



# Comparison of activity, structure, and dynamics of SF-1 and LRH-1 complexed with small molecule modulators

Received for publication, December 16, 2022, and in revised form, May 29, 2023. Published, Papers in Press, June 14, 2023.  
<https://doi.org/10.1016/j.jbc.2023.104921>

Michael L. Cato<sup>1,†</sup>, Emma H. D'Agostino<sup>1,†</sup>, Racheal M. Spurlin<sup>2</sup>, Autumn R. Flynn<sup>2</sup>, Jeffery L. Cornelison<sup>2</sup>, Alyssa M. Johnson<sup>2</sup>, Rei A. Fujita<sup>1</sup>, Sarah M. Abraham<sup>1</sup>, Nathan T. Jui<sup>2</sup>, and Eric A. Ortlund<sup>1,\*</sup>

From the <sup>1</sup>Department of Biochemistry, Emory University School of Medicine, Atlanta, Georgia, USA; <sup>2</sup>Department of Chemistry, Emory University, Atlanta, Georgia, USA

Reviewed by members of the JBC Editorial Board. Edited by Brian Strahl

Steroidogenic factor-1 (SF-1) is a phospholipid-sensing nuclear receptor expressed in the adrenal glands, gonads, and hypothalamus which controls steroidogenesis and metabolism. There is significant therapeutic interest in SF-1 because of its oncogenic properties in adrenocortical cancer. Synthetic modulators are attractive for targeting SF-1 for clinical and laboratory purposes due to the poor pharmaceutical properties of its native phospholipid ligands. While small molecule agonists targeting SF-1 have been synthesized, no crystal structures have been reported of SF-1 in complexes with synthetic compounds. This has prevented the establishment of structure–activity relationships that would enable better characterization of ligand-mediated activation and improvement in current chemical scaffolds. Here, we compare the effects of small molecules in SF-1 and its close homolog, liver receptor homolog-1 (LRH-1), and identify several molecules that specifically activate LRH-1. We also report the first crystal structure of SF-1 in complex with a synthetic agonist that displays low nanomolar affinity and potency for SF-1. We use this structure to explore the mechanistic basis for small molecule agonism of SF-1, especially compared to LRH-1, and uncover unique signaling pathways that drive LRH-1 specificity. Molecular dynamics simulations reveal differences in protein dynamics at the pocket mouth as well as ligand-mediated allosteric communication from this region to the coactivator binding interface. Our studies, therefore, shed important insight into the allostery driving SF-1 activity and show potential for modulation of LRH-1 over SF-1.

Steroidogenic factor-1 (SF-1; NR5A1) is 1 of 48 human members of the nuclear receptor (NR) family, composed of ligand-regulated transcription factors that control the expression of genes involved in critical biological processes such as development, metabolism, and reproduction (1–3). SF-1 belongs to the NR5A subfamily, a class of monomeric, phospholipid-binding NRs (4, 5). SF-1 is critical for development of the adrenal glands, gonads, and ventromedial nucleus of the hypothalamus (VMH), where it also plays

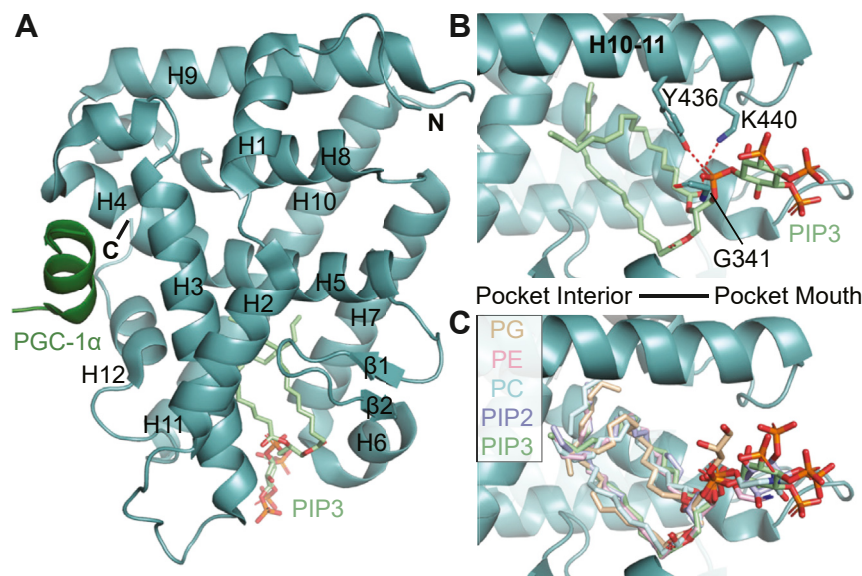
critical roles in steroidogenesis and energy homeostasis (6–12). In the adrenal glands and gonads, SF-1 controls the expression of steroidogenic genes as well as the expression of receptors for adrenocorticotrophic and follicle-stimulating hormones to stimulate steroidogenesis (13–16). In the VMH, SF-1 regulates energy homeostasis, although the precise genes controlled by SF-1 are unclear (17, 18). Loss of SF-1 in this tissue leads to obesity due to increased food intake, decreased energy expenditure, and impaired glucose tolerance and insulin sensitivity (19). SF-1 has also been implicated in adrenocortical carcinoma, a rare and deadly cancer (20–27). Therefore, SF-1 is an attractive therapeutic target for disease states characterized by dysregulation of steroidogenesis and energy homeostasis as well as adrenocortical carcinoma.

Liver receptor homolog-1 (LRH-1; NR5A2) is the only other human NR5A and is largely expressed in tissues of endodermal origin (5). It is well characterized in the liver and intestine, where it regulates processes such as lipid and glucose homeostasis, as well as steroidogenesis (5, 28). Accordingly, this receptor has been successfully targeted to alleviate symptoms associated with diabetes (29) and inflammatory bowel disease (30) *in vivo*.

The endogenous ligands for the NR5As are believed to be phospholipids (PLs) (Fig. 1). Early crystal structures of both SF-1 and LRH-1 revealed bacterial phospholipids in the binding pocket, with the polar head engaging hydrophilic residues at the mouth of the pocket and the fatty acyl tails protruding into the largely hydrophobic pocket interior (31–34). Subsequent studies identified phosphatidylcholine species which could activate both NR5As (35, 36), and dietary phospholipid 1,2-dilauroyl-sn-glycero-3-phosphocholine (DLPC) was able to activate LRH-1 *in vivo* (29). Both NR5As can also bind phosphatidylinositols with high affinity (Fig. 1) and numerous studies have found that these lipids can regulate SF-1 activity (34, 37–43). However, PLs do not make ideal therapeutic candidates as they are challenging to work with in the laboratory setting and are rapidly metabolized and remodeled *in vivo*, limiting their efficacy. There is significant interest in the development of synthetic NR5A ligands, and recent progress has produced high affinity modulators (44–47). Cross-reactivity with SF-1 has been reported for

<sup>†</sup> These authors contributed equally to this work.

\* For correspondence: Eric A. Ortlund, [eortlund@emory.edu](mailto:eortlund@emory.edu).



**Figure 1. Various phospholipid species bind SF-1 with a similar binding pose.** A, structure (PDB entry 4QJR) of SF-1 LBD (teal) bound to PIP3 (light green). The fragment of coactivator peroxisome proliferator-activated receptor gamma coactivator 1-alpha (PGC-1α) is shown in dark green. B, PIP3 binding orientation is shown, with residue side chains (and main chain of G341) critical for contacting the phospholipid polar head group shown as sticks. C, the binding orientation of different phospholipids is shown. Each phospholipid is oriented with the polar head at the mouth of the pocket and fatty acyl tails occupying the pocket cavity. Phospholipids included are phosphatidylglycerol ("PG," tan, PDB entry 1YOK), phosphatidylethanolamine ("PE," pink, PDB entry 1ZDT), phosphatidylcholine ("PC," cyan, PDB entry 3F7D), phosphatidylinositol 4,5-bisphosphate ("PIP2," purple, PDB entry 4QK4), and phosphatidylinositol 3,4,5-trisphosphate ("PIP3," green, PDB entry 4QJR). O = red, N = blue, P = orange, C = variable.

some of these small molecules, and there have been several promising antagonists developed (44, 45, 48–51). Nevertheless, no structure has been reported of SF-1 in complex with a synthetic modulator, preventing the establishment of a structure–activity relationship to further compound development.

We used a library of small molecule agonists designed to target different regions of the LRH-1 binding pocket to explore SF-1 agonism as it compares to that of LRH-1. We found that deep pocket contacts are particularly useful for LRH-1 specificity, as numerous small molecules with moieties exclusively targeting this region only activated LRH-1. However, a small molecule ("6N-10CA") that simultaneously contacts deep pocket and mouth residues demonstrated low nanomolar affinity and potency for SF-1. The stability induced by this molecule enabled us to solve the first crystal structure of SF-1 in a complex with a synthetic agonist. While the structure shows a co-occupying bacterial PL within the crystal, the electron density for 6N-10CA is unambiguous. Molecular dynamics (MD) simulations confirm that 6N-10CA makes critical contacts at the targeted polar sites but indicate that structural features at two loops situated at the pocket mouth destabilize ligand orientation in the SF-1 binding pocket. This may explain why small molecules consistently show greater affinity for LRH-1. Simulations also suggest that small molecule agonists drive greater allosteric communication from the pocket mouth to the site of coregulator binding in LRH-1. Our mechanistic studies thus reveal how SF-1 differs in its response to small molecules and offer insight into the development of selective SF-1 and LRH-1 synthetic ligands.

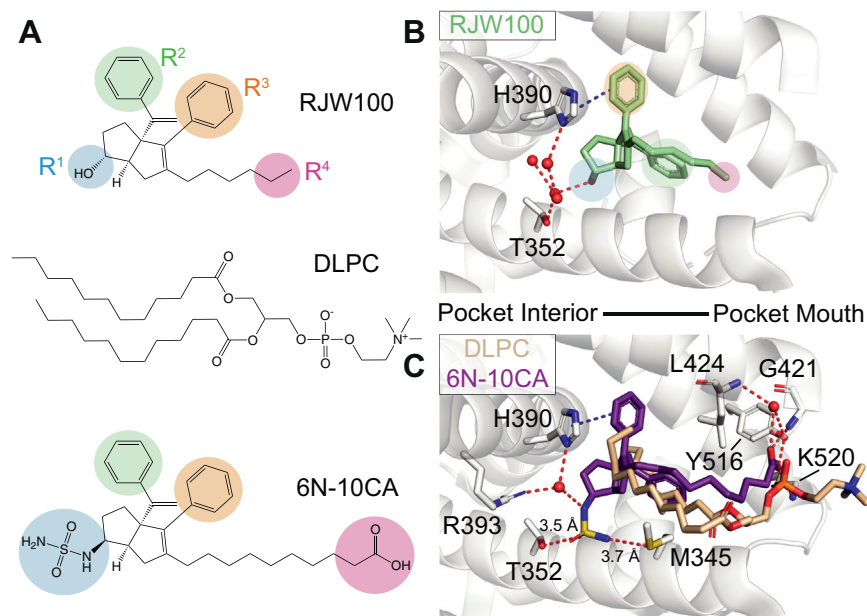
## Results

### NR5A synthetic agonist design

One of the earliest NR5A synthetic agonists was RJW100, which had a bicyclic hexahydropentalene (6HP) core and activated SF-1 and LRH-1 *in vitro* (Fig. 2A) (45). Structural studies revealed that the R<sup>1</sup> hydroxyl group indirectly engages deep pocket polar residues through a conserved water network (Fig. 2B) (52). With this structural information, we synthesized compounds using the RJW100/6HP scaffold that targeted polar residues deep within the pocket. Replacing the R<sup>1</sup> hydroxyl with a sulfamide radically improved affinity and potency by enabling direct and indirect contacts with deep pocket polar residues ("6N," Fig. 2) (38, 44). We next sought to increase activation by lengthening the alkyl chain (R<sup>4</sup> position; Fig. 2A) and appending polar groups to access PL-coordinating residues at the pocket mouth (Figs. 1B and 2C) (46). This strategy was successful, resulting in improved activation and potency toward LRH-1. The best compound from this series ("10CA") activates LRH-1 to decrease disease-associated weight loss in an *in vivo* model of ulcerative colitis (30, 46).

