions and metal coordination interactions that are identified and displayed? in *Schrodinger Knowledge Base*
28. Ogata, K., Wodak, S.J. (2002) Conserved water molecules in MHC class-I molecules and their structural and functional roles. *Protein Engineering* **15**, 697-705
29. Barillari, C., Taylor, J., Viner, R., Essex, J.W. (2007) Cassification of Water Molecules in Protein Binding Sites. *J Am Chem Soc* **129**, 2577-2587
30. Klebe, G. (2015) Applying thermodynamic profiling in lead finding and optimization. *Nat Rev Drug Discov* **14**, 95-110
31. Stec, J. (2010) *Tandem Reaction Sequences on a Zirconocene Template*Ph.D. Thesis, University of Southhampton
32. Otwinowski, Z., Minor, W. (1997) Processing of X-Ray Diffraction Data Collected in Oscillation Mode. *Methods in Enzymology* **276**, 307-326
33. Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta crystallographica. Section D, Biological crystallography* **66**, 213-221
34. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta crystallographica. Section D, Biological crystallography* **60**, 2126-2132
35. Schrodinger, LLC. (2010) The PyMOL Molecular Graphics System, Version 1.3r1.
36. Binkowski, T. A. (2003) CASTp: Computed Atlas of Surface Topography of proteins. *Nucleic acids research* **31**, 3352-3355
37. Laskowski, R. A., and Swindells, M. B. (2011) LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. *J Chem Inf Model* **51**, 2778-2786
38. Goswami, D., Devarakonda, S., Chalmers, M.J., Pascal, B.D., Spiegelman, B.M., Griffin, P.R. (2013) Time Window Expansion for HDX Analysis of an Intrinsically Disordered Protein. *J. Am. Soc. Mass Spectrom.* **24**, 1584-1592
39. Pascal, B. D., Willis, S., Lauer, J. L., Landgraf, R. R., West, G. M., Marciano, D., Novick, S., Goswami, D., Chalmers, M. J., and Griffin, P. R. (2012) HDX workbench: software for the analysis of H/D exchange MS data. *J Am Soc Mass Spectrom* **23**, 1512-1521
40. Case, D., and Kollman, P. (2012) AmberTools 12. *University of California, San Francisco*
41. Pérez, A., Marchán, I., Svozil, D., Sponer, J., Cheatham, T. E., Laughton, C. A., and Orozco, M. (2007) Refinement of the AMBER force field for nucleic acids: improving the description of α/γ conformers. *Biophysical journal* **92**, 3817-3829
42. Wang, J., Wang, W., Kollman, P. A., and Case, D. A. (2001) Antechamber: an accessory software package for molecular mechanical calculations. *J. Am. Chem. Soc* **222**, U403
43. Case, D., Babin, V., Berryman, J., Betz, R., Cai, Q., Cerutti, D., Cheatham Iii, T., Darden, T., Duke, R., and Gohlke, H. (2014) Amber 14.
44. Ryckaert, J.-P., Ciccotti, G., and Berendsen, H. J. (1977) Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *Journal of Computational Physics* **23**, 327-341

45. Roe, D. R., and Cheatham III, T. E. (2013) PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data. *Journal of Chemical Theory and Computation* **9**, 3084-3095
46. Ortlund, E. A., Lee, Y., Solomon, I. H., Hager, J. M., Safi, R., Choi, Y., Guan, Z., Tripathy, A., Raetz, C. R., McDonnell, D. P., Moore, D. D., and Redinbo, M. R. (2005) Modulation of human nuclear receptor LRH-1 activity by phospholipids and SHP. *Nature structural & molecular biology* **12**, 357-363
47. Benod, C., Carlsson, J., Uthayaruban, R., Hwang, P., Irwin, J. J., Doak, A. K., Shoichet, B. K., Sablin, E. P., and Fletterick, R. J. (2013) Structure-based discovery of antagonists of nuclear receptor LRH-1. *The Journal of biological chemistry* **288**, 19830-19844

Table 1: X-ray data collection and refinement statistics.

Data collection	LRH-1 + RR-RJW100 +Tif2	LRH-1 + SR-RJW100 + Tif2
Space group	P422*	$P4_32_12$
Cell dimensions□□		
a, *b*, *c* (Å)	46.4 46.4, 220.5	46.3, 46.3, 220.0
□ α,β,γ□ (°)	90, 90, 90	90, 90, 90
Resolution (Å)	50 – 1.85 (1.92-1.85)	50 – 1.93 (2.00-1.93)
R_{pim}	0.03 (0.22)	0.02 (0.231)
$I / \sigma I$	28.5 (1.98)	33.5 (2.06)
CC1/2 in highest shell	0.878	0.861
Completeness (%)	96.8 (86.1)	99.7 (97.9)
Redundancy	8.9 (6.8)	11.1 (8.2)
Refinement		
Resolution (Å)	1.85	1.93
No. reflections	21026	19073
R_{work} / R_{free} (%)	19.8 / 24.1	20.5 / 22.8
No. atoms		
Protein	4038	4062
Water	92	51
B-factors		
Protein	38.1	46.9
Ligand	34.8	59.2
Water	36.8	43.5
R.m.s. deviations		
Bond lengths (Å)	0.004	0.004
Bond angles (°)	0.578	0.588
Ramachandran favored (%)	98	98
Ramachandran outliers (%)	0	0
PDB accession code	5L11	5SYZ

*Data were processed in P422 but refined in $P4_32_12$
Values in parentheses indicate highest resolution shell

Table 2. Binding pocket dimensions of LRH-1 in various liganded states

PDB ID	Ligand	Width (Å)	Volume (Å^3)
5SYZ	SR-RJW100	7.2	1083.0
3PLZ	GSK8470	7.3	957.6
5L11	RR-RJW100	7.4	1061.1
1PLD	apo	8.0	965.8
1PLE	*E. coli* lipids	10.0	1368.7
4DOS	DLPC	10.3	1553.9
4RWV	PIP3	12.0	1233.3

Table 3. B-factors for waters near the ligand and their ligating atoms.

