

Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

Paul M. Musille

Date

Structural and Functional Studies of the Orphan Nuclear Receptor LRH-1

By
Paul M. Musille
Doctor of Philosophy

Graduate Division of Biological and Biomedical Sciences
Molecular and Systems Pharmacology

Eric A. Ortlund, PhD
Advisor

T.J. Murphy, PhD
Committee Member

Graeme Conn, PhD
Committee Member

Rita Nahta, PhD
Committee Member

Jing Chen, PhD
Committee Member

Accepted:

Lisa A. Tedesco, PhD
Dean of the Graduate School

Date

Structural and Functional Studies of the Orphan Nuclear Receptor LRH-1

By

Paul M. Musille
B.S., Gannon University, 2008

Advisor: Eric A. Ortlund, PhD

An abstract of a dissertation submitted to the Faculty of the James T. Laney School of Graduate Studies of Emory University in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Graduate Division of Biological and Biomedical Sciences
Molecular and Systems Pharmacology
2014

Abstract

The NR5A subfamily family of nuclear receptors are important regulators of pluripotency, lipid and glucose homeostasis, and steroidogenesis. Liver receptor homologue 1 (LRH-1; NR5A2) has therapeutic potential for the treatment of metabolic and neoplastic disease; however, a poor understanding of its ligand regulation has hampered the pursuit of these proteins as pharmaceutical targets. It was previously thought that LRH-1 was an intractable therapeutic target due to its orphan classification, but structural and biochemical studies combined with extensive small molecule screening have since shown that LRH-1 is a tractable drug target. LRH-1 is known to bind phospholipids (PLs) but the role of PLs in controlling LRH-1 activation remains highly debated. An improved understanding of the molecular determinants of LRH-1 activation may enhance efforts to target LRH-1 with therapeutics. The data presented here use a variety of structural and biochemical techniques to probe how phospholipid ligands, coregulator proteins and evolutionary changes alter LRH-1's activation state.

Structural and Functional Studies of the Orphan Nuclear Receptor LRH-1

By

Paul M. Musille
B.S., Gannon University, 2008

Advisor: Eric A. Ortlund, PhD

A dissertation submitted to the Faculty of the James T. Laney School of Graduate Studies of
Emory University in partial fulfillment of the requirements for the degree of Doctor of
Philosophy

Graduate Division of Biological and Biomedical Sciences
Molecular and Systems Pharmacology
2014

Table of Contents

Chapter 1: Introduction	1
Introduction	2
<i>Phospholipids</i>	2
<i>Nuclear Receptors: lipid regulated transcription factors</i>	4
Case Studies	6
<i>SF-1</i>	9
<i>PPARs</i>	11
<i>USP</i>	13
PL transport and PL dependent coactivation	13
<i>PPAR and PC-TP</i>	14
Structural Analysis of PL binding proteins.....	14
<i>What does it take to bind to PLs as a ligand?</i>	14
<i>Shuttlers versus transcription factors</i>	15
<i>Parallels in the immune system</i>	15
<i>Comparison to the PL PI/PC transporter Sec14</i>	17
<i>PL presentation as a model for PL dependent signaling</i>	17
Closing Remarks	18
LRH-1 Knowledge Gaps.....	20
Chapter 2: Antidiabetic phospholipid-nuclear receptor complex reveals the mechanism for phospholipid-driven gene regulation.....	22
Introduction.....	23
Experimental Procedures	24
Results.....	30
<i>Structure of the activated LRH-1–DLPC–TIF2 NR box 3 complex</i>	30