Recently, we combined these two design strategies into a hybrid compound ("6N-10CA") that simultaneously accesses deep pocket and pocket mouth interactions (Fig. 2, A and C). The modifications have a synergistic effect on affinity and potency, and 6N-10CA more effectively alters LRH-1 target gene expression in HepG2 cells than its single component agonists (47).

We have also developed a library of additional RJW100 derivatives and shown that some of these activate SF-1 in reporter assays (44). However, while the majority of RJW100 derivatives bind SF-1, they consistently do so with considerably



**Figure 2. Structure and binding pose of LRH-1 agonists.** A, chemical structures of LRH-1 small molecule agonist RJW100 (top), phospholipid agonist DLPC (middle), and small molecule agonist 6N-10CA (bottom). Major substituents of synthetic ligands are highlighted. B, binding orientation of RJW100 (light green, PDB entry 5L11) is shown, with major substituents indicated as in (A). C, binding orientations of 6N-10CA (purple, PDB entry 7TT8) and DLPC (tan, PDB entry 4DOS). Interacting residue side chains (and main chains of G421 and L424) are shown as sticks (O = red, N = blue, S = yellow, P = orange, C = white). Waters are represented as red spheres, hydrogen bonds are represented as red dotted lines, and  $\pi$ - $\pi$  stacking with H390 is represented as a blue dotted line. Only interactions with 6N-10CA are shown for simplicity.

lower affinity than LRH-1 (38). While our work has focused on LRH-1, the NR5A binding pockets are strikingly similar. We therefore sought to use our small molecule library to identify a useful tool for targeting SF-1, gain a better mechanistic understanding of its ligand-mediated activity, and explore what features of small molecule binding/agonism drive LRH-1 specificity.

#### Biochemical characterization of small molecules

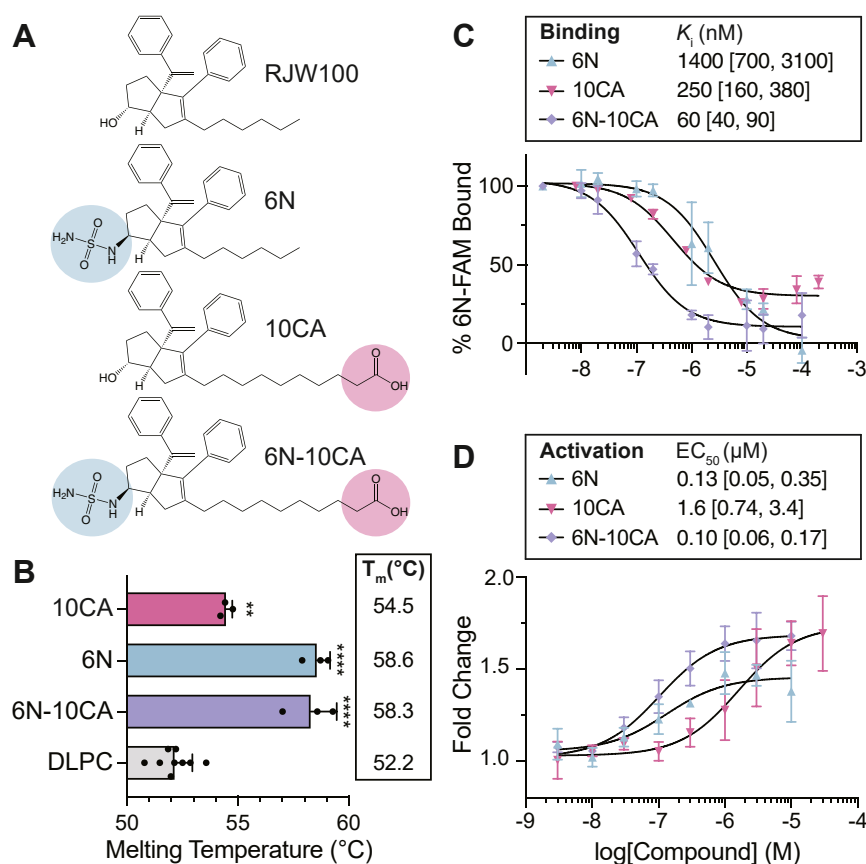
We began by comparing the functional effects of 6N-10CA with its constituent compounds, 6N and 10CA (Fig. 3A). Thermal shift assays revealed that all three compounds stabilize SF-1, relative to the DLPC phospholipid control (Fig. 3B). 6N and 6N-10CA both resulted in a positive shift in melting temperature of  $\sim 6.5$  °C, whereas 10CA shifted the melting temperature by  $\sim 2$  °C. FP competition assays (38) show that the 6N and 10CA modifications had a synergistic effect on SF-1 binding, resulting in a  $\sim 20$ -fold increase in affinity over 6N and a  $\sim 4$ -fold improvement over 10CA (Fig. 3C). The 60 nM  $K_i$  of 6N-10CA is the highest affinity reported for an SF-1 synthetic agonist, although it does not reach the picomolar affinity of 6N-10CA for LRH-1 (47). Luciferase reporter assays (Fig. 3D) revealed that 6N-10CA had a similar potency as 6N and was 16-fold more potent than 10CA, with an  $EC_{50}$  of 100 nM. 6N-10CA had a maximum activation of  $\sim 1.7$ , which was similar to 10CA and higher than that of 6N (max =  $\sim 1.4$ ). Taken together, these data demonstrate that, like LRH-1, incorporation of both sulfamide and carboxylic acid moieties improves small molecule binding and activation for SF-1.

#### Exploration of NR5A specificity

One of the greatest challenges in NR5A small molecule development has been the ability to specifically target one receptor. The sequences between the ligand binding domains (LBDs) of LRH-1 and SF-1 are highly conserved (LBD sequence identity =  $\sim 61\%$ ; Figure 4A), and the few non-identical residues in the pocket interior largely preserve hydrophobic character (Fig. 4B). Previous studies have noted that a few 6HP-derived small molecules are capable of specifically activating LRH-1, which suggests that specificity with this class of compounds is possible (44, 45). We used a library of small molecule agonists to identify moieties that specifically activate one NR5A over the other (Fig. 4C). Out of the six compounds tested, three activated LRH-1 without activating SF-1 (Fig. 4D). These included 2N, which was shown previously to specifically activate LRH-1 over SF-1 (44) as well as 4N and 7N. We also note that 3N reduced SF-1 activity, indicating potential utility for antagonist development. Previous studies have shown that LRH-1 is highly sensitive to modification and removal of the  $R^2$  substituent (44, 53). We note the same phenomenon here for SF-1, as even the highly efficacious 6N-10CA is rendered ineffective when the  $R^2$  group is removed (Fig. 4D). Therefore, deep pocket residues are an attractive region of the binding pocket for promoting specificity for LRH-1, as nearly half of small molecule variants tested drove activity of LRH-1 without activating SF-1.

Recent studies have suggested that the inclusion of the 10CA carboxylic acid tail improves small molecule potency and efficacy for LRH-1 (47). We were interested to see whether we could incorporate this moiety into small molecules that specifically activate LRH-1 without promoting cross-reactivity





**Figure 3. 6N-10CA displays nanomolar affinity and potency for SF-1.** A, chemical structures of the previously published small molecules RJW100, 6N, 10CA, and 6N-10CA. The 6N sulfamide (blue) and 10CA carboxylic acid (pink) moieties are highlighted. B, thermal melting assay comparing effects of small molecules on SF-1 thermal stability as compared to activating PL DLPC.  $T_m$  = melting temperature. Data are shown as means  $\pm$  SD from three (small molecule) or nine (DLPC) independent experiments. Data analyzed relative to DLPC control: One-Way ANOVA with Dunnett multiple comparison's test,  $**p < 0.01$ ,  $****p < 0.0001$ . C, FP competition assay comparing compound binding affinity.  $K_i$  = inhibition constant. 95% confidence intervals are shown in brackets. Data shown as means  $\pm$  SD from two (6N and 10CA) or three (6N-10CA) independent experiments. D, luciferase reporter assay showing effects of small molecules on SF-1 activity.  $EC_{50}$  = half maximal effective concentration. 95% confidence intervals are shown in brackets. Data were normalized relative to DMSO control and are shown as means  $\pm$  SD from three biological replicates.

with SF-1. We, therefore, synthesized 2N and 4N variations (termed “2N-10CA” and “4N-10CA”) and tested them in reporter assays. Consistent with previous studies, the inclusion of the carboxylic acid tail improved agonism, and specificity for LRH-1 was largely preserved for both small molecules (Fig. 4E). 2N-10CA and 4N-10CA thus represent valuable tools for specifically activating LRH-1 over SF-1 and indicate that unfavorable deep pocket contacts supersede agonistic effects of SF-1 pocket mouth interactions. The only major difference deep in the pocket is the absence of a methionine (M345) present in LRH-1, which is a leucine (L265 M) in SF-1 (Fig. 4B). This residue has been shown to be critical for certain ligand-mediated effects in LRH-1, and mutation of this residue decreases efficacy of 2N in LRH-1 reporter assays (44, 54). We therefore hypothesized that this residue may be driving the specificity of LRH-1-specific ligands. However, 2N-10CA failed to activate the L265M SF-1 mutant, indicating that this residue is not sufficient to drive ligand-mediated activation (Fig. S1).

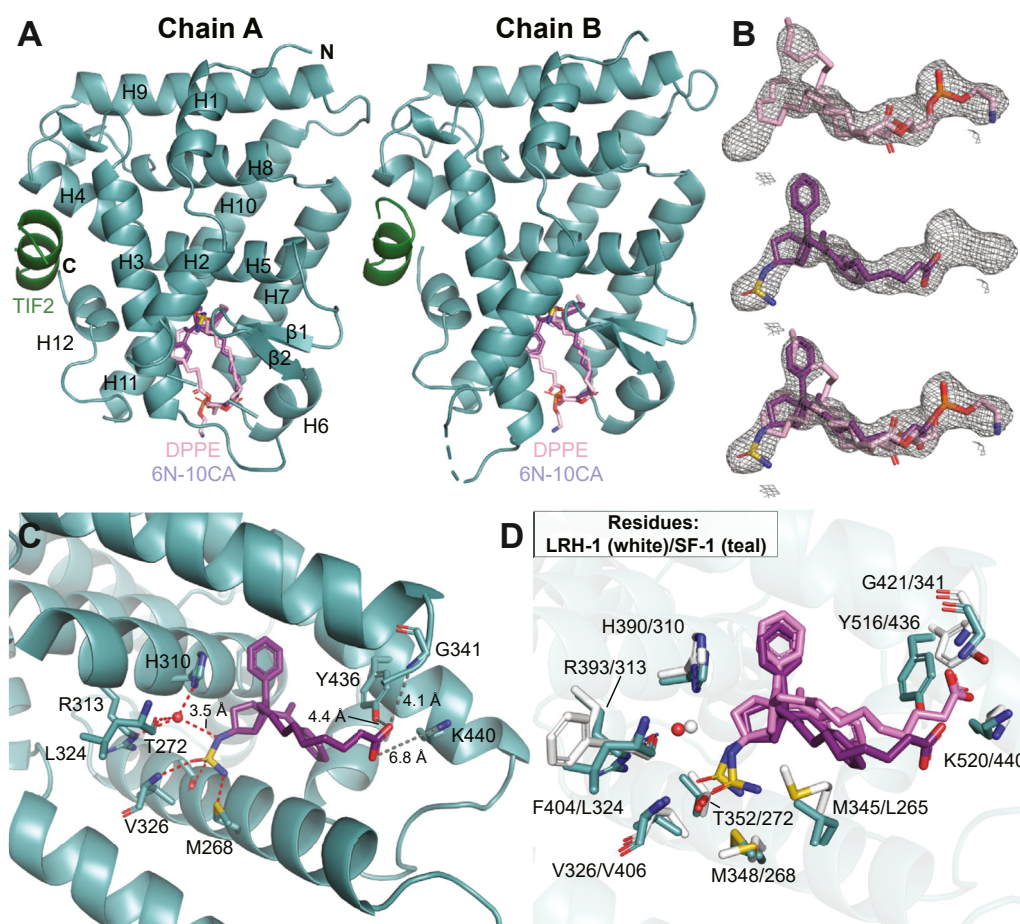
#### Crystal structure of SF-1-6N-10CA

To visualize the 6N-10CA binding mode and understand its mechanism of action, we solved the structure of the SF-1 LBD bound to 6N-10CA and a fragment of the coactivator