PDB 5L11

Water	Ligating atom 1	Ligating atom 2	Ligating atom 3	Ligating atom 4*
W1 (22.5)	RJW, *O* (30.8)	T352, *OG1* (22.9)	W2 (23.3)	W5 (22.8)
W2 (23.2)	V406, *N* (24.2)	W1 (22.5)	W9 (24.3)	--
W5 (22.8)	L386, *O* (25.2)	D389, *OD2* (28.6)	W1 (22.5)	--
W9 (24.3)	H390, *ND1* (23.1)	R393, *NH1* (28.6)	W2 (23.3)	--

PDB 5SYZ

Water	Ligating atom 1	Ligating atom 2	Ligating atom 3	Ligating atom 4*
W17 (34.5)	ENO, *O* (58.8)	T352, *OG1* (28.9)	W16 (25.9)	W4 (32.5)
W16 (25.9)	L386, *O* (26.3)	D389, *OD2* (34.3)	W17	--
W3 (32.0)	H390, *ND1* (30.7)	R393, *NH1* (33.0)	F404, *O* (31.9)	W4
W4 (32.5)	ENO, *O* (58.8)	V406, *N* (32.8)	W3	--

PDB 3PLZ, Chain A

Water	Ligating atom 1	Ligating atom 2	Ligating atom 3	Ligating atom 4*
W41 (19.5)	T352, *OG1* (14.4)	R393, *NH2* (21.5)	W16 (18.1)	W13
W16 (18.1)	D389, *OD2* (19.4)	R393, *NH2* (21.5)	W41 (19.4)	
W13(21.2)	V406, *N* (20.4)	W41	W31	
W31 (21.1)	F404, *O* (19.2)	R393, *NE* (22.0)	H390, *ND1* (15.6)	W13

PDB 3PLZ, Chain B

Water	Ligating atom 1	Ligating atom 2	Ligating atom 3	Ligating atom 4*
W25 (15.9)	T352, *OG1* (11.7)	W14 (13.8)	W129 (14.2)	
W14 (13.8)	L386, *O* (13.8)	D389, *OD2* (18.7)	W25 (15.9)	W52 (18.1)
W129 (14.2)	V406, *N* (20.4)	W25 (15.9)	W52 (18.1)	
W52 (18.1)	F404, *O* (13.6)	R393, *NH2* (22.7)	H390, *ND1* (12.1)	

*When a fourth ligating atom is identified, it is likely that the waters intermittently interact with the four atoms.

Figure legends

Figure 1. Chemical structures of LRH-1 agonists. A. GSK8470, the parent compound. B. RJW100 enantiomers. C. RJW100 analog lacking the hydroxyl group (named 18a), assayed in Figure 8.

Figure 2. Crystal structure of LRH-1 ligand binding domain bound to RR-RJW100 and a fragment of the Tif2 coactivator. A. Overall structure, with α-helices shown in light blue and β-sheets in slate. The Tif2 peptide (green) is bound at the AFS. The ligand (yellow) is bound at a single site in the binding pocket. *Dotted line*, region of disorder in the protein backbone that could not be modeled. B. Omit map (F_O-F_C, contoured at 2.5 σ) showing that a single enantiomer of RJW100 is bound in the structure. C and D. Superposition of RR-RJW100 (yellow) with the ligand coordinates from (*C)* DLPC (purple, PDB 4DOS) or *(D)* PIP3 (blue, PDB 4RWV) show the very different binding mode of RR-RJW100 compared to the PL ligands. E. DLPC expands the width at the mouth of the pocket by approximately 3 Å compared to RR-RJW100. The width was measured from T341 to N419 (alpha carbons). F. Superposition of four PL-bound LRH-1 crystal structures (PDBs 4DOS (20), 1YUC (46), 4RWV (18), and 4PLE (21), purple) and three structures of LRH-1 bound to synthetic agonists (PDB 3PLZ (23) and the two structures from this paper, grey) showing the shift of helix 6 by the PL ligands. One PL ligand is shown to illustrate that the shift of H6 is likely due to a stearic clash with the PL headgroup. The number in the top right corner indicates the average distance that H6 shifts when PLs are bound relative to LRH-1

in the apo state or when synthetic ligands are bound (+/- SEM, determined by measuring distances between S418 alpha carbons). The apo LRH-1 structure used for measurements was PDB 4PLD (21).

Figure 3. Differential effects on LRH-1 dynamics when RR-RJW100 is bound *versus* DLPC. A. HDX was used to probe differential effects on protein dynamics by the two ligands. The scale refers to the difference in percent deuterium incorporation for RR-RJW100-bound LRH-1 minus DLPC-bound LRH-1). The scale reflects the difference in average percent deuterium incorporation for RR-RJW100-*versus* DLPC- bound LRH-1 LBD. For example, negative numbers reflect slower deuterium incorporation (less motion) for RR-RJW100 *versus* DLPC. Results are mapped onto PDB 4DOS (20). B. DSF curves showing that RJW100 increases overall LRH-1 thermostability compared to DLPC. Each point represents the mean +/- SEM of values for three independent experiments, each conducted in triplicate.

Figure 4. A very different binding mode of RR-RJW100 compared to a closely related synthetic agonist. A. Close-view of the bottom of the receptor, showing the shift in H3 induced by both synthetic agonists compared to Apo LRH-1. B. Superposition of coordinates for GSK8470 (cyan, from PBD 3PLZ) and RR-RJW100 (yellow). C. The RR-RJW100 hydroxyl group was predicted to interact with residues H390 and R393, but it is over 6 Å away from these residues in our structure.

Figure 5. Crystal structure of LRH-1 bound to an RJW100 enantiomer with *endo* stereochemistry. A. The overall structure of the LRH-1 ligand binding domain with SR-RJW100 (violet) and a peptide derived from the Tif2 coactivator (green). *Dotted line*, region of disorder in the protein backbone that could not be modeled. B. Omit map (F_O-F_C, contoured at 2.5 σ) showing that a single enantiomer of RJW100 is bound in the structure. C. Superposition of SR-RJW100 with RR-RJW100 (yellow) showing a very similar position in the binding pocket.

Figure 6. Residues interacting with LRH-1 agonists. Close views of the binding pockets from the structures of LRH-1 bound to (*A)* RR-RJW100 or (*B)* SR-RJW100, depicting sidechains of amino acid residues that interact with each ligand. In *A*, residues that also interact with GSK8470 are shown in cyan, while unique interactions made by RR-RJW100 are shown in grey. Portions of the electron density maps are shown to highlight the interactions with T352 through water (F_O-F_C, contoured to 1σ).

Figure 7. π-π-stacking with residue H390 differs among agonists. A-C. *Left*, views of the different types of π-π stacking utilized by *(A)* GSK8470, *(B)* RR-RJW100, and *(C)* SR-RJW100. *Right*, MDS monitoring the distances between ring centroids (x-axis) and angle between the ring planes (y-axis) for the ligand phenyl group and H390 at each time increment of the 200 ns MDS. The red and blue boxes indicate when face-to-face and edge-to-face π-π stacking occurred, respectively. The numbers in the top right corners indicate the percentage of time spent π-π stacking during the MDS.