<i>Phospholipid tail composition drives differential receptor dynamics</i>	34
<i>Generating apo LRH-1</i>	37
<i>Ligand Binding Alters Co-regulator Preference</i>	37
<i>β-sheet–helix 6 mobility is essential for PL driven transactivation</i>	42
<i>Structure of the apo LRH-1–SHP NR Box1 complex</i>	42
<i>Empty receptor binds with high affinity to SMRT</i>	47
Discussion	47
Chapter 3: Divergent sequence tunes ligand sensitivity in phospholipid-regulated hormone receptors	50
Introduction.....	51
Experimental Procedures	54
Results	58
<i>Overall Structure</i>	58
<i>Apo mLRH-1 adopts a destabilized active orientation</i>	61
<i>Sequence divergence allows to NR5A receptors adopt multiple conformations to achieve the active state</i>	63
<i>Do mLRH-1 and Ftz-F1 bind PLs despite their ability to crystallize empty?</i>	63
<i>mLRH-1 is specifically stabilized versus hLRH-1</i>	67
Discussion	70
Chapter 4: Unexpected Allosteric Network Drives Nuclear Receptor-Phospholipid Signaling	74
Introduction.....	75
Experimental Procedures	77
Results.....	80
<i>Structure of the apo LRH-1 LBD–TIF complex</i>	80

<i>Structure of the LRH-1 LBD–E. coli PL–TIF2 complex</i>	80
<i>Co-regulator binding interactions are altered by ligand status</i>	84
Discussion	85
Chapter 5: Discussion	87
Phospholipids as signaling molecules.....	88
LRH-1 is an untapped pharmacological target	88
LRH-1’s mode of activation is unique among NRs.....	88
Sequence Divergence throughout evolution has altered the structural mechanisms driving LRH-1 activation	89
LRH-1 structures as invaluable tools to understand the detailed molecular mechanism driving receptor transactivation and transrepression	90
Future Directions	90
Concluding Remarks.....	91
Appendix: Structure of Glycerol Dehydrogenase From <i>Serratia</i>	94
Introduction.....	95
Experimental Procedures	96
Results and Discussion	105
References	114

List of Figures

Figure 1.1 Structures of major phospholipid species.....	3
Figure 1.2 Crystal structures of soluble PL signaling proteins.....	16
Figure 1.3 Phospholipid mediated transcription control.....	19
Figure 1.4 Published LRH-1 structures before 2010.....	21
Figure 2.1 DLPC binds directly to LRH-1 and promotes activation through unique interactions.....	31
Figure 2.2 Ligplot representation of bound PLs.....	35
Figure 2.3 DLPC alters LRH-1 stability and structural dynamics.....	36
Figure 2.4 DLPC binding affects LRH-1 dynamics and co-regulator preference.....	38
Figure 2.5 Characterization of apo LRH-1.....	41
Figure 2.6 Structure of apo LRH-1 identifies a novel mobile activation function region.....	44
Figure 2.7 Percent deuterium incorporation over time.....	45
Figure 2.8 Structure of apo LRH-1 Displays Disorder around LBP.....	46
Figure 3.1 Overall Structure of apo mLRH-1.....	59
Figure 3.2 Opening to the LBP in the human and mLRH-1.....	62
Figure 3.3 NR5A receptors have multiple modes for achieving the active conformation in the absence of ligand.....	64
Figure 3.4 Phospholipid analysis of mLRH-1 and Ftz-F1.....	65
Figure 3.5 Apo mLRH-1 is more active than wt LRH-1 by enhancing stability in the alternate activation region.....	68
Figure 3.6 Coordination at the mouth of LRH-1's LBP is important for receptor activation.....	72
Figure 4.1 Structure of the apo LRH-1 LBD-TIF complex.....	81
Figure 4.2 Structure of the LRH-1 LBD- <i>E. coli</i> PL-TIF2 complex.....	82
Figure 4.3 AF-2 charge clamp engagement is dictated by ligand coregulator combination.....	83
Figure 5.1 LRH-1 Activation States Model.....	93

Appendix Figure 1 Protein Purification	97
Appendix Figure 2 Crystals of <i>Serratia</i> glycerol dehydrogenase.....	99
Appendix Figure 3 Peptide Identification.....	104
Appendix Figure 4 Overall Structure.....	106
Appendix Figure 5 Sequence Confirmation	108
Appendix Figure 6 Superpositions of <i>Serratia</i> glycerol dehydrogenase	109
Appendix Figure 7 Conserved NAD Binding Site	111
Appendix Figure 8 Active Site of <i>Serratia</i> Glycerol Dehydrogensase.....	112