transcriptional intermediate factor 2 (TIF2) to a resolution of 2.59 Å (Figs. 5, S2, and Table S1). There is electron density indicating the presence of both 6N-10CA and a PL in the pocket of both monomers in the asymmetric unit, indicating partial occupancy of both ligands in the crystal (Figs. 5, A and B and S2A). Modeling of both ligands is supported by the strong electron density centered on the sulfamide group of 6N-10CA and phosphate moiety of the PL. We modeled in dipalmitoylphosphoethanolamine (DPPE; 16:0/16:0), which has been co-crystallized with SF-1 before (32, 33), to satisfy the PL density. Although we modeled 6N-10CA in both chains, the density for the 6HP core is more evident in chain B, and the density corresponding to the sulfamide in chains A and B becomes visible in omit maps at  $\sigma = \sim 4$  and  $\sigma = \sim 6$ , respectively. Nevertheless, the relative positioning of 6N-10CA in both chains is virtually identical after refinement (Fig. S2B).

The 6N-10CA interactions deep within the pocket are highly similar to those of 6N-10CA and LRH-1 (Fig. 5, C and D) (47). The sulfamide moiety directly interacts with the side chains of T352, M268, and backbone of V326. It also engages a water-mediated hydrogen bond network with H310, R313, and L324 (Fig. 5C). The mediatory water has been modeled in previous structures of LRH-1 bound to 6N (44) and 6N-10CA





**Figure 5.** SF-1 LBD is bound to synthetic agonist 6N-10CA and phospholipid DPPE. **A**, chains A and B of the SF-1-6N-10CA-TIF2 structure (PDB entry 8DAF). 6N-10CA and DPPE are shown co-occupying the ligand binding pocket, with the tail of DPPE overlapping with the 6N-10CA tail. Fragment of TIF2 coactivator protein is shown in green. **B**,  $F_o - F_c$  omit map after deletion of all ligands in chain B of the model (contoured at  $\sigma = 2.0$ ). **C**, binding orientation of 6N-10CA in chain B, with interacting residue side chains (and main chains of L324, V326, and G341) shown as sticks (O = red, N = blue, S = yellow, C = teal or purple). Waters are represented as red spheres and hydrogen bonds are represented as red dotted lines. Note that the 6N-10CA carboxylic acid tail was not within hydrogen bond distance of indicated residues. **D**, comparison of the structures of SF-1 (protein = teal, ligand = purple) and LRH-1 (PDB entry 7TT8; protein = white, ligand = pink) bound to 6N-10CA (chain B). The red sphere represents the mediatory water in the SF-1 structure, while the white sphere represents the mediatory water in the LRH-1 structure.

(47) and was visible in omit maps in both chain A and B at  $\sigma = \sim 4$  and  $\sigma = \sim 3$ , respectively. These are all interactions made by the sulfamide group with LRH-1. However, in LRH-1, 6N-10CA engages the backbone of phenylalanine (F404, LRH-1 numbering) rather than leucine (L324, SF-1 numbering) as well as the nearby methionine noted earlier (M345, LRH-1 numbering). The DPPE molecules in chains A and B have the same orientation as PLs complexed with SF-1 in previous structures, with the phosphates coordinated by Y436, K440, and the backbone of G341 (Fig. S2C). The strong density for DPPE interferes with proper refinement of the carboxylic acid tail position, pulling it away from the residues that coordinate this moiety in LRH-1 (Fig. 5, C and D). These residues correspond to phosphate-coordinating residues in SF-1 (Fig. 1B): G341, Y436, K440.

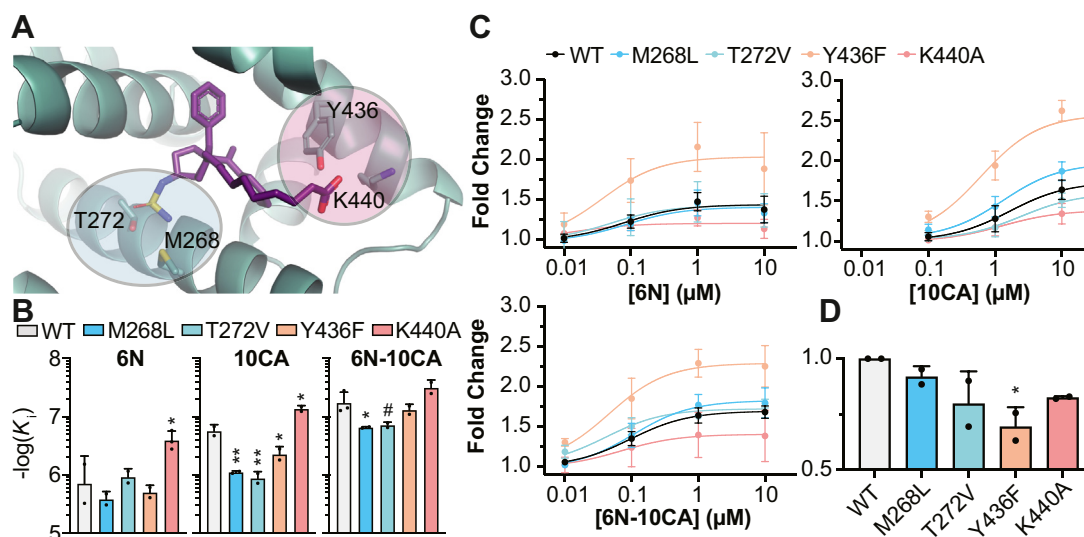
#### Mutational analysis of residue contacts

We mutated residues that made direct contact with 6N-10CA deep within the pocket (M268 and T272) and at the

mouth (Y436 and K440) to explore their roles in ligand binding and activation (Fig. 6A). Deep pocket mutations decreased SF-1 binding affinity for both 10CA and 6N-10CA (Fig. 6B), suggesting that ligands are sensitive to disruption of water-mediated and direct contacts deep in the pocket. Mutation of pocket mouth residue Y436 had minimal impact on 6N and 6N-10CA binding but decreased 10CA binding affinity, indicating that the sulfamide anchoring group can counteract reduced mouth pocket contacts. The K440A mutation consistently increased affinity of each agonist, although the increased affinity for 6N-10CA was statistically insignificant. This potentially highlights that this lysine is dispensable for ligand binding.

Examination of these mutations in reporter assays revealed that deep pocket mutations had little impact on compound potency or efficacy (Fig. 6C). However, the Y436F pocket mouth mutation consistently increased ligand-mediated activation. This mutant demonstrated diminished basal activity (Fig. 6D), presumably due to decreased capacity for PL ligand binding, which may have driven the greater fold changes





**Figure 6. Deep pocket and pocket mouth residues are responsible for small molecule binding and agonism.** A, residues mutated for FP competition and luciferase reporter analyses in (B–D) are indicated. B, FP competition assay was used to assess the effect of indicated mutations on ligand binding.  $K_i$  = inhibition constant. Data shown as means + SD. Data compared to wild-type (WT) control: One-Way ANOVA with Dunnett multiple comparison's test, # $p < 0.1$ , \* $p < 0.05$ , \*\* $p < 0.01$ . C, luciferase reporter assays were used to assess the effect of indicated mutations on ligand-mediated activation of SF-1. Data were normalized relative to DMSO control and are shown as means ± SD from three biological replicates. D, luciferase reporter assay showing effects of mutations on basal SF-1 activity. Data normalized and analyzed relative to WT SF-1 activity and shown as means + SD from two biological replicates: One-Way ANOVA with Dunnett multiple comparison's test, \* $p < 0.05$ .

observed. In contrast, the K440A mutation, while consistently improving compound binding, diminished efficacy of each small molecule, reflecting an important role that this residue plays for communicating ligand binding status. This indicates that the two critical residues at the SF-1 pocket mouth have distinct and important roles for compound binding and efficacy and that these roles are not limited to small molecules that mimic phospholipid contacts.