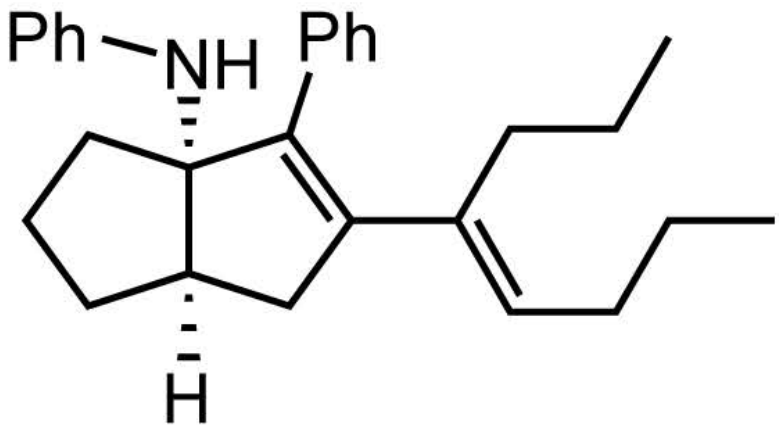
Figure 8. Importance of protein-ligand interactions on ligand binding and activity. A. The analog 18a, lacking the hydroxyl group, does not stabilize wild-type LRH-1 in DSF assays. B. Introduction of the T352V mutation to LRH-1 ablates the stabilizing effects of RR-RJW00 and Endo-RJW100. Purified LRH-1 LBD, initially bound to DLPC for homogeneity, was incubated with either DMSO (control) or synthetic agonist dissolved in DMSO. C. Compound 18a, is a significantly weaker agonist in luciferase reporter assays. D-F. Luciferase reporter assays measuring LRH-1 activity, using the SHP-luc reporter. Values have been normalized to constitutive Renilla luciferase signal and are presented as fold change *versus* wild-type LRH-1 + DMSO. The A349F mutation introduces a bulky aromatic side chain, which blocks the binding pocket and prevents binding of synthetic ligands (47). This was used as a negative control. G. Snapshots from MDS using T352V LRH-1 with GSK8470 bound. H. Plot of distances between ring centroids (x-axis) and angle between the ring planes (y-axis) for the GSK8470 phenyl group and H390 at each time increment of the 200 ns with T352V LRH-1 (as described for Figure 7). For *A-F*, each bar (or point, for panels D-F) represents the mean +/- SEM for three independent experiments, each

conducted in triplicate. * $p < 0.05$ (significance was determined by two-way ANOVA followed by Tukey's multiple comparisons test).

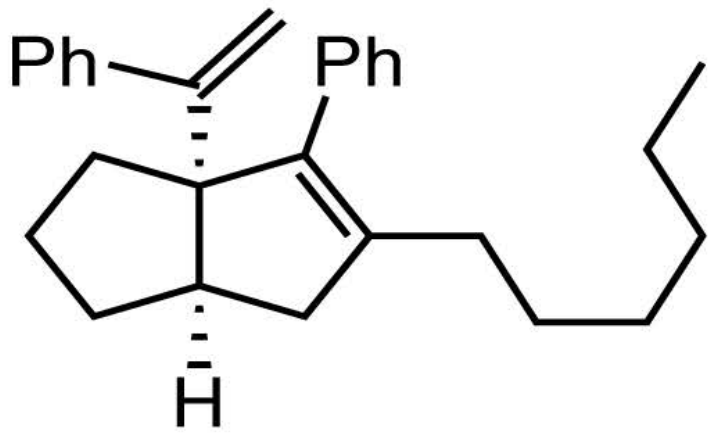
Figure 9. Impaired π-π stacking of the RR-RJW100 styrene group in MDS. A. RR-RJW100 (yellow) was artificially aligned with GSK8470 (cyan) for MDS to study its ability to π-π-stack with residue H390. Red dotted lines indicate water molecules that are within hydrogen bonding distances of the ligand or residue T352. B. Plot of distances between ring centroids (x-axis) and angle between the ring planes (y-axis) for the ligand phenyl group and H390 at each time increment of the 200 ns MDS for the styrene phenyl ring of repositioned RR-RJW100 and residue H390 over the course of the simulation. The red and blue boxes indicate when face-to-face and edge-to-face π-π stacking occurred, respectively. The numbers in the top right corners indicate the percentage of time π-π stacking during the MDS.