List of Tables

Table 2.1 Data collection and refinement statistics	33
Table 2.2 Coactivator Peptide Recruitment	40
Table 2.3 Proteolysis protection data showing spectral counts for trypsin fragments.....	43
Table 3.1 Data collection and refinement statistics for the mILRH-1 crystal structure.....	60
Table 3.2 Coactivator Peptide Recruitment.....	69
Table 4.1 Data collection and refinement statistics	79
Appendix Table 1 Crystallographic Statistics.....	101

List of Abbreviations

2-(bis(2-hydroxyethyl)- amino)-2-(hydroxymethyl)propane-1,3-diol (Bis-Tris)

20-hydroxyecdysone (20E)

activation function 1 (AF-1)

activation function helix (AF-H)

atomic mass units (amu)

collisionally induced dissociation (CID)

diacylglycerol (DAG)

dilauroylphosphatidylcholine (DLPC, PC 12:0/12:0)

dipalmitoylphosphatidylcholine (DPPC; PC 16:0–16:0)

diundecanoylphosphatidylcholine (DUPC, PC 11:0/11:0)

DNA binding domain (DBD)

Drosophila FTZ-f1 (dmFTZ-f1)

ecdysone receptor (EcR)

electrospray injection mass spectrometry (ESI/MS)

estrogen receptor (ER)

fatty acid synthase (FAS)

free fatty acids (FFAs)

G protein-coupled receptors (GPCRs)

general transcription factors (GTFs)

human LRH-1 (hLRH-1)

hydrogen-deuterium exchange mass spectroscopy (HDX-MS)

inositol trisphosphate (IP3)

ligand binding domain (LBD)

ligand binding pocket (LBP)

liver receptor homolog 1 (LRH-1)
maltose binding protein (MBP)
mouse LRH-1 (mLRH-1)
mouse-loop LRH-1 (mlLRH-1)
nuclear receptor (NR)
nuclear receptor co-repressor 1 (NCoR)
nuclear receptor co-repressor 2 (SMRT)
nuclear receptor coactivator 2 (TIF)
peroxisome proliferator activated receptor gamma coactivator 1- α (PGC-1 α)
peroxisome proliferator-activated receptor alpha (PPAR α)
phosphatidic acid (PA)
phosphatidyl glycerol (PG)
phosphatidylcholine (PC)
phosphatidylcholine transfer protein (PC-TP)
phosphatidylinositol (PI)
phosphatidylserine (PS)
phospholipase A (PLA)
Phospholipids (PLs)
PI-bisphosphate (PIP2)
PI-trisphosphate (PIP3)
polyethylene glycol (PEG)
Ras homolog gene family, member A (RhoA)
retinoid X receptor (RXR)
root-mean-square deviation (rmsd)
small heterodimer partner (SHP)
Southeast Regional Collaborative Access Team (SER-CAT)

sphingomyelin (SM)

sphingosine-1-phosphate (S1P)

steroidogenic acute regulatory protein (StAR)-related lipid transfer (START)

steroidogenic factor 1 (SF-1)

transcriptional intermediary factor-2 (TIF2)

Ultraspiracle protein (USP)

Chapter 1: Introduction¹

¹This chapter has been slightly modified from the published manuscript: Musille PM, Kohn JA, Ortlund EA. (2013) Phospholipid-driven gene regulation. *FEBS Lett.* 587(8):1238-46.

Introduction

Phospholipids

Phospholipids (PLs) are ubiquitous to all forms of life serving as the major constituent of the membranes that isolate and protect cells from their external environment, and segregate organelles from the greater cellular milieu. PLs are composed of two hydrophobic tails, donated by a diacylglycerol (DAG), and a hydrophilic head group containing a phosphate, which is frequently conjugated to an additional hydrophilic metabolite (Figure 1.1). This amphipathic, bipartite structure drives their spontaneous assembly into bilayers, which compartmentalize the cell and harbor an assortment of proteins, glycans, and other lipids that play critical roles in cell structure, function, metabolism, and signaling.