#### MD analysis of ligand contacts

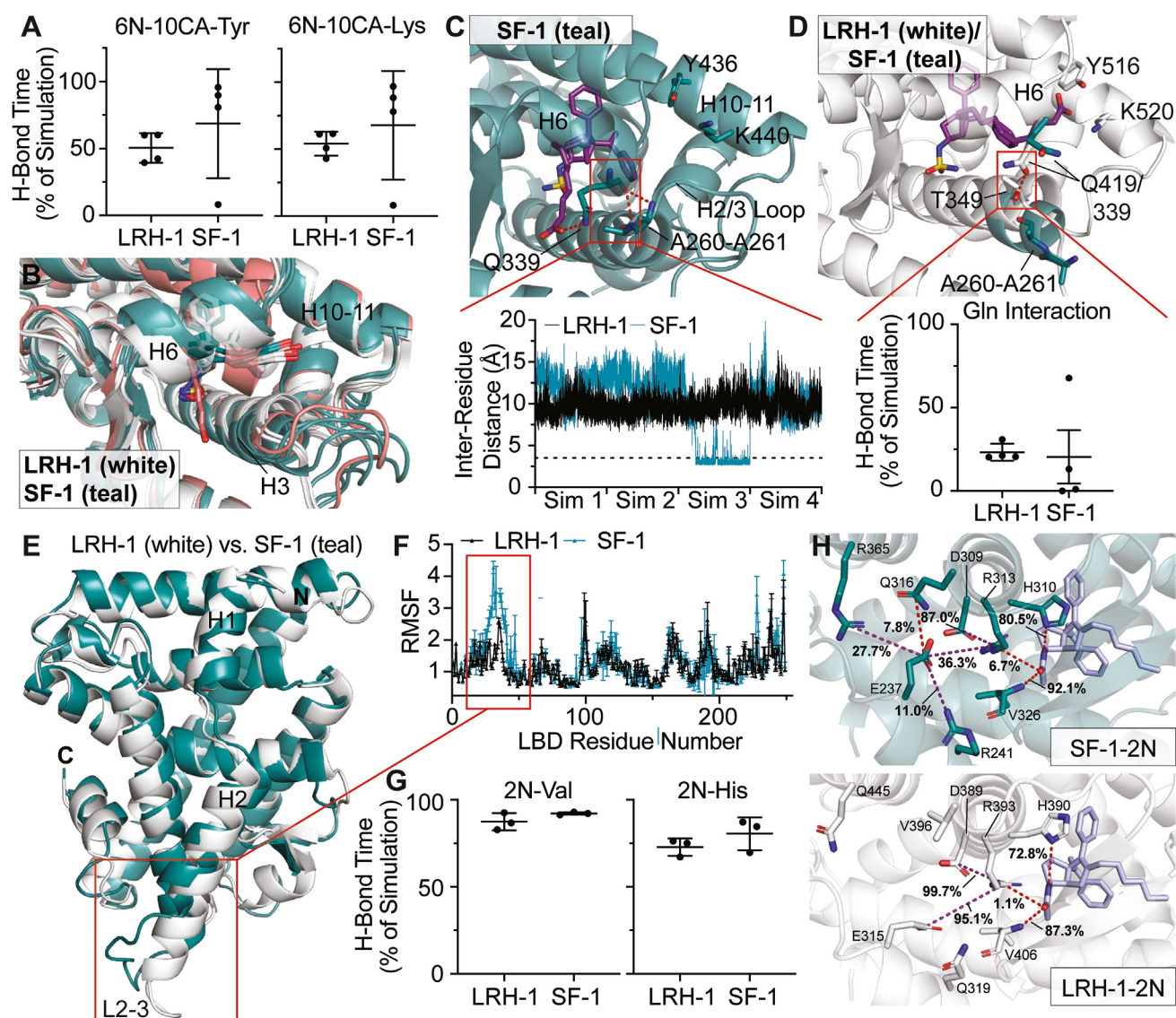
We used MD simulations (4 × 500-ns) to compare ligand dynamics in SF-1 and LRH-1. While the starting positioning of the 6N-10CA tail was considerably different between LRH-1 and SF-1 complexes (Fig. 5D), the carboxylic acid quickly made and maintained hydrogen bond contacts with tyrosine and lysine residues at the mouth of the pocket for three of the four simulations (Fig. 7A). We noticed that the tail moved radically out of position early in one of the simulations (Fig. 7, A and B), and closer examination revealed that an inter-helical hydrogen bond was responsible for trapping the carboxylic acid in an orientation away from coordinating atoms at the pocket mouth (Fig. 7C). Dynamics of LRH-1 in this region is restricted by an inter-helical hydrogen bond driven by a threonine absent in SF-1 (Fig. 7D) and a longer helix 2 (Fig. 7E) observed in previously solved structures of the LBD, which results in a smaller root mean square fluctuation (RMSF) of corresponding LRH-1 residues (Fig. 7F) (32, 34, 35, 37, 41, 47, 52, 55). Therefore, as seen in LRH-1, 6N-10CA is capable of strong engagement with SF-1 residues at the mouth of the pocket, but can be repositioned away from these residues due to motions at the pocket mouth that are unique to SF-1.

Simulations (3 × 500-ns) with LRH-1-specific 2N revealed that this small molecule directly engages the backbone of a

valine (LRH-1 V326; SF-1 V406) for nearly the entire simulation in both LRH-1 and SF-1 (Fig. 7G). Initial simulations indicated that 2N had substantially reduced hydrogen bonding with the deep pocket histidine (LRH-1 H390; SF-1 H310) in SF-1 complexes (Fig. S3). However, we found that this was driven by different starting rotamers of a nearby arginine (LRH-1 R393; SF-1 R313) (Fig. S3). Simulations with similar starting positions for this arginine indicated that 2N engaged in hydrogen bonds with both the nearby valine and histidine for comparable durations as LRH-1 simulations (Fig. 7G). Nevertheless, we noticed that this arginine is stabilized by an additional salt bridge with a nearby glutamate in LRH-1 for considerably longer than in SF-1 (Fig. 7H). This is likely driven by the presence of several nearby residues unique to SF-1 that compete for this side chain (Fig. 7H). It is possible that at longer time scales, this results in greater propensity for repositioning of R313 in SF-1, consequently destabilizing local contacts with the nearby histidine. This may explain why 2N shows specificity for LRH-1 in binding (38) and activity (44).

#### MD analysis of ligand-driven allostery

Activating ligands drive a conformation on the activation function surface (AF-2) on the LBD that promotes coactivator association. Coactivators drive target gene expression by recruiting transcriptional machinery and chromatin remodeling factors to gene loci (4). Previous simulations, paired with hydrogen-deuterium exchange mass spectrometry, have highlighted the importance of a region corresponding to β1/2, H6, and H7 on LRH-1 as an alternative activation function surface (AF-B) that communicates ligand-binding status to the AF-2 (35, 55). We sought to similarly use simulations (3 × 500-ns) to explore whether this region also plays a critical role in SF-1 activity. A comparison of average structures revealed that

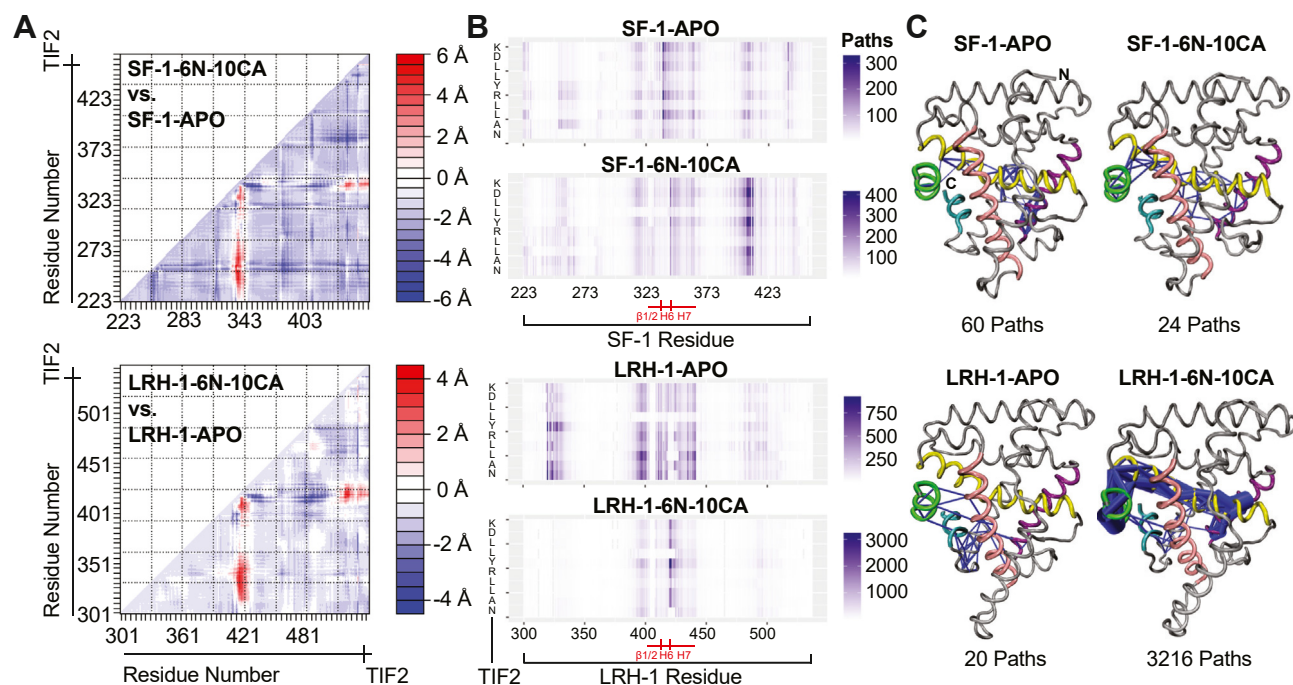


**Figure 7. MD simulations reveal differences in pocket mouth dynamics and small molecule binding.** A, percentage of time that 6N-10CA spent in hydrogen bond with mouth residues Y516 (LRH-1)/Y436 (SF-1) and K520 (LRH-1)/K440 (SF-1) during MD simulations. Data shown as means  $\pm$  SD from four 500-ns simulations. B, average atomic positions of 6N-10CA bound to LRH-1 (white) and SF-1 (teal). All four simulations are shown independently, with the repositioned 6N-10CA (and corresponding SF-1 LBD) shown in pink. C, top: Representative image of hydrogen bond contact between SF-1 L2-3 and helix 6. Bottom: Inter-residue distance between L2-3 alanine and helix 6 glutamine for LRH-1 (black) and SF-1 (blue) complexes. The dotted line corresponds to 3.5 Å. D, top: Representative image of inter-helical hydrogen bond in starting complex of LRH-1 (white). Corresponding residues on SF-1 are shown in blue. Bottom: Percentage of time that 6N-10CA spent in hydrogen bond with mouth residues Y516 (LRH-1)/Y436 (SF-1) and K520 (LRH-1)/K440 (SF-1) during MD simulations. The outlier seen in SF-1 corresponds to the simulation where the loop was repositioned (see panels A–C). E, LBD of LRH-1 (white) and SF-1 (teal) prior to MD simulations. L2-3 is highlighted with a red box. F, RMSF analysis of LRH-1-6N-10CA and SF-1-6N-10CA. Data represented as means  $\pm$  SD from four 500-ns simulations. G, percentage of time that 2N spent in hydrogen bond with H390 (LRH-1)/H310 (SF-1) and V406 (LRH-1)/V326 (SF-1) during MD simulations. Data shown as means  $\pm$  SD from three 500-ns simulations. H, representative images of critical deep pocket interactions from SF-1-2N (top) and LRH-1-2N (bottom) at the start of MD simulations. Important hydrogen bonds and electrostatic interactions are represented by red and purple dotted lines, respectively. Interaction duration is represented as the percentage of frames over three simulations.

the AF-B is mobilized by 6N-10CA in both LRH-1 and SF-1 complexes (Fig. 8A). To explore allosteric paths of communication driven by ligands, we conducted suboptimal paths analysis. This analysis identifies correlated motion within chains of residues, which reflects allosteric communication between two distant sites on the protein (56–58). This has previously detected allosteric patterns that are correlated with activity of different ligand-bound states of LRH-1 (44, 47, 55, 59). We examined paths between all residues on the LBD and all residues on the co-complexed peptide fragment of

coactivator TIF2 (Fig. 8B). Relative to the apo state, 6N-10CA dramatically increased paths of communication in LRH-1 between the region corresponding to the loop between helix 6 and 7 (“L6-7”) and TIF2. However, the SF-1 L6-7 was comparatively insensitive to 6N-10CA. Overlaying the paths between the glycine situated on this loop (G421, LRH-1 numbering; G341, SF-1 numbering) and TIF2 revealed that the effects of activating ligands are considerably more subtle in SF-1 (Fig. 8C). The number of paths and directionality of those paths is much more divergent between apo and ligand-bound





**Figure 8. 6N-10CA drives stronger AF-B allosteric communication in LRH-1.** A, all-residue (C $\alpha$ ) difference distance analysis of structures averaged after three 500-ns simulations. Inter-residue distances (Å) were determined for complexes and difference distance matrices were then calculated by subtracting distances from two complexes. The resulting matrices reveal structural differences between indicated complexes. B, matrices showing a number of paths between all residues on LRH-1-LBD and SF-1-LBD and their corresponding TIF2 peptides. Residues corresponding to the TIF2 peptide are on the vertical axis. Residues corresponding to SF-1 or LRH-1 are on the horizontal axis. Indicated in red are the regions that roughly correspond to the AF-B (B1/2-H6-H7). C, pathway analysis of LRH-1 or SF-1 LBD bound to ligands is shown (H3 = pink, H4/H5 = yellow, H7 = purple, H12 = cyan, TIF2 coactivator peptide = green). Paths between AF-B and TIF2 peptide are represented by blue lines. The number of paths between two nodes (C $\alpha$ ) is represented by edge weight. Note the SF-1-6N-10CA simulations were run with reoriented R313 (described in Fig. S3).

states in LRH-1. This result suggests that while pocket mouth-contacting ligands such as 6N-10CA mobilize the AF-B region of both receptors, this region is not as critical for communicating ligand binding status for SF-1.