Figure 1

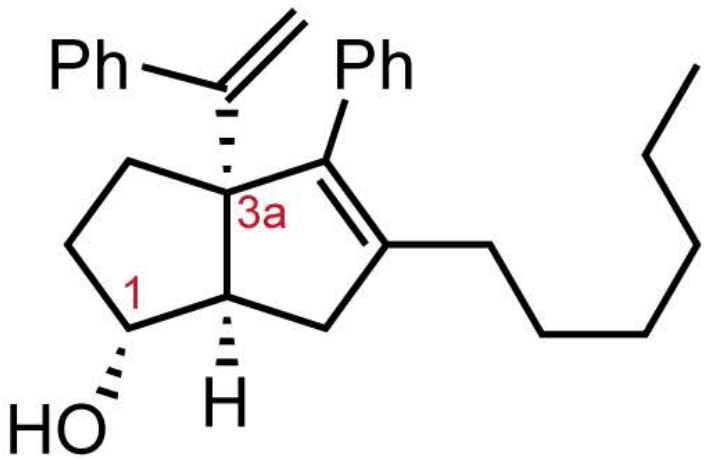
A



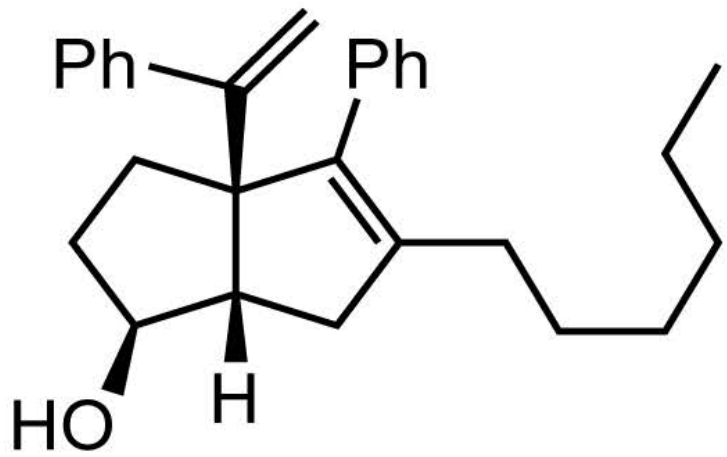
C



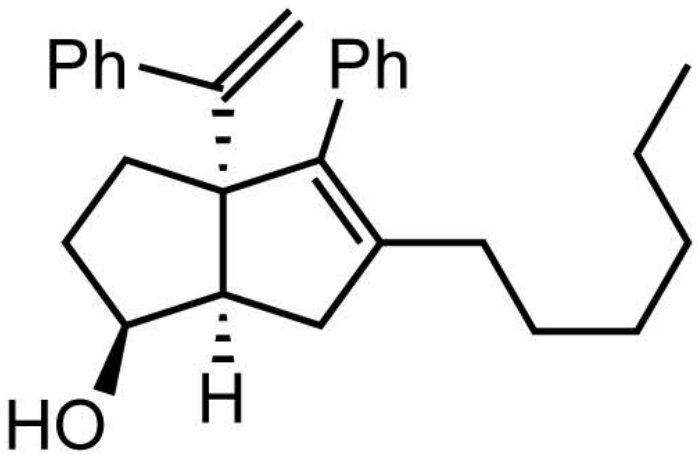
B



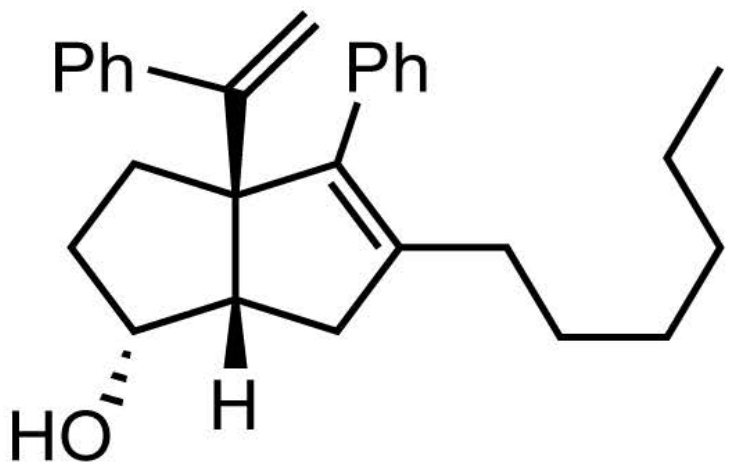
Exo (RR-RJW100)



Exo (SS-RJW100)



Endo (SR-RJW100)



Endo (RS-RJW100)

Figure 2

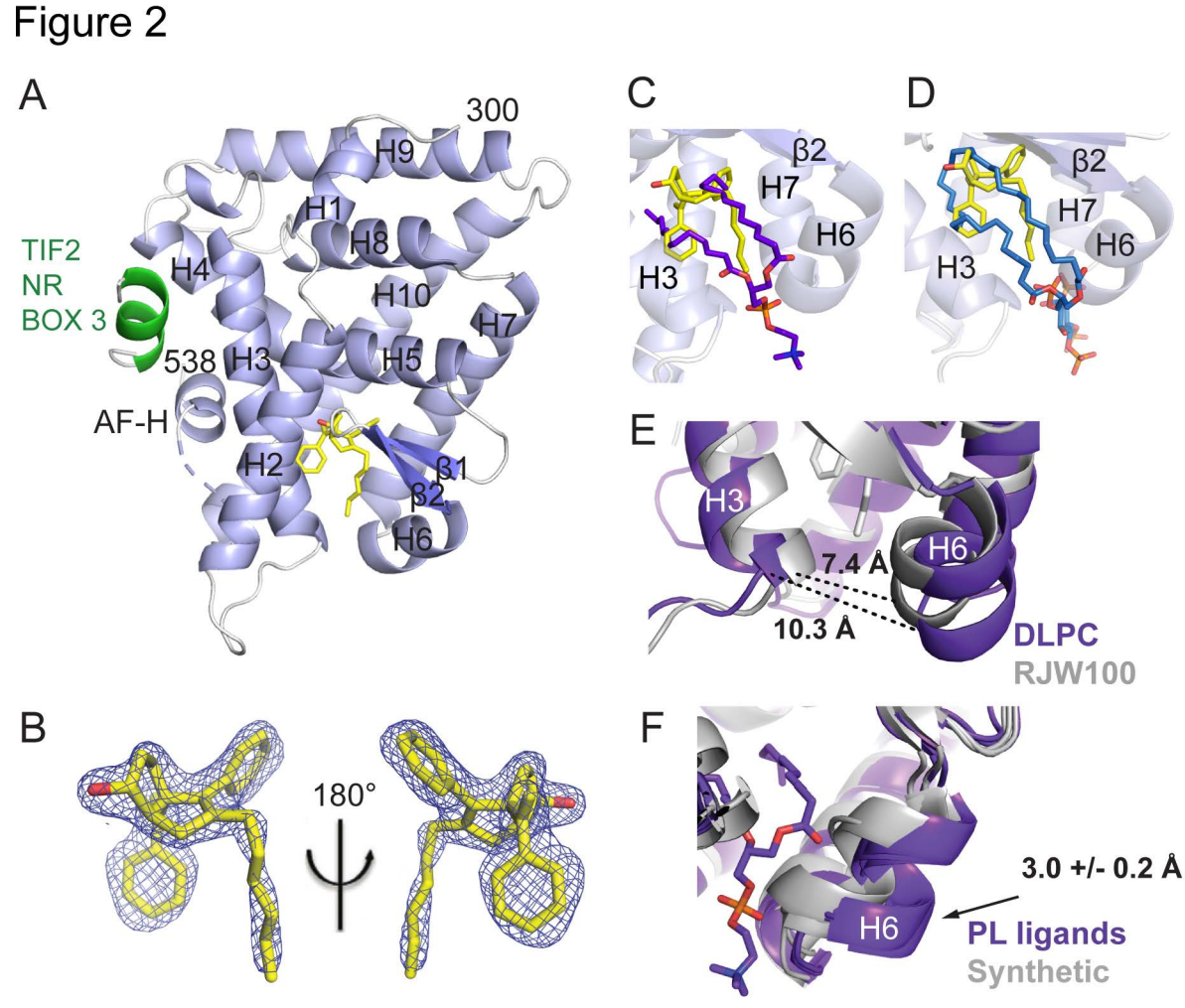

A
300
H9
H1
H8
TIF2
NR
BOX 3
H4
H10
H7
538
H3
H5
AF-H
H2
β1
β2
H6
B
180°
C
β2
H7
H6
H3
D
β2
H7
H6
H3
E
H3
H6
7.4 Å
10.3 Å
DLPC
RJW100
F
3.0 +/- 0.2 Å
H6
PL ligands
Synthetic