PLs as signaling molecules

Though best known for their role in membrane construction, PLs play integral roles in a number of cellular signaling cascades at and within the membrane bilayer (1). Arguably the most familiar of these are the IP₃/DAG and Akt cascades. In the former, membrane-bound PI-bisphosphate (PIP₂) is cleaved by PLC to yield inositol trisphosphate (IP₃) and DAG; IP₃ is released into the cytoplasm and triggers the release of Ca²⁺ from the endoplasmic reticulum, while DAG remains in the plasma membrane and activates PKC (2). PI-trisphosphate (PIP₃) is instrumental in recruiting Akt to the plasma membrane, where it is activated by PDK-1 (3). In more recent years, additional PL derivatives have been implicated in cell signaling. Lysophospholipids, single-chain PLs that include sphingosine-1-phosphate (S1P) and lysophosphatidic acid (LPA), were found to bind and activate G protein-coupled receptors (GPCRs) upstream of Ras homolog gene family, member A (RhoA) activation, affecting numerous signaling responses (4). Furthermore, a family of tail-oxidized PLs are now known to play central roles in the regulation of the plasma membrane and the innate immune system (5).

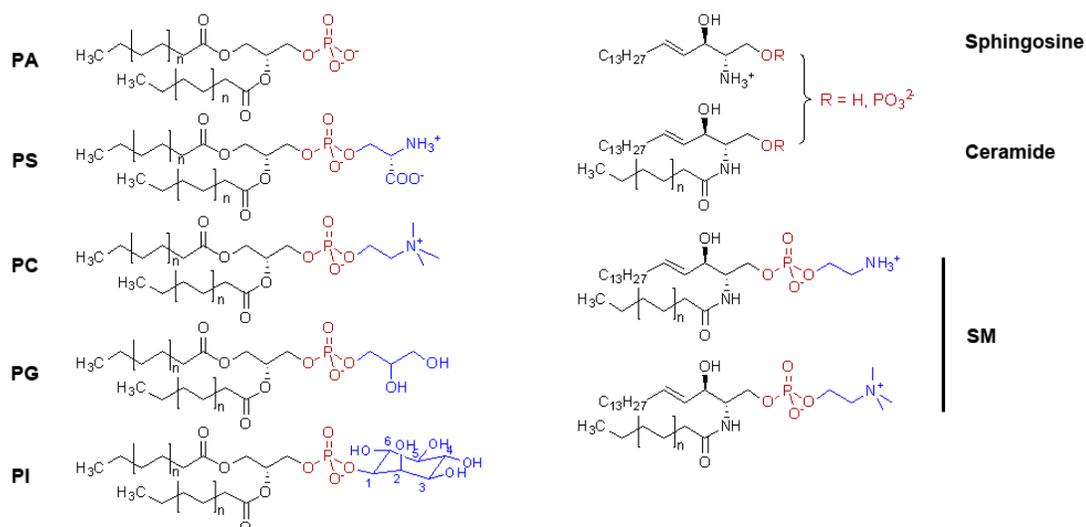


Figure 1.1 Structures of major phospholipid species.

PLs consist of a hydrophobic diacyl tail (black), a phosphate (red), and a polar head group (blue).

PA: phosphatidic acid; PS: phosphatidylserine; PC: phosphatidylcholine; PG: phosphatidyl glycerol; PI: phosphatidylinositol; SM: sphingomyelin.

PLs have therefore emerged as key players in the signal cascades that control many vital biological processes.

PLs outside the membrane

A significant fraction of the cellular PL pool resides outside of the membrane, particularly inside the nucleus. While some of this subpopulation may have structural roles as part of chromatin or the nuclear lamin (6), it is now evident that there is a PL signaling system distinct from that which occurs within the membrane bilayer (7). PIs again are at the core of the known nuclear lipid signaling pathways (8), and while the nature of nuclear PLs remains enigmatic, it is now understood that PI and PIPs have important functions in the regulation of protein-chromatin interactions (9). The close association of PLs with DNA (10) suggests that, in addition to their roles in cell structure and signal transduction, PLs play a role in driving gene expression and regulation.