## Discussion

Here, we explored SF-1 activity using several small molecules that have been designed to enhance the activity of its close homolog LRH-1 (30, 44, 46, 47). Lead agonists 6N, 10CA, and 6N-10CA were able to bind and activate SF-1, with 6N-10CA showing the greatest affinity. As evidenced by this cross-reactivity with SF-1 and the nearly identical ligand binding pockets, one of the most challenging tasks for NR5A small molecule agonist development has been specifically targeting one NR5A. Here, we show numerous compounds that specifically activate LRH-1. Appending a 10CA tail to these molecules increases LRH-1 activity without promoting SF-1 cross-reactivity, providing useful tools for targeting LRH-1 where activation of SF-1 is undesirable.

It has been challenging to improve SF-1 modulators without the mechanistic insight provided by an SF-1-small molecule structure. However, the high affinity and stabilization provided by 6N-10CA enabled the solution of the first SF-1 crystal structure with any non-PL ligand. This structure showed co-occupancy with a bacterial phospholipid. Despite the incomplete ligand exchange, the SF-1-6N-10CA structure yielded valuable information about the molecular interactions between

this agonist class and SF-1. The interactions in the pocket were nearly identical to those of 6N and 6N-10CA with LRH-1, with the sulfamide moiety forming an extensive hydrogen bond network with M268, H310, R313, L324, and V326 (44, 47). Only absent was the interaction with a nearby methionine (M345, LRH-1 numbering), as this residue is leucine in SF-1. While deep pocket interactions were nearly identical to those of LRH-1, the electron density of the bacterial PL hindered proper refinement of the 10CA tail.

We then explored how mutating residues directly contacted by 6N-10CA affected agonist binding and efficacy. We found that while deep pocket mutations slightly decrease the binding of carboxylic acid-containing small molecules, these mutations have minimal effects on activation. In contrast, pocket mouth mutations consistently either decreased (Y436F) or increased (K440A) binding affinity of small molecules. These mutations had the opposite effect on activity, with Y436F and K440A increasing and decreasing small molecule efficacy, respectively. The effect of Y436F may be the result of decreased basal activity in cells, which is likely driven by diminished hydrogen bond contact with activating PLs. The effect of K440A is more puzzling, as the mutation improved binding but decreased efficacy in reporter assays. This may be the consequence of the alanine mutation destabilizing allosteric motions at the pocket mouth while permitting increased access to the pocket interior. Altogether, these data suggest that while Y436 is critical for small

## Structure/activity of SF-1- and LRH-1-modulator complexes

molecule binding, K440 is important for communicating ligand binding status.

While the crystal structure confirmed a ligand orientation similar to that of LRH-1, it did not explain why 6HP-derived ligands consistently show greater affinity for LRH-1. This is the case even with 6N-10CA, which has an affinity for SF-1 over two orders of magnitude lower than for LRH-1 (Fig. 3C and reference (47)). We used MD simulations to elucidate how these small molecules show preference for LRH-1. While 6N-10CA makes stable contact with Y436 and K440, 25% of simulations show radical repositioning of the tail. LRH-1 has a long H2 and an H3-H6 interhelical bond which restricts local motion, while SF-1 is considerably more mobile due decreased helical character in this region. We note that this flexibility can result in closure of the SF-1 pocket mouth, trapping the hydrocarbon tail in an orientation away from the mouth residues. Interestingly, the same glutamine on H6 is responsible for the different inter-helical interactions visible in LRH-1 and SF-1 (Fig. 7, C and D). Previous studies have established that this loop ("L2-3") on SF-1 is responsive to ligand-binding status. The headgroups of PIP2 and PIP3 promote stabilization of SF-1 L2-3 (37) and recent work by Bryant *et al.* showed that the acyl chain composition of PLs also influences order of L2-3 in crystal structures, which correlates with coactivator binding (41). This site can therefore sense favorable contacts at the pocket mouth and pocket interior. As was noted in this study (41), there are several mutations on L2-3 that have clinical significance, indicating that this site is critical for normative SF-1 activity (60–63). Therefore, L2-3 appears to be important for PL-sensing, and this region is stabilized by the large head group of PIP3. This, along with the propensity for our ligands to reposition in this region, suggests that synthesizing 10CA variants with larger, PIP3-like moieties is an attractive route for improved SF-1 binding and agonism. We predict that larger moieties would prevent repositioning that is permitted by the small carboxylic acid tail in 6N-10CA and hydrocarbon tail on other ligands tested here.

While we identified several LRH-1-specific small molecules, the molecular basis for their specificity is not clear. The pocket interiors of LRH-1 and SF-1 are nearly identical, and the introduction of the nearby methionine (M345 in LRH-1) into the SF-1 pocket does not sensitize the receptor to 2N-10CA. While differences in pocket mouth motions may decrease the binding of ligands with a 10CA tail, these motions likely have a smaller effect on ligands such as 2N, which primarily contact deep pocket residues. A deeper investigation of MD simulations revealed that while residues proximal to small molecules are highly similar, those that lie directly outside of the binding pocket are not. We found that the positioning of the nearby arginine (R313 in SF-1) dramatically affects deep pocket interactions. This arginine is stabilized by a glutamate in LRH-1 (E315) for nearly the entire simulation. The corresponding glutamate in SF-1 (E237) interacts with R313 for a considerably shorter duration because of its propensity to engage with other interacting residues (*e.g.*, R241, Q316, and R365) not found in LRH-1. It is possible that at longer timescales, this results in destabilized deep pocket interactions. Therefore, we

hypothesize that the molecular bases for LRH-1 specificity are different pocket-mouth dynamics and destabilized deep pocket interactions.

While we did not identify small molecules that showed specificity for SF-1, our findings suggest ideas for future designs. The rigid pocket mouth of LRH-1 may be incompatible with small molecules with inflexible tails and/or certain large polar moieties. Additionally, the small molecules tested here are racemic mixtures. Our recent investigations found critical differences in binding between enantiomers of parent compound RJW100 (54). Interestingly, while LRH-1 showed a clear preference for one enantiomer in reporter assays, SF-1 showed similar sensitivity to both enantiomers. Therefore, chiral separation may result in small molecules that are specific for SF-1.

MD simulations also revealed that 6N-10CA promotes mobility of the region corresponding to  $\beta$ 1/2-H6-H7 in both SF-1 and LRH-1. This site is proximal to the binding pocket and includes a glycine on L6-7 that makes direct contact with the carboxylic acid tail on synthetic ligands, as well as the polar head group of PLs (Figs. 1B and 2C). This region ("AF-B") was originally identified as a site on LRH-1 critical for communication driven by activating PL DLPC (35). More recent studies have identified this site's role in differentiating coactivator and ligand identity (30, 44, 47, 54, 59). Pathway analysis revealed strong communication between this region and the TIF2 coactivator in LRH-1. However, 6N-10CA drove considerably weaker communication between L6-7 and TIF2 in SF-1, suggesting that this region may play less of a role for SF-1 activity. More direct methods, such as hydrogen-deuterium exchange mass spectrometry, are necessary to further explore ligand-driven allosteric communication in SF-1. Our simulations thus reveal different roles for both loops at the pocket mouth in ligand differentiation by NR5As.

With a highly effective agonist and the first synthetic small molecule-bound crystal structure of SF-1 in hand, new avenues of therapeutic development are now feasible. 6N-10CA would be a useful tool for elucidating SF-1's role in altering gene expression in the VMH. These studies may also enable improvements in agonist efficacy. 6N-10CA has a potency nearly two times less and an affinity two orders of magnitude lower than LRH-1, indicating differences in the mechanism of activation and room for improvement toward SF-1 (47). The ability to directly compare SF-1- and LRH-1-6N-10CA structure and dynamics also offers the potential to develop compounds that specifically activate LRH-1, as noted above. While global activation of LRH-1 is unlikely to be detrimental, unnecessary activation of SF-1 in the adrenal glands may lead to the overproduction of adrenal steroids. Thus, our studies provide useful information and tools for LRH-1 agonism in the settings of diabetes, obesity, and inflammatory bowel disease. Finally, SF-1 antagonists are sought for the treatment of adrenocortical cancer, and this structure may yield strategies to modify 6N-10CA for receptor destabilization and antagonism. We note that both receptors are sensitive to the removal of the R<sup>2</sup> group, which is positioned proximal to the AF-2. Modification of this moiety may drive NR5A inverse agonism, and this is currently being pursued in

our lab. Altogether, our studies offer novel insight into the mechanisms underlying ligand-mediated activation of SF-1 and provide useful tools for specifically targeting its close homolog LRH-1.

## Experimental procedures

### Chemical synthesis

Synthesis of 6N-10CA has been previously reported (38, 47). Synthetic methods for 1-7N have also been previously published along with RJW100 (44). The carboxylic acid derivate (10CA) was synthesized as previously described (46). Synthesis of 6N-10CA (No R<sup>2</sup>), 2N-10CA, and 4N-10CA are detailed in the supplemental information. Note that while all compounds tested in these studies are racemic mixtures, the enantiomers used for figures and MD simulations are based on evidence from previous crystal structures (44, 47).

### Cell culture

HeLa cells were cultured under standard conditions (5% CO<sub>2</sub>, 37°C) in phenol red-free MEM $\alpha$  + 10% fetal bovine serum (FBS) – charcoal/dextran treated and were verified to be *mycoplasma* free with the LookOut *Mycoplasma* PCR Detection Kit.

### Protein purification of wild-type SF-1

*E. coli* strain BL21(DE3)-pLysS was transformed with the SF-1 LBD (amino acids 218–461) in the pLIC-His vector and cultured at 37 °C to OD<sub>600</sub> of 0.6 in Lysogeny Broth medium in the presence of chloramphenicol and ampicillin. Protein expression was induced with 0.5 mM isopropyl-1-thio-D-galactopyranoside for 4 h at 32 °C. Cell pellets were lysed in 125 ml NiA (500 mM NaCl, 25 mM imidazole, 5% glycerol, 20 mM Tris HCl pH 7.4, 0.5 mM TCEP) with lysozyme, phenylmethylsulfonyl fluoride, and DNase followed by sonication. Lysate was clarified by centrifugation in a Sorvall RC 6+ centrifuge at 16,000g for 45 min. Supernatant was flowed over a 5 ml HisTrap FF column (GE Healthcare) and protein was eluted with NiB (NiA with 500 mM imidazole). To homogenize the lipid population, SF-1 was incubated overnight with DLPC. Size exclusion chromatography (SEC) into assay buffer (150 mM NaCl, 20 mM Tris HCl pH 7.4, 5% glycerol) was used as a final purification step. Protein was concentrated to ~3 mg/ml, flash frozen, and stored at –80 °C for use in assays.

### Protein purification of CysLite SF-1

CysLite SF-1 (amino acids 218–461, C247S, C412S) in the pLic-His vector was used for crystallization. This protein was purified as described for wildtype SF-1 through the HisTrap column; after elution from the HisTrap column, the 6X-His tag was cleaved overnight using tobacco etch virus protease. Cleaved protein was flowed over a second HisTrap column and the flowthrough was collected, concentrated to ~3 mg/ml, flash frozen, and stored at –80 °C for use in crystallization.