Figure 3

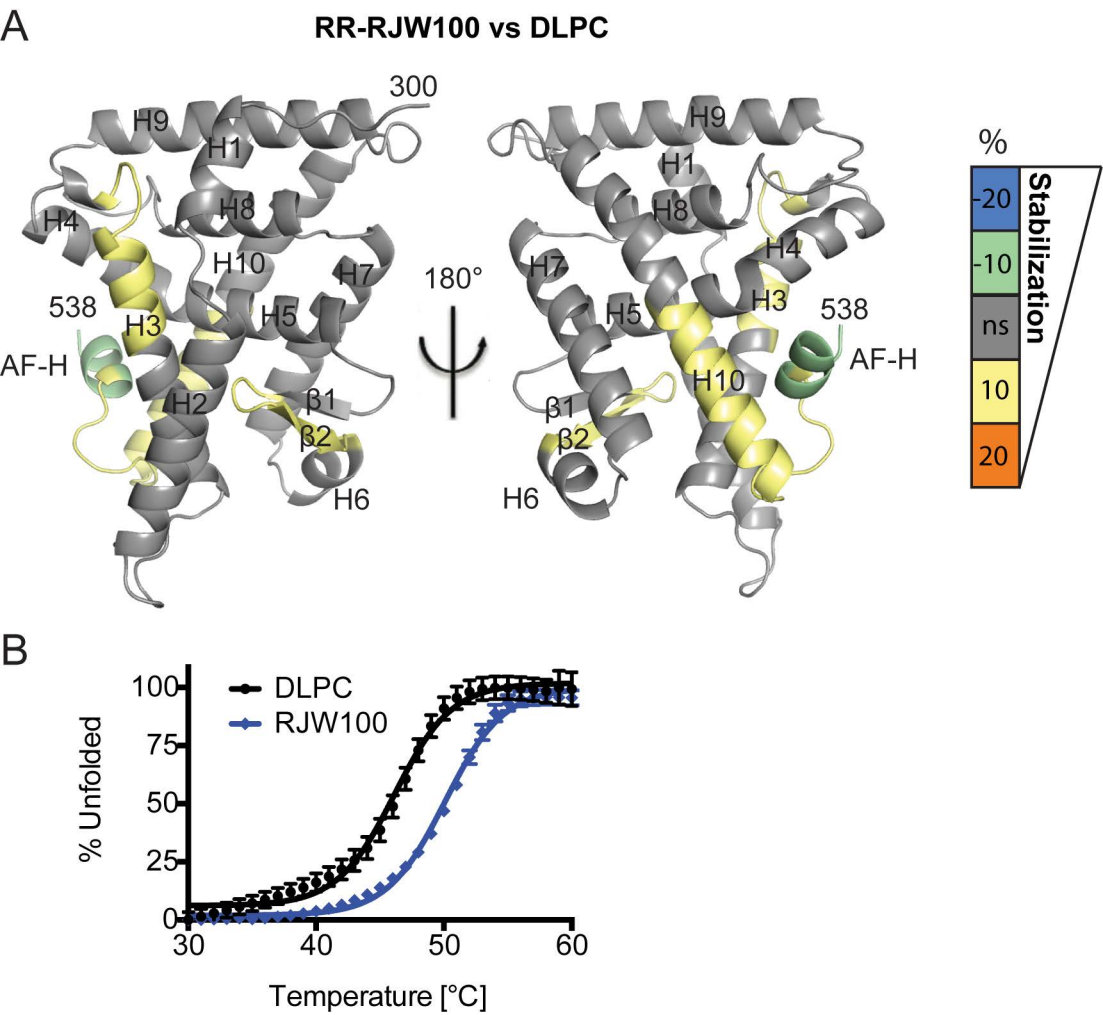

A
RR-RJW100 vs DLPC
300
H9
H1
H8
H4
H10
H7
538
H3
H5
AF-H
H2
β1
β2
H6
180°
%
-20
-10
ns
10
20
Stabilization
B
100
75
50
25
0
% Unfolded
DLPC
RJW100
30
40
50
60
Temperature [°C]