PLs are a new class of hormone

Ernest Starling coined the term “hormone” in 1905, long before the isolation of the first nuclear receptor (NR) in 1958, to describe a substance that is able to travel throughout an organism serving as a chemical messenger to alter cell behavior. PLs have long been thought of as synthesis material for some hormones, but new evidence suggests they are transmitting their own unique signals to alter transcriptional patterns. The vast majority of evidence for direct PL-mediated transcription is among the NR family of transcription factors.

Nuclear Receptors: lipid regulated transcription factors

Nuclear receptor structure and function

NRs are a family of ligand regulated transcription factors that are activated by a diverse group of lipophilic ligands including fatty acids, cholesterol derivatives, steroid hormones, vitamins, dietary components, and xenobiotics (11-14). These ligands, primarily derived from

lipids, act as messengers by transmitting chemical information that reflects the body's nutritional and endocrine states (15). This allows for the coordination of growth, reproduction, and homeostasis, and allows the body to appropriately respond to events, such as eating a meal, exercise, or stress.

NRs share a highly conserved multi-domain architecture including a variable N-terminal domain, often referred to as the activation function 1 (AF-1), a DNA binding domain (DBD), a flexible linker region, and a ligand binding domain (LBD) that contains a ligand sensitive transcriptional switch, the AF-2 (12; 13). Ligand dependent NR activation is centered on the LBD, a helical bundle containing a lipophilic cavity that can accommodate ligands. The hydrophobic pockets within NRs typically vary in size and shape to match their cognate hormone (13; 14). A mobile ligand sensing helix, termed the activation function helix (AF-H), responds to a bound ligand by rotating and packing against the LBD. This repositioning completes the AF-2 surface, enabling interaction with coactivator proteins contained in chromatin modifying complexes that promote gene transcription (12). In the absence of ligand, NRs preferentially interact with corepressor complexes which displace the "active AF-H" from the body of the protein resulting in transcriptional repression (12). Similarly, NR antagonists alter AF-H positioning to either prevent coactivator binding or promote binding of corepressor proteins to inhibit transcription.

NRs ligands are invariably hydrophobic and freely diffuse across membranes to allow for long-range signal transmission. In this way, hormones affect diverse groups of gene programs involved in pathophysiology ranging from diabetes to cancer, making NRs ideal targets for pharmacological intervention. As such, NR-targeting drugs have a myriad of uses ranging from cancer treatments, and contraceptives, to treating allergic reactions and metabolic disorders and represent a major industrial and academic investment in basic research and drug development (14; 16; 17).

PL-driven NR activation

To date, four NRs have been identified as PL-binding proteins: liver receptor homolog 1 (LRH-1) and steroidogenic factor 1 (SF-1), members of the NR5a class of steroidogenic factor-like NRs; peroxisome proliferator-activated receptor alpha (PPAR α), a member of the NR1 thyroid hormone receptor-like family of receptors; and ultraspiracle (USP), the insect homolog of the retinoid X receptor. This review will focus on the compelling evidence for PLs role in regulating these receptors, as well as a family of PL transporters that stimulate NR transactivation.

Case Studies

LRH-1

LRH-1 is a member of the NR5, or Ftz-f1, subfamily of NR's, and regulates the expression of genes involved in development, lipid and glucose homeostasis, steroidogenesis, and cell proliferation (18; 19). During the early stages of development, LRH-1 is responsible for maintaining levels of OCT-4, considered to be a master regulator of pluripotency (20). Disruption of the LRH-1 gene in mice leads to the loss of Oct4 expression in the epiblast, causing lethality at embryonic day 6.5 (21). Over expression of LRH-1 is sufficient to reprogram murine somatic cells to pluripotent cells without simultaneous overexpression of OCT-4. This makes LRH-1 the only known transcription factor that can replace OCT-4 in the cellular reprogramming identifying it as a new stem cell factor (22). It is unknown what role LRH-1 plays in OCT4 regulation beyond development, however, the receptor was recently shown to regulate OCT4 expression in human cancer stem cells (23).