### Differential scanning fluorimetry

Purified SF-1 LBD, pre-exchanged with DLPC (0.2 mg/ml), was combined with small molecules overnight at 4 °C in assay buffer. SYPRO orange dye was added to the complexes the next day, at a final dilution of 1:1000. Complexes were heated at a rate of 0.5 °C/minute on a StepOne Plus thermocycler, using the ROX filter for fluorescence detection. The melting temperature (T<sub>m</sub>, 50% unfolding) was calculated using the Boltzman equation GraphPad Prism (version 9). Assays were conducted with nine technical replicates over three experiments.

### Fluorescence polarization

FP assays were performed as described previously (38, 47, 53). Briefly, experiments were conducted in black, polystyrene, non-binding surface 384-well plates (Corning Inc., Corning, NY) with 30  $\mu$ l volumes in assay buffer (150 mM NaCl, 20 mM Tris-HCl, 5% glycerol, pH 7.4). Binding affinity for 6N-FAM was determined using 10 nM 6N-FAM and protein concentrations ranging from 1<sup>–10</sup> – 5<sup>–5</sup> M. Plates were incubated overnight at 4 °C and centrifuged at 2000g for 2 min before polarization measurement. Polarization was monitored on a Neo plate reader (Biotek, Winooski, VT) at an excitation/emission wavelength of 485/528 nm. Plates were incubated overnight at 4 °C and centrifuged at 2000g for 2 min before polarization measurement. Nine technical replicates were conducted over three experiments and compiled binding data were baseline-corrected to wells with no protein and fit with a one-site binding curve in GraphPad Prism version 7 (GraphPad, Inc, La Jolla, CA). 6N conjugated to fluorescein amidite (FAM) (10 nM/well – 0.8 times the affinity of SF-1 for 6N-FAM) was incubated with SF-1 LBD (25 nM/well – 60% of the forward binding B<sub>max</sub>). Unlabeled compounds were added at concentrations indicated in figures with DMSO in each well held constant at 6.7% v/v. Each experiment was performed two or three times with four technical replicates each. Technical replicates were averaged and normalized independently prior to final data analysis. Using GraphPad Prism (version 9), data were fit to a one-site, fit K<sub>i</sub> curve, assuming a final probe concentration of 10 nM and probe affinity determined with forward binding assays (WT: 12 nM, M268L: 19.3 nM, T272V: 10.9 nM, Y436F: 9.9 nM, K440A: 5.5 nM).

### Luciferase reporter

Reporter assays were conducted as described previously (47, 52, 53). Briefly, HeLa cells were seeded at ~7500 to 10,000 cells per well in 96-well plates (white-walled, clear bottom) in MEM $\alpha$  + 10% FBS – charcoal/dextran treated. Once cells reached 80 to 90% confluence, they were transfected with LRH-1 (in pCI vector, 5 ng/well) or SF-1 (in pcDNA vector, 5 ng/well), a reporter plasmid with an NR5A response element derived from the *SHP* promoter cloned upstream of *Firefly* Luciferase (in pGL3-Basic vector, 50 ng/well), and a plasmid expressing *Renilla* Luciferase constitutively from a *CMV* promoter (1 ng/well). Cells were



## Structure/activity of SF-1- and LRH-1-modulator complexes

transfected with FuGENE at a ratio of 3:1 (FuGENE:DNA). Twenty-four hours after transfection, compounds were diluted in Opti-MEM and introduced to cells at the final concentrations indicated in figures (final DMSO concentration was 0.37%). Luciferase signal was measured after ~24 h using the DualGlo kit (Promega) with a BioTek Neo plate reader. Each experiment was conducted with three biological replicates, each with three technical replicates averaged prior to data analysis. Firefly Luciferase signal for each well was divided by the well's Renilla Luciferase signal intensity and then normalized relative to the DMSO control. Data were analyzed with GraphPad Prism (version 9) using a stimulating dose-response curve (Hill slope = 1). Data were excluded from analysis for cells treated with 3e-5 M of 6N and 6N-10CA, as the final signal showed a drastic decrease in overall signal, potentially indicating cell toxicity or compound insolubility.

### Sequence alignment

Sequences were aligned with Clustal Omega (64–66).

### Crystallization

Cleaved SF-1 CysLite was incubated overnight with 4-fold molar excess compound 6N, and SEC was used to remove bacterial lipids and excess compound and exchange into crystallization buffer (150 mM NaCl, 100 mM ammonium acetate (pH 7.4), 1 mM DTT, 1 mM EDTA, 2 mM 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonic acid (CHAPS)). An additional 2-fold molar excess of 6N was added to SF-1 collected after SEC. SF-1 was concentrated to ~5 mg/ml, 4-fold molar excess of the TIF2 peptide (<sup>3</sup>H<sub>3</sub>N-KENALLRYLLDKDD-CO<sub>2</sub><sup>-</sup>) was added, and the complex was incubated at room temperature for 2 h. Crystals were seeded with LRH-1-RJW100 crystals, grown as previously described (52). Crystals were grown using hanging drop vapor diffusion at 4 °C with 0.05 mM Na acetate (pH 4.6), 0 to 25% glycerol, and 5 to 11% PEG 4000, with 2 to 4 µl drops. The crystals contained bacterial phospholipid rather than 6N. Thus, soaking of 6N-10CA, which had a considerably higher affinity than 6N, was used to exchange the phospholipids. 6N-10CA (100 mM DMSO) was diluted to 2.5 mM in mother liquor, and 0.5 µl volumes were added to drops containing crystals for a 2-days soak. Crystals were flash-frozen in liquid nitrogen using mother liquor with a cryoprotectant of 30% glycerol.

### Structure determination

Data were collected remotely from the Southeast Regional Collaborative Access Team (SER-CAT) at the Advanced Photon Source (Argonne National Laboratories) using the 22ID beamline. Data were processed and scaled using HKL2000 (67) and phased by molecular replacement using Phaser-MR in Phenix (68) with PDB entry 1ZDT as the search model. Coot (69) and Phenix.refine (68, 70) were used for model building and refinement, respectively. Additional refinement was performed with PDB Redo (71). Figures were

constructed using Pymol (72). Hydrogen bonds in all figures were shown if indicated atoms were within ~3.5 Å of one another unless otherwise indicated. Omit maps were generated by removing all ligand atoms from the model and running an additional round of refinement.

### Mutagenesis

All mutations other than L265M were introduced using the QuikChange II Site-Directed Mutagenesis Kit (Agilent). The Q5 Site-Directed Mutagenesis Kit (New England Biolabs) was used to generate the L265M mutant. All constructs were sequenced prior to use.

### MD simulations

MD simulations were conducted similarly to previous studies (47). LRH-1 and SF-1 models were prepared using SF-1-TIF2-6N-10CA (chain B of PDB entry 8DAF) and LRH-1-TIF2-6N-10CA (PDB entry 7TT8) as starting structures. For SF-1 complexes, PDB entry 4QK4 (residues 239–265) was used to replace missing residues. SF-1 residues 221 to 222 and TIF2 residue 741 were removed to keep LRH-1 and SF-1 complexes consistent. Surface cysteines mutated to serine for crystallography were mutated back to cysteine for simulations. For LRH-1 complexes, PDB entry 4PLE (residues 321–345) was used to replace missing residues. LRH-1 residues 299 to 300 were removed to keep LRH-1 and SF-1 complexes consistent.

For 6N-10CA-bound structures, PDB entry 8DAF (chain B) and PDB entry 7TT8 were used as starting models. For 2N-bound constructs, the position of 2N from the previously solved crystal structure of LRH-1-2N (PDB entry 6OR1) was used to guide the ligand into the pocket of LRH-1 and SF-1. Although the density in the crystal structure was satisfied by various orientations for R313, initial simulations indicated differences in histidine interactions driven by side chain orientation. Therefore, we repeated SF-1-2N simulations with this residue reoriented to better reflect the R393 rotamer in LRH-1 complexes (see Fig. S3 for representative images). Note that we also repositioned R313 in the SF-1-6N-10CA complex used for average structure comparison and pathway analyses (Fig. 8) so that the starting deep pocket interactions in LRH-1 and SF-1 complexes were similar. Apo complexes were generated by removing 6N-10CA from the prepared LRH-1- and SF-1-6N-10CA complexes used in initial MD simulations. Maestro (Schrödinger, LLC) (73) was used to optimize hydrogen bond assignments, add N- and C-terminal caps to LRH-1, SF-1, and TIF2, and run initial minimization on the structure. The complexes were then solvated in an octahedral box of TIP3P water with a 10 Å buffer around the protein complex. Na<sup>+</sup> and Cl<sup>-</sup> ions were added to neutralize the protein and achieve physiological buffer conditions (150 mM NaCl). Parameters for 6N, 10CA, and 6N-10CA were obtained using Antechamber (74) in AmberTools20. Systems were set up using the xleap tool in AmberTools20 of Amber 2020 (75), with ff14SB (76) (protein), GAFF2 (77) (ligand), and TIP3P (78) (water) forcefields.

Minimizations and simulations were performed with Amber20. For minimization, 5000 steps of steepest descent were used, followed by 5000 steps of conjugate gradient minimization. Minimizations were first performed with 500 kcal/mol·Å<sup>2</sup> restraints on all protein and ligand atoms. Restraints were then removed on all atoms except the ligand and TIF2 peptides, and the protocol was repeated. Restraints were then removed from TIF2 atoms, and the protocol was repeated. Ligand restraints were then lowered to 100 kcal/mol·Å<sup>2</sup> for an additional round of minimization. Restraints were then removed from all atoms for two final rounds of minimization. Minimized systems were heated from 0 to 300 K with a 100 ps MD run, with constant volume periodic boundaries and 10 kcal/mol·Å<sup>2</sup> restraints on all protein and ligand atoms. A 10 ns equilibration was then performed with 10 kcal/mol·Å<sup>2</sup> restraints on all protein and ligand atoms. Restraints were then lowered to 1 kcal/mol·Å<sup>2</sup> for ligand and protein atoms and a 10 ns equilibration was performed. A final 10 ns round of equilibration was then run with only ligand atoms restrained (1 kcal/mol·Å<sup>2</sup>). Note that the water molecule critical for ligand engagement deep within the pocket was restrained along with ligand atoms. Finally, production trajectories of 500 ns were obtained for unrestrained complexes in the NPT ensemble. All bonds between heavy atoms and hydrogens were fixed with the SHAKE algorithm (79). A cutoff distance of 10 Å was used to evaluate long-range electrostatics with particle mesh Ewald and for van der Waals forces. Three (2N-bound and apo complexes, as well as the 6N-10CA-bound SF-1 complex prepared for pathway analysis) or four (6N-10CA-bound LRH-1 complexes and SF-1-6N-10CA used for initial ligand binding analysis) 500-ns simulations were run. Frames from MD simulations were concatenated with CPPTRAJ (80) for comparison of averaged 6N-10CA-bound complexes, as well as pathway analysis. Every fifth frame was used for data analysis. We excluded one 500-ns simulation from LRH-1-6N-10CA average structure and pathway analyses to keep analysis consistent with SF-1 complexes. Water, Na<sup>+</sup>, and Cl<sup>-</sup> molecules, along with N- and C-terminal caps, were removed for data analysis.