Figure 4

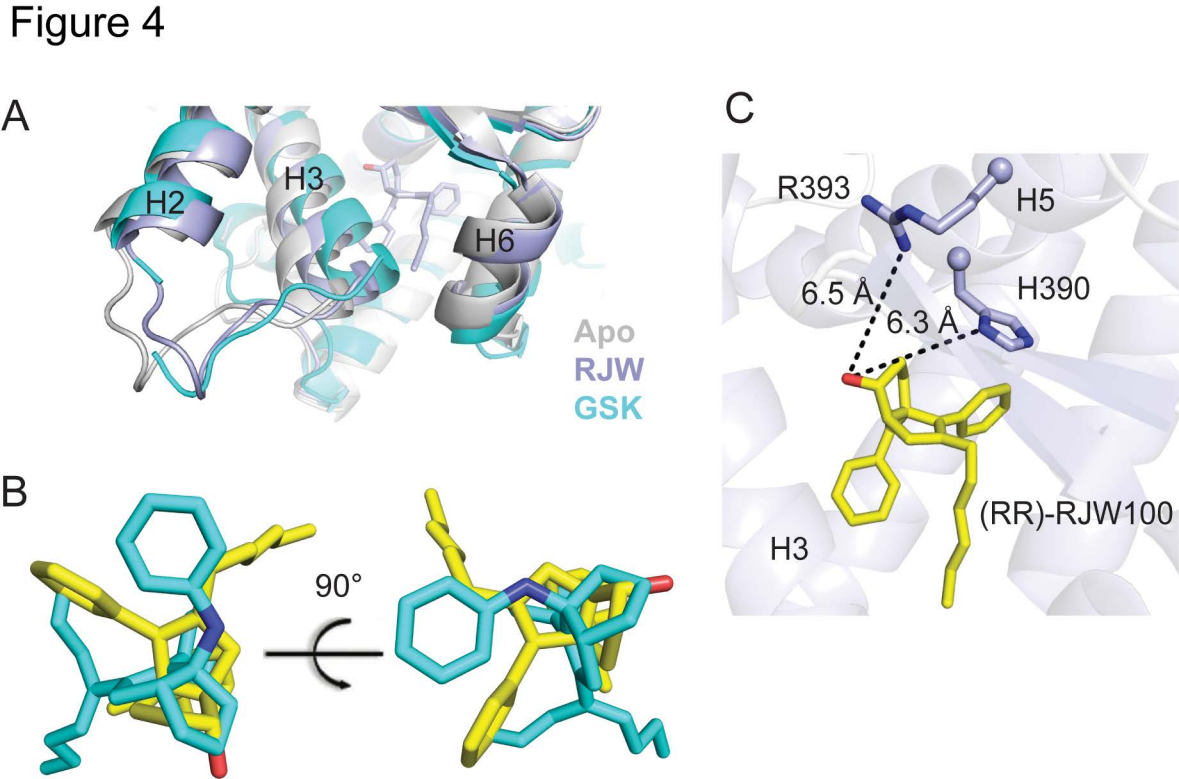


Figure 5

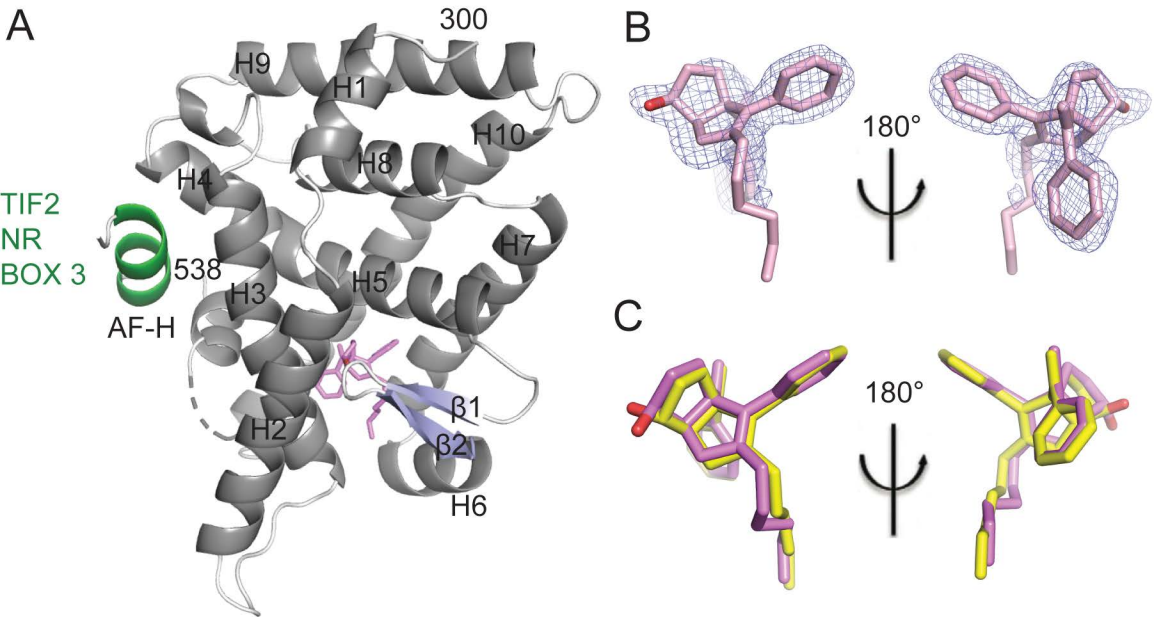

A
300
H9
H1
H10
H8
H4
TIF2
NR
BOX 3
538
H7
H3
H5
AF-H
β1
H2
β2
H6
B
180°
C
180°

Figure 6

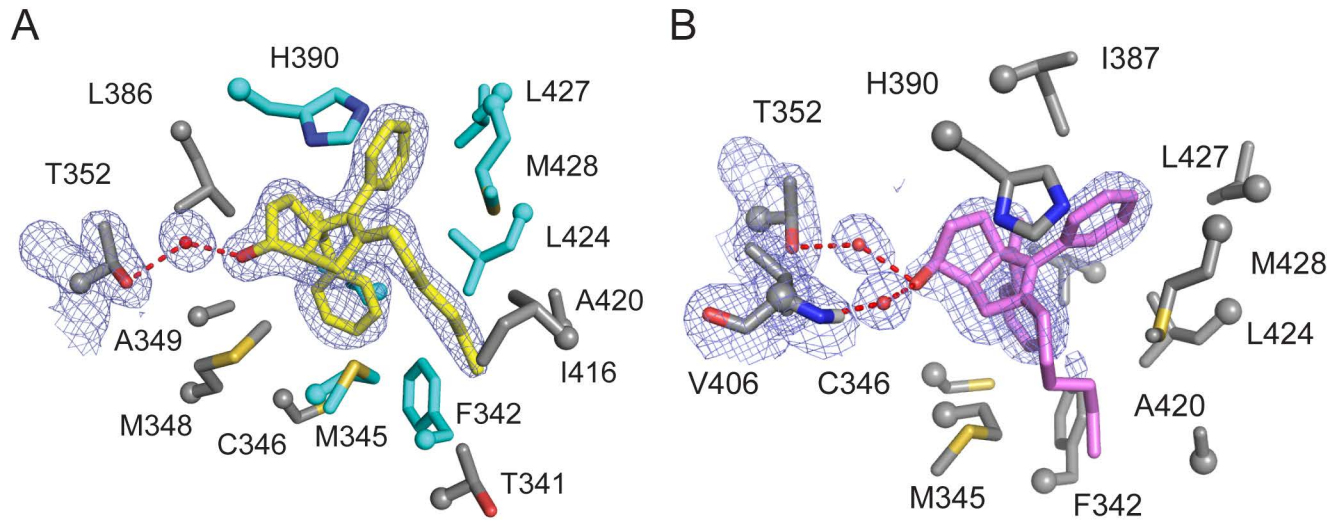

A
H390
L386
L427
T352
M428
L424
A420
A349
I416
M348
C346
M345
F342
T341
B
I387
H390
T352
L427
M428
L424
V406
C346
A420
M345
F342