In adults, LRH-1 is expressed in liver, pancreas, intestine, brain and sex glands such as the ovaries and placenta (18; 24). In the liver, LRH-1 is a master regulator of lipid homeostasis (19) regulating bile acid and cholesterol flux through regulation of CYP7A1, which catalyzes the rate-limiting step in bile acid synthesis (18). LRH-1 also regulates the transcription a number of other lipid, bile, and cholesterol synthesis enzymes and transporters required in the processes of lipid

transport to the liver and elimination (25-32). Recently, LRH-1 has been identified as a direct transcriptional regulator of glucokinase, responsible for glucose capture in the liver (33). Disruption of the LRH-1 gene in healthy livers not only disrupted lipogenesis but also resulted in reduced glycogen synthesis and glycolysis in response to acute and prolonged glucose exposure. Taken together, these studies demonstrate LRH-1's influences on metabolic homeostasis by linking PL levels to glucose and lipid metabolism.

LRH-1 is also expressed in preadipocytes and adipocytes surrounding estrogen receptor positive breast cancer cells (24). Here, in conjunction with GATA and protein kinase A, LRH-1 drives the expression of *CYP19* (aromatase), increasing the local estrogen concentration to fuel tumor growth (24; 34). Additionally, LRH-1 appears to take part in a positive feedback loop with active estrogen receptor further enhancing these effects (35).

In the colon, LRH-1 plays a markedly different role in cancer development and progression. Here, LRH-1 has been shown to synergize with the beta-catenin/TCF transcriptional complex to enhance the expression of cell proliferation, growth and survival genes such as cyclin's D1 and E1 (21). Additionally, LRH-1 has also been found to be overexpressed in gastric cancer (36).

Bound E. coli PLs offer the first clue that LRH-1 may be PL regulated.

In 2003, the crystal structure of mouse LRH-1 was reported, showing the receptor held in an active conformation in the absence of a ligand or co-regulatory peptide (37). This structure suggested that LRH-1 might act in a ligand-independent manner, discouraging efforts to pursue LRH-1 as a drug target despite its therapeutic potential. In 2005, however, subsequent crystal structures of human LRH-1 all revealed a large $>1,400 \text{ \AA}^3$ ligand binding pocket (LBP) occupied by a diverse array of PLs including PG, PE, and a rare phosphatidylglycerol-phosphoglycerol (38-40). Mutations designed to reduce PL binding showed decreased transcriptional activity in reporter gene assays and a decrease in the ability to recruit coregulators and coregulator

fragments both *in vitro* and in cells (39; 41). These exciting new findings showed for the first time that LRH-1 might be regulated by PLs.

LRH-1 – PIP interactions

To identify plausible mammalian PL ligands, Krylova et al. assessed binding of LRH-1 to immobilized PLs which revealed that LRH-1 bound to a range of PLs, but bound most strongly to PIP2 and PIP3 species (40). Lipid binding was confirmed through non-denaturing mass spectrometry (40). LBP mutations designed to prevent lipid binding decreased the ability of LRH-1 to bind these immobilized lipids (40). Notably, this assay did not show PC binding for either LRH-1 or SF-1 (40), both of which were later shown to be activated by PC in cells and bind PC *in vitro* (41; 42).

DLPC

Recently, Lee et al. showed that both human and mouse LRH-1 are specifically activated by the exogenous medium chain phosphatidylcholine isoforms – diundecanoyl (DUPC, PC 11:0/11:0) and dilauroyl (DLPC, PC 12:0/12:0) phosphatidylcholine (43). These medium chain PC agonists selectively activate the receptor in luciferase assays, increase the ability of LRH-1 to interact with the coactivators and increase the production of LRH-1 target genes (43). Moreover, DLPC lowers serum lipid levels and reduces blood glucose levels in diabetic mice in a LRH-1 dependent manner (43). The X-ray crystal structure of the LRH-1–DLPC complex in combination with hydrogen-deuterium exchange assays confirmed that DLPC interacts directly with LRH-1 and revealed the mechanism dictating DLPC-driven transcriptional activation (41). Unlike other NRs that rely on intra-protein interactions to coordinate activation, LRH-1 relies on intramolecular contacts between distal residues in the LBP and the PL to sense and transmit ligand status to the AF-H (41). Additionally, generation and characterization of apo LRH-1, showed that ligand free LRH-1 LBD has a highly destabilized structure that is profoundly stabilized by lipids (41). DLPC simultaneously enhanced co-activator peptide recruitment while