CPPTRAJ (80) was used for hydrogen bond, distance, and RMSF analyses as well as the construction of average structures. Note that hydrogen bond analysis was conducted with the default angle cutoff of 135° and distance cutoff of 3.5 Å. Frames with any interacting atoms with noted residues were counted in calculations. Note that only side chains were considered for Figure 7H (other than interaction with the nearby valine backbone). Bio3D (81) was used to create difference distance matrices comparing average structures between apo and 6N-10CA-bound complexes. Dynamic networks were constructed from trajectories using the NetworkView plugin (57) in VMD (82) and the Carma program (83). Networks are constructed by defining all protein Cα atoms as nodes, using Cartesian covariance (calculated in Carma) to measure communication within the network. Hydrogen atoms were excluded from network construction, and edges between neighboring residues were disallowed. Pairs

of nodes that reside within a 4.5 Å cutoff for 75% of the simulation are connected *via* an edge. Suboptimal paths between nodes were identified using the Floyd–Warshall algorithm (84). Matrices visualizing the number of paths between every residue on LRH-1 or SF-1 and the corresponding TIF2 peptide was determined using a loop script previously described (47), using a cutoff value of 30 in the subopt executable file. Communication with residue Y<sub>6</sub> on the TIF2 peptide (N<sub>1</sub>A<sub>2</sub>L<sub>3</sub>L<sub>4</sub>R<sub>5</sub>Y<sub>6</sub>L<sub>7</sub>L<sub>8</sub>D<sub>9</sub>K<sub>10</sub>) was used as the sink for pathway analysis, as this residue displayed strong communication with the AF-B in LRH-1-6N-10CA complexes. Suboptimal paths were visualized with VMD (82).

### Data availability

Crystal structure atomic coordinates and experimental data are available in the Protein Data Bank (PDB ID: 8DAF).

**Supporting information**—This article contains supporting information (47, 53).

**Acknowledgments**—The authors are grateful to the beamline staff at Argonne National Laboratory, Southeast Regional Collaborative Team, for support during remote collection of crystal diffraction data. We would also like to thank Anamika B. Patel, PhD (Ortlund Lab, Emory University) for her help looping crystals.

**Author contributions**—M. L. C., E. H. D., N. T. J., and E. A. O. methodology; M. L. C., E. H. D., R. M. S., and E. A. O. writing—original draft; M. L. C. and E. H. D. formal analysis; M. L. C., E. H. D., R. M. S., A. R. F., J. L. C., A. M. J., R. A. F., S. M. A., and N. T. J. investigation; E. A. O. data curation.

**Funding and additional information**—This work was supported in part by the National Institutes of Health under the following awards: F31DK122745 (M. L. C.), T32GM008367 to 27 (E. H. D.), R01DK115213 (E. A. O. and N. T. J.), and R56DK115213 (E. A. O.). E. H. D. was supported by the National Science Foundation Graduate Research Fellowship (DGE-1444932). E. A. O. was supported by The Litwin IBD Pioneers initiative and two Loan Repayment Assistance Programs from Eli Lilly and Company. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

**Conflict of interest**—The authors declare that they have no conflicts of interest with the contents of this article.

**Abbreviations**—The abbreviations used are: 6HP, hexahydropentalene; AF-B, alternative activation function surface; AF-2, activation function surface; CHAPS, 3-[(3-cholamidopropyl)dime-thylammonio]-1propanesulfonic acid; DLPC, 1,2-dilauroyl-sn-glycero-3-phosphocholine; DMSO, Dimethyl sulfoxide; DPPE, dipalmitoylphosphoethanolamine; EC50, Half maximal effective concentration; FAM, fluorescein amidite; FBS, fetal bovine serum; Ki, inhibition constant; LBD, ligand binding domain; LRH-1, liver receptor homolog-1; MD, molecular dynamics; NR, nuclear receptor; PGC-1α, peroxisome proliferator-activated receptor gamma coactivator 1-α; PL, phospholipid; RMSF, root mean square fluctuation; SEC, size exclusion chromatography; SF-1, steroidogenic factor-1; TIF2, transcriptional intermediate factor 2; Tm,

melting temperature; VMH, ventromedial nucleus of the hypothalamus.

## References

- Mangelsdorf, D. J., Thummel, C., Beato, M., Herrlich, P., Schutz, G., Umesono, K., *et al.* (1995) The nuclear receptor superfamily: the second decade. *Cell* **83**, 835–839
- Robinson-Rechavi, M., Escriva Garcia, H., and Laudet, V. (2003) The nuclear receptor superfamily. *J. Cell Sci.* **116**, 585–586
- Evans, R. M., and Mangelsdorf, D. J. (2014) Nuclear receptors, RXR, and the big bang. *Cell* **157**, 255–266
- Weikum, E. R., Liu, X., and Ortlund, E. A. (2018) The nuclear receptor superfamily: a structural perspective. *Protein Sci.* **27**, 1876–1892
- Yazawa, T., Imamichi, Y., Miyamoto, K., Khan, M. R., Uwada, J., Umezawa, A., *et al.* (2015) Regulation of steroidogenesis, development, and cell differentiation by Steroidogenic factor-1 and liver receptor homolog-1. *Zoolog. Sci.* **32**, 323–330
- Luo, X., Ikeda, Y., and Parker, K. L. (1994) A cell-specific nuclear receptor is essential for adrenal and gonadal development and sexual differentiation. *Cell* **77**, 481–490
- Parker, K. L., and Schimmer, B. P. (1997) Steroidogenic factor 1: a key determinant of endocrine development and function. *Endocr. Rev.* **18**, 361–377
- Sadovsky, Y., Crawford, P. A., Woodson, K. G., Polish, J. A., Clements, M. A., Tourtellotte, L. M., *et al.* (1995) Mice deficient in the orphan receptor steroidogenic factor 1 lack adrenal glands and gonads but express P450 side-chain-cleavage enzyme in the placenta and have normal embryonic serum levels of corticosteroids. *Proc. Natl. Acad. Sci. U. S. A.* **92**, 10939–10943
- Hammer, G. D., and Ingraham, H. A. (1999) Steroidogenic factor-1: its role in endocrine organ development and differentiation. *Front. Neuroendocrinol.* **20**, 199–223
- Wong, M., Ikeda, Y., Luo, X., Caron, K. M., Weber, T. J., Swain, A., *et al.* (1997) Steroidogenic factor 1 plays multiple roles in endocrine development and function. *Recent Prog. Horm. Res.* **52**, 167–182. discussion 182–184
- Shinoda, K., Lei, H., Yoshii, H., Nomura, M., Nagano, M., Shiba, H., *et al.* (1995) Developmental defects of the ventromedial hypothalamic nucleus and pituitary gonadotroph in the Ftz-F1 disrupted mice. *Dev. Dyn.* **204**, 22–29
- Ikeda, Y., Luo, X., Abbud, R., Nilson, J. H., and Parker, K. L. (1995) The nuclear receptor steroidogenic factor 1 is essential for the formation of the ventromedial hypothalamic nucleus. *Mol. Endocrinol.* **9**, 478–486
- Val, P., Lefrancois-Martinez, A. M., Veysiere, G., and Martinez, A. (2003) SF-1 a key player in the development and differentiation of steroidogenic tissues. *Nucl. Recept.* **1**, 8
- Mello, M. P., Franca, E. S., Fabbri, H. C., Maciel-Guerra, A. T., and Guerra-Junior, G. (2011) Multifunctional role of steroidogenic factor 1 and disorders of sex development. *Arq Bras Endocrinol. Metabol.* **55**, 607–612
- Lynch, J. P., Lala, D. S., Peluso, J. J., Luo, W., Parker, K. L., and White, B. A. (1993) Steroidogenic factor 1, an orphan nuclear receptor, regulates the expression of the rat aromatase gene in gonadal tissues. *Mol. Endocrinol.* **7**, 776–786
- Morohashi, K. I., and Omura, T. (1996) Ad4BP/SF-1, a transcription factor essential for the transcription of steroidogenic cytochrome P450 genes and for the establishment of the reproductive function. *FASEB J.* **10**, 1569–1577
- Dhillon, H., Zigman, J. M., Ye, C., Lee, C. E., McGovern, R. A., Tang, V., *et al.* (2006) Leptin directly activates SF1 neurons in the VMH, and this action by leptin is required for normal body-weight homeostasis. *Neuron* **49**, 191–203
- Majdic, G., Young, M., Gomez-Sanchez, E., Anderson, P., Szczepaniak, L. S., Dobbins, R. L., *et al.* (2002) Knockout mice lacking steroidogenic factor 1 are a novel genetic model of hypothalamic obesity. *Endocrinology* **143**, 607–614
- Kinyua, A. W., Yang, D. J., Chang, I., and Kim, K. W. (2016) Steroidogenic factor 1 in the ventromedial nucleus of the hypothalamus regulates age-dependent obesity. *PLoS One* **11**, e0162352
- Lalli, E. (2010) Adrenocortical development and cancer: focus on SF-1. *J. Mol. Endocrinol.* **44**, 301–307
- Duregon, E., Volante, M., Giorcelli, J., Terzolo, M., Lalli, E., and Papotti, M. (2013) Diagnostic and prognostic role of steroidogenic factor 1 in adrenocortical carcinoma: a validation study focusing on clinical and pathologic correlates. *Hum. Pathol.* **44**, 822–828
- Creemers, S. G., Hofland, L. J., Korpershoek, E., Franssen, G. J., van Kemenade, F. J., de Herder, W. W., *et al.* (2016) Future directions in the diagnosis and medical treatment of adrenocortical carcinoma. *Endocr. Relat. Cancer* **23**, R43–R69
- Doghman, M., Karpova, T., Rodrigues, G. A., Arhatte, M., De Moura, J., Cavalli, L. R., *et al.* (2007) Increased steroidogenic factor-1 dosage triggers adrenocortical cell proliferation and cancer. *Mol. Endocrinol.* **21**, 2968–2987
- Ruggiero, C., Doghman-Bouguerra, M., Sbiera, S., Sbiera, I., Parsons, M., Ragazzon, B., *et al.* (2017) Dosage-dependent regulation of VAV2 expression by steroidogenic factor-1 drives adrenocortical carcinoma cell invasion. *Sci. Signal.* **10**, eaal2464
- Almeida, M. Q., Soares, I. C., Ribeiro, T. C., Fragoso, M. C., Marins, L. V., Wakamatsu, A., *et al.* (2010) Steroidogenic factor 1 overexpression and gene amplification are more frequent in adrenocortical tumors from children than from adults. *J. Clin. Endocrinol. Metab.* **95**, 1458–1462
- Pianovski, M. A., Cavalli, L. R., Figueiredo, B. C., Santos, S. C., Doghman, M., Ribeiro, R. C., *et al.* (2006) SF-1 overexpression in childhood adrenocortical tumours. *Eur. J. Cancer* **42**, 1040–1043
- Sbiera, S., Schmul, S., Assie, G., Voelker, H. U., Kraus, L., Beyer, M., *et al.* (2010) High diagnostic and prognostic value of steroidogenic factor-1 expression in adrenal tumors. *J. Clin. Endocrinol. Metab.* **95**, E161–E171
- Stein, S., and Schoonjans, K. (2015) Molecular basis for the regulation of the nuclear receptor LRH-1. *Curr. Opin. Cell Biol.* **33**, 26–34
- Lee, J. M., Lee, Y. K., Mamrosh, J. L., Busby, S. A., Griffin, P. R., Pathak, M. C., *et al.* (2011) A nuclear-receptor-dependent phosphatidylcholine pathway with antidiabetic effects. *Nature* **474**, 506–510
- Mays, S. G., D'Agostino, E. H., Flynn, A. R., Huang, X., Wang, G., Liu, X., *et al.* (2022) A phospholipid mimetic targeting LRH-1 ameliorates colitis. *Cell Chem Biol.* **29**, 1174–1186.e1177
- Ortlund, E. A., Lee, Y., Solomon, I. H., Hager, J. M., Safi, R., Choi, Y., *et al.* (2005) Modulation of human nuclear receptor LRH-1 activity by phospholipids and SHP. *Nat. Struct. Mol. Biol.* **12**, 357–363
- Wang, W., Zhang, C., Marimuthu, A., Krupka, H. I., Tabrizi, M., Shelloe, R., *et al.* (2005) The crystal structures of human steroidogenic factor-1 and liver receptor homolog-1. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 7505–7510
- Li, Y., Choi, M., Cavey, G., Daugherty, J., Suino, K., Kovach, A., *et al.* (2005) Crystallographic identification and functional characterization of phospholipids as ligands for the orphan nuclear receptor steroidogenic factor-1. *Mol. Cell* **17**, 491–502
- Krylova, I. N., Sablin, E. P., Moore, J., Xu, R. X., Waite, G. M., MacKay, J. A., *et al.* (2005) Structural analyses reveal phosphatidyl inositols as ligands for the NR5 orphan receptors SF-1 and LRH-1. *Cell* **120**, 343–355
- Musille, P. M., Pathak, M., 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. *Nat. Struct. Mol. Biol.* **19**, 532–532
- Sablin, E. P., Blind, R. D., Krylova, I. N., Ingraham, J. G., Cai, F., Williams, J. D., *et al.* (2009) Structure of SF-1 bound by different phospholipids: evidence for regulatory ligands. *Mol. Endocrinol.* **23**, 25–34
- Blind, R. D., Sablin, E. P., Kuchenbecker, K. M., Chiu, H. J., Deacon, A. M., Das, D., *et al.* (2014) The signaling phospholipid PIP3 creates a new interaction surface on the nuclear receptor SF-1. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 15054–15059
- D'Agostino, E. H., Flynn, A. R., Cornelison, J. L., Mays, S. G., Patel, A., Jui, N. T., *et al.* (2020) Development of a versatile and sensitive direct ligand binding assay for human NR5A nuclear receptors. *ACS Med. Chem. Lett.* **11**, 365–370
- Sablin, E. P., Blind, R. D., Uthayaruban, R., Chiu, H. J., Deacon, A. M., Das, D., *et al.* (2015) Structure of liver receptor homolog-1 (NR5A2) with PIP3 hormone bound in the ligand binding pocket. *J. Struct. Biol.* **192**, 342–348