Figure 7

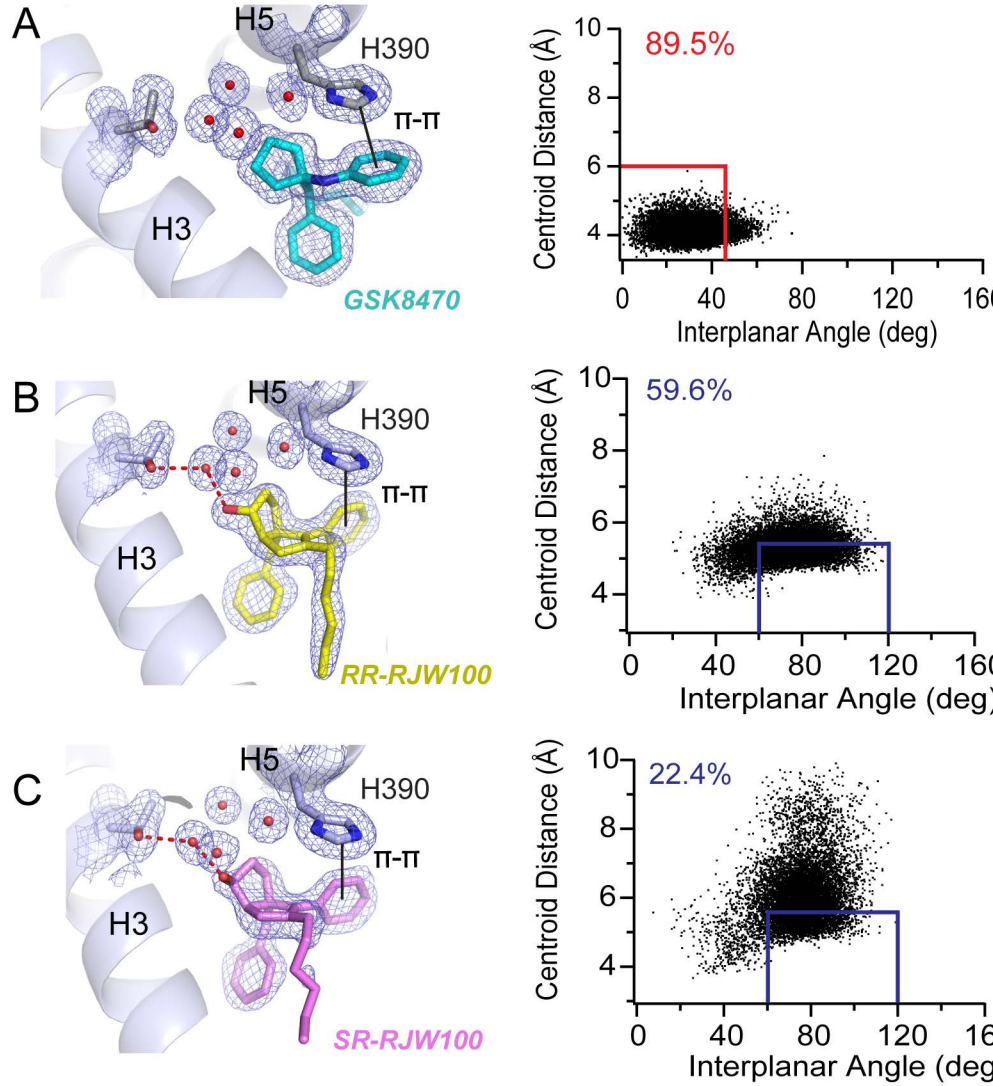

A
H5
H390
π-π
H3
GSK8470
Centroid Distance (Å)
89.5%
Interplanar Angle (deg)
B
H5
H390
π-π
H3
RR-RJW100
59.6%
C
H5
H390
π-π
H3
SR-RJW100
22.4%