40. Lin, B. C., Suzawa, M., Blind, R. D., Tobias, S. C., Bulun, S. E., Scanlan, T. S., *et al.* (2009) Stimulating the GPR30 estrogen receptor with a novel tamoxifen analogue activates SF-1 and promotes endometrial cell proliferation. *Cancer Res.* **69**, 5415–5423
41. Bryant, J. M., Malabanan, M. M., Vanderloop, B. H., Nichols, C. M., Haratipour, Z., Poon, K. T., *et al.* (2021) The acyl chains of phosphoinositide PIP3 alter the structure and function of nuclear receptor steroidogenic factor-1. *J. Lipid Res.* **62**, 100081
42. Mullaney, B. C., Blind, R. D., Lemieux, G. A., Perez, C. L., Elle, I. C., Faergeman, N. J., *et al.* (2010) Regulation of *C. elegans* fat uptake and storage by acyl-CoA synthase-3 is dependent on NR5A family nuclear hormone receptor nhr-25. *Cell Metab.* **12**, 398–410
43. Blind, R. D., Suzawa, M., and Ingraham, H. A. (2012) Direct modification and activation of a nuclear receptor-PIP(2) complex by the inositol lipid kinase IPMK. *Sci. Signal.* **5**, ra44
44. Mays, S. G., Flynn, A. R., Cornelison, J. L., Okafor, C. D., Wang, H., Wang, G., *et al.* (2019) Development of the first low nanomolar liver receptor homolog-1 agonist through structure-guided design. *J. Med. Chem.* **62**, 11022–11034
45. Whitby, R. J., Stec, J., Blind, R. D., Dixon, S., Leesnitzer, L. M., Orband-Miller, L. A., *et al.* (2011) Small molecule agonists of the orphan nuclear receptors steroidogenic factor-1 (SF-1, NR5A1) and liver receptor homolog-1 (LRH-1, NR5A2). *J. Med. Chem.* **54**, 2266–2281
46. Flynn, A. R., Mays, S. G., Ortlund, E. A., and Jui, N. T. (2018) Development of hybrid phospholipid mimics as effective agonists for liver receptor homolog-1. *ACS Med. Chem. Lett.* **9**, 1051–1056
47. Cato, M. L., Cornelison, J. L., Spurlin, R. M., Courouble, V. V., Patel, A. B., Flynn, A. R., *et al.* (2022) Differential modulation of nuclear receptor LRH-1 through targeting buried and surface regions of the binding pocket. *J. Med. Chem.* **65**, 6888–6902
48. Madoux, F., Li, X., Chase, P., Zastrow, G., Cameron, M. D., Konkright, J. J., *et al.* (2008) Potent, selective and cell penetrant inhibitors of SF-1 by functional ultra-high-throughput screening. *Mol. Pharmacol.* **73**, 1776–1784
49. Del Tredici, A. L., Andersen, C. B., Currier, E. A., Ohrmund, S. R., Fairbairn, L. C., Lund, B. W., *et al.* (2008) Identification of the first synthetic steroidogenic factor 1 inverse agonists: pharmacological modulation of steroidogenic enzymes. *Mol. Pharmacol.* **73**, 900–908
50. Doghman, M., Cazareth, J., Douguet, D., Madoux, F., Hodder, P., and Lalli, E. (2009) Inhibition of adrenocortical carcinoma cell proliferation by steroidogenic factor-1 inverse agonists. *J. Clin. Endocrinol. Metab.* **94**, 2178–2183
51. Whitby, R. J., Dixon, S., Maloney, P. R., Delerive, P., Goodwin, B. J., Parks, D. J., *et al.* (2006) Identification of small molecule agonists of the orphan nuclear receptors liver receptor homolog-1 and steroidogenic factor-1. *J. Med. Chem.* **49**, 6652–6655
52. Mays, S. G., Okafor, C. D., Whitby, R. J., Goswami, D., Stec, J., Flynn, A. R., *et al.* (2016) Crystal structures of the nuclear receptor, liver receptor homolog 1, bound to synthetic agonists. *J. Biol. Chem.* **291**, 25281–25291
53. Cornelison, J. L., Cato, M. L., Johnson, A. M., D'Agostino, E. H., Melchers, D., Patel, A. B., *et al.* (2020) Development of a new class of liver receptor homolog-1 (LRH-1) agonists by photoredox conjugate addition. *Bioorg. Med. Chem. Lett.* **30**, 127293
54. Mays, S. G., Stec, J., Liu, X., D'Agostino, E. H., Whitby, R. J., and Ortlund, E. A. (2020) Enantiomer-specific activities of an LRH-1 and SF-1 dual agonist. *Sci. Rep.* **10**, 22279
55. Musille, P. M., Kossmann, B. R., Kohn, J. A., Ivanov, I., and Ortlund, E. A. (2016) Unexpected allosteric network contributes to LRH-1 Co-regulator selectivity. *J. Biol. Chem.* **291**, 1411–1426
56. Vanwart, A. T., Eargle, J., Luthey-Schulten, Z., and Amaro, R. E. (2012) Exploring residue component contributions to dynamical network models of allostery. *J. Chem. Theor. Comput.* **8**, 2949–2961
57. Sethi, A., Eargle, J., Black, A. A., and Luthey-Schulten, Z. (2009) Dynamical networks in tRNA: protein complexes. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 6620–6625
58. Gasper, P. M., Fuglestad, B., Komives, E. A., Markwick, P. R., and McCammon, J. A. (2012) Allosteric networks in thrombin distinguish procoagulant vs. anticoagulant activities. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 21216–21222
59. Mays, S. G., Okafor, C. D., Tuntland, M. L., Whitby, R. J., Dharmarajan, V., Stec, J., *et al.* (2017) Structure and dynamics of the liver receptor homolog 1-PGC1alpha complex. *Mol. Pharmacol.* **92**, 1–11
60. Knarston, I. M., Robevska, G., van den Bergen, J. A., Eggers, S., Croft, B., Yates, J., *et al.* (2019) NR5A1 gene variants repress the ovarian-specific WNT signaling pathway in 46,XX disorders of sex development patients. *Hum. Mutat.* **40**, 207–216
61. Suntharalingham, J. P., Buonocore, F., Duncan, A. J., and Achermann, J. C. (2015) DAX-1 (NR0B1) and steroidogenic factor-1 (SF-1, NR5A1) in human disease. *Best Pract. Res. Clin. Endocrinol. Metab.* **29**, 607–619
62. Biason-Lauber, A., and Schoenle, E. J. (2000) Apparently normal ovarian differentiation in a prepubertal girl with transcriptionally inactive steroidogenic factor 1 (NR5A1/SF-1) and adrenocortical insufficiency. *Am. J. Hum. Genet.* **67**, 1563–1568
63. Philibert, P., Paris, F., Lakhal, B., Audran, F., Gaspari, L., Saad, A., *et al.* (2013) NR5A1 (SF-1) gene variants in a group of 26 young women with XX primary ovarian insufficiency. *Fertil. Steril.* **99**, 484–489
64. McWilliam, H., Li, W., Uludag, M., Squizzato, S., Park, Y. M., Buso, N., *et al.* (2013) Analysis tool web services from the EMBL-EBI. *Nucl. Acids Res.* **41**, W597–W600
65. Goujon, M., McWilliam, H., Li, W., Valentin, F., Squizzato, S., Paern, J., *et al.* (2010) A new bioinformatics analysis tools framework at EMBL-EBI. *Nucl. Acids Res.* **38**, W695–W699
66. Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., *et al.* (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539
67. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Met. Enzymol.* **276**, 307–326
68. Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., *et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221
69. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132
70. Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., *et al.* (2012) Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D Biol. Crystallogr.* **68**, 352–367
71. Joosten, R. P., Long, F., Murshudov, G. N., and Perrakis, A. (2014) The PDB\_REDO server for macromolecular structure model optimization. *IUCr* **1**, 213–220
72. The PyMOL Molecular Graphics System, Version 2.5. (2015). Schrödinger, LLC, New York, NY
73. Maestro (2021). Schrödinger, LLC, New York, NY
74. Wang, J. M., Wang, W., and Kollman, P. A. (2001) Antechamber: an accessory software package for molecular mechanical calculations. *Abstr. Pap. Am. Chem. Soc.* **222**, U403
75. Case, D. A., Belfon, K., Ben-Shalom, I. Y., Brozell, S. R., Cerutti, D. S., Cheatham, T. E. I., *et al.* (2020) AMBER 2020, University of California, San Francisco
76. Maier, J. A., Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K. E., and Simmerling, C. (2015) ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. *J. Chem. Theor. Comput.* **11**, 3696–3713
77. Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A., and Case, D. A. (2004) Development and testing of a general amber force field. *J. Comput. Chem.* **25**, 1157–1174
78. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935
79. Ryckaert, J.-P., Ciccotti, G., and Berendsen, H. J. C. (1977) Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* **23**, 327–341

## Structure/activity of SF-1- and LRH-1-modulator complexes

80. Roe, D. R., and Cheatham, T. E., 3rd. (2013) PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data. *J. Chem. Theor. Comput.* **9**, 3084–3095
81. Grant, B. J., Rodrigues, A. P., ElSawy, K. M., McCammon, J. A., and Caves, L. S. (2006) Bio3d: an R package for the comparative analysis of protein structures. *Bioinformatics* **22**, 2695–2696
82. Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD: visual molecular dynamics. *J. Mol. Graph* **14**, 27–38
83. Glykos, N. M. (2006) Software news and updates. Carma: a molecular dynamics analysis program. *J. Comput. Chem.* **27**, 1765–1768
84. Floyd, R. W. (1962) Algorithm 97: shortest path. *Commun. ACM* **5**. <https://doi.org/10.1145/367766.368168>