Figure 8

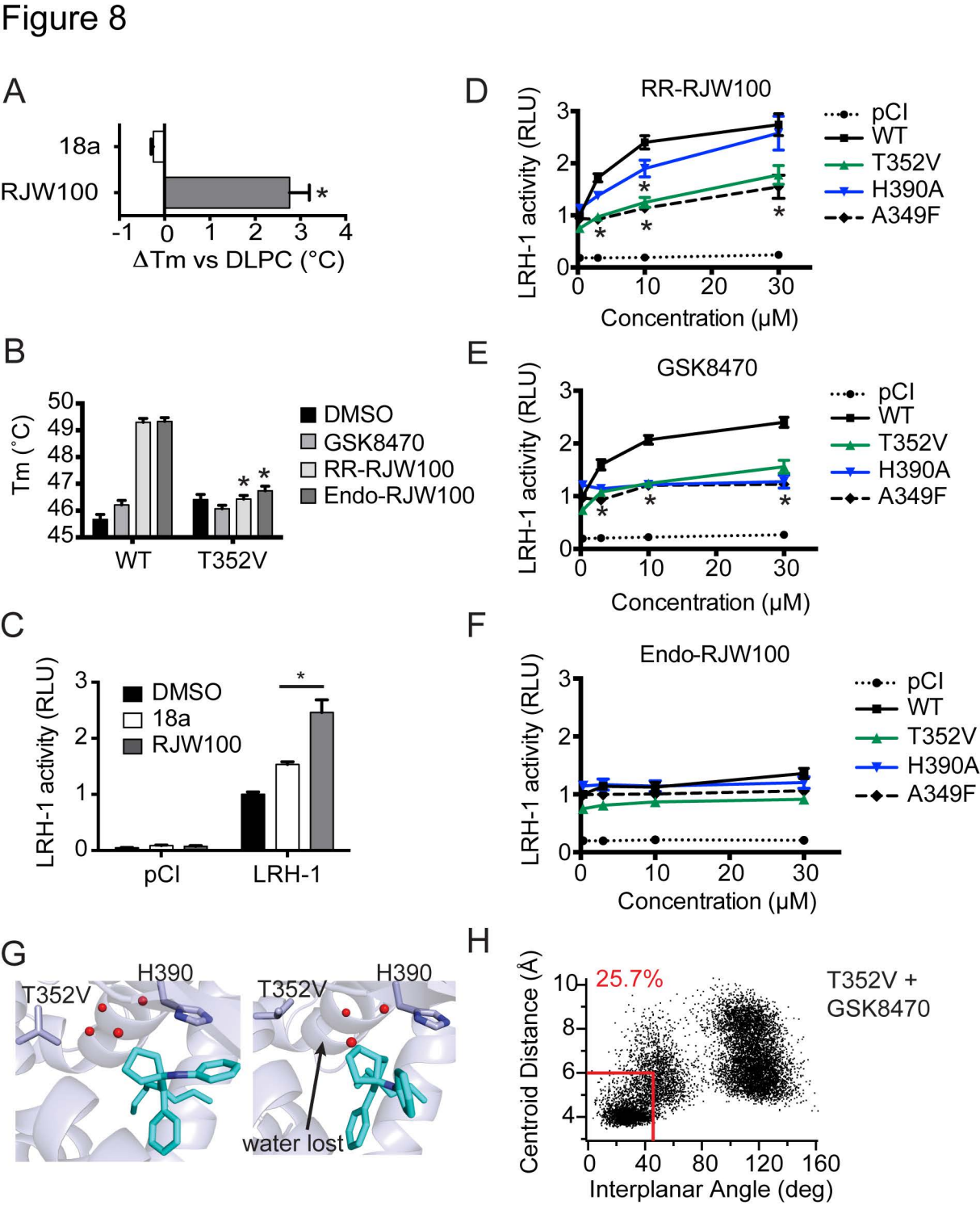
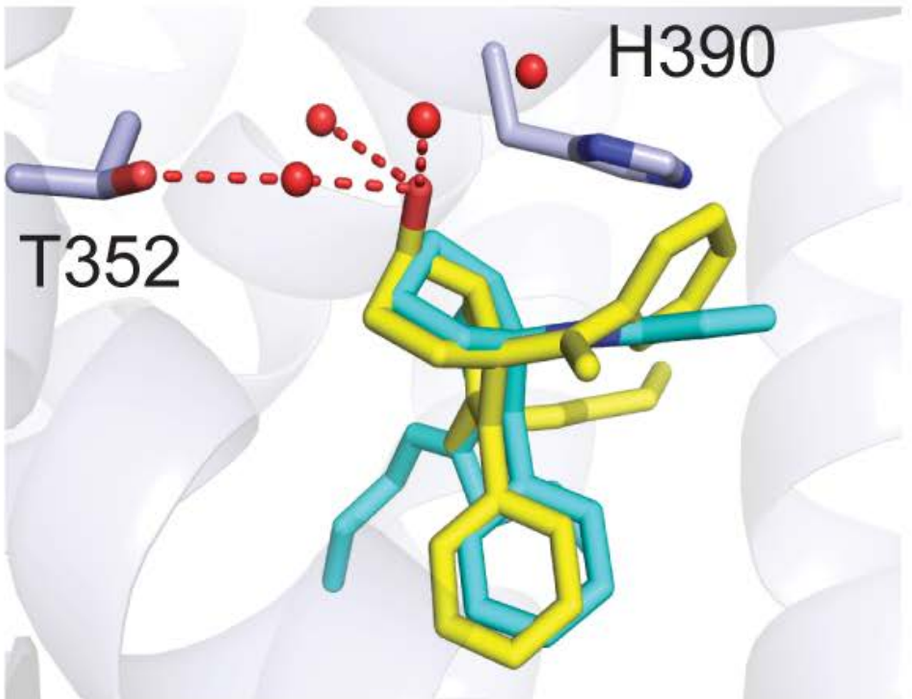
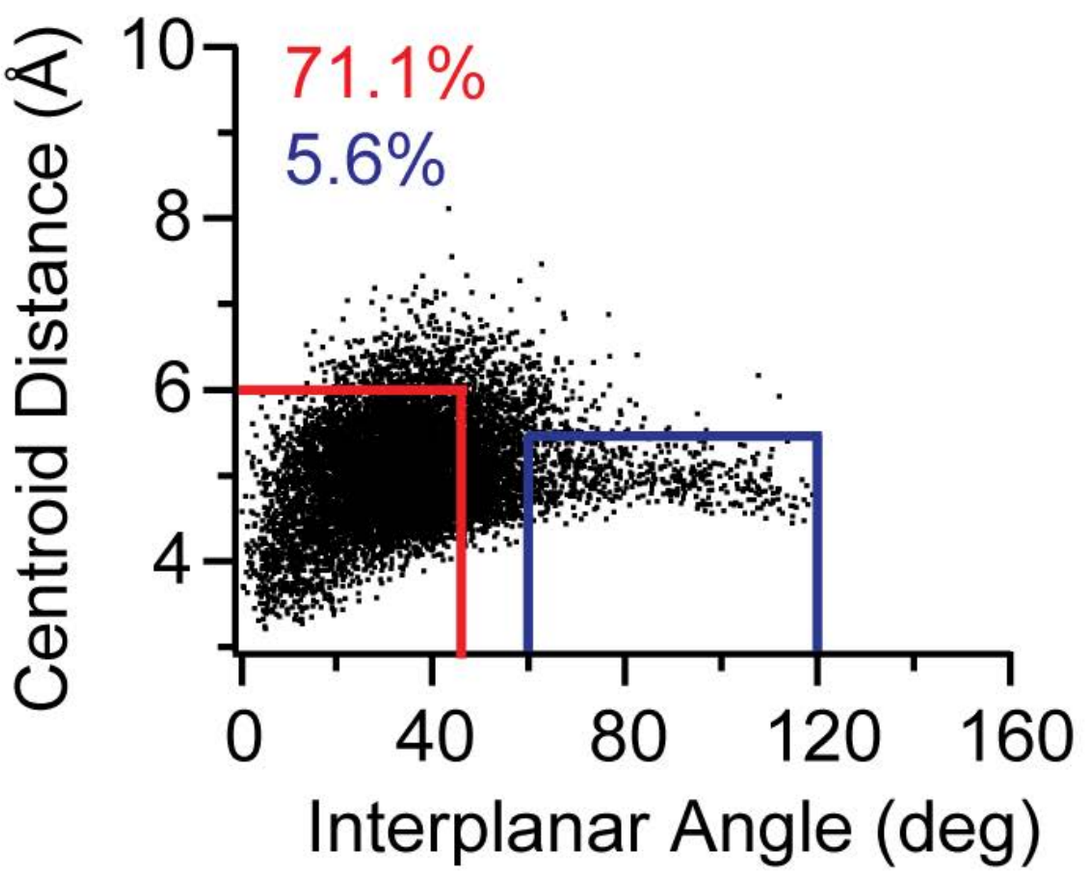

A
18a
RJW100
*
-1 0 1 2 3 4
ΔTm vs DLPC (°C)
B
Tm (°C)
50 49 48 47 46 45
DMSO
GSK8470
RR-RJW100
Endo-RJW100
WT
T352V
C
LRH-1 activity (RLU)
3 2 1 0
DMSO
18a
RJW100
*
pCI
LRH-1
D
RR-RJW100
LRH-1 activity (RLU)
3 2 1 0
pCI
WT
T352V
H390A
A349F
0 10 20 30
Concentration (μM)
E
GSK8470
LRH-1 activity (RLU)
3 2 1 0
pCI
WT
T352V
H390A
A349F
0 10 20 30
Concentration (μM)
F
Endo-RJW100
LRH-1 activity (RLU)
3 2 1 0
pCI
WT
T352V
H390A
A349F
0 10 20 30
Concentration (μM)
G
T352V
H390
T352V
H390
water lost
H
Centroid Distance (Å)
10 8 6 4
25.7%
T352V +
GSK8470
0 40 80 120 160
Interplanar Angle (deg)

Figure 9

A


H390
T352

B


71.1%
5.6%
Centroid Distance (Å)
10
8
6
4
0
40
80
120
160
Interplanar Angle (deg)