disfavoring repressor peptide interaction (41). These recent results show for the first time that LRH-1 is able to dynamically respond to a PL ligand.

SF-1

SF-1, another member of the Ftz-F1 NR5A subfamily, is a key regulator of steroidogenesis and the development of steroidogenic organs, such as the adrenal cortex and gonads (44). It is expressed primarily in these tissues, and in tissues along the steroid hormone regulatory axes, including the hypothalamus and pituitary gland (45; 46). Genes involved in nearly all stages of steroid biosynthesis are regulated by SF-1, including those that encode HMG-CoA synthase (47), cholesterol transporters (48-50), 3 β steroid dehydrogenase, and many of the cytochrome P450 enzymes that catalyze the conversion of cholesterol into steroid hormones (51).

Dysfunction of SF-1 has been linked to a number of human disorders (52; 53). Mutations in SF-1 have been detected in patients with disorders in sexual development (54-57), ovarian insufficiency (55), and adrenal failure (56), while SF-1 dysregulation has been linked to endometriosis (58) and adrenocortical carcinoma (59). Like LRH-1, SF-1 makes an alluring drug target, yet a robust understanding of its ligand-binding properties is only now emerging.

However, some headway has been made in identifying synthetic compounds that act upon SF-1. In 2008, a number of inverse agonists for SF-1 were identified (60-62). Not only could these compounds inhibit SF-1-dependent gene transcription in luciferase assays, they also inhibited StAR expression in human adrenocortical cells (60), suggesting a possible therapeutic value in the treatment of adrenocortical cancers. Isoquinolone-derived inverse agonists were subsequently shown to inhibit the expression of CYP21 and CYP17 mRNA *in vitro*, with a concurrent reduction in the secretion of aldosterone, cortisol, and DHEA-S, and inhibition of adrenocortical carcinoma cell proliferation (5; 63). These results indicate that pharmacological modulation of SF-1 may be a viable strategy in treating adrenocortical carcinomas, and possibly

other human diseases. However, more research is needed to understand the intricacies of ligand-driven SF-1 activity, before its full potential as a drug target can be realized.

E. coli PL binding from early structural studies

The first crystal structures of SF-1 were reported in 2005, showing the LBD in complex with copurified *E. coli* medium chain PG and PE species (38; 40; 64). The binding of SF-1 to immobilized eukaryotic PLs was tested along with LRH-1, and it was found that SF-1 could bind to an array of PL species, including PA, PI, PIP₂, and PIP₃, with a preference for PIPs phosphorylated at the 3- and 5-carbons (40). Coactivator recruitment was enhanced by PEs (38; 64) and PCs (64) identifying diverse PLs as activating ligands *in vitro*.

PA versus sphingosine

The discovery that SF-1 could bind exogenous PLs intensified the search for its endogenous ligands. By 2007, mass spectrometry experiments had identified sphingosine, lysoSM, PA, PE, and PI bound to SF-1 that had been immunoprecipitated from human adrenocarcinoma cells (65; 66). Further analysis showed that sphingosine acts as a SF-1 antagonist, blocking cAMP-stimulated CYP17 reporter gene activity and coactivator recruitment, which could be negated by inhibiting the acid ceramidases that produce sphingosine from ceramide, or by introducing mutations into the LBP that abrogated sphingosine binding (65). Subsequently, it was found that PA activated SF-1-dependent CYP17 expression and transcriptional activity, SF-1 heterocomplex assembly, and steroidogenesis. These effects could be inhibited by sphingosine or by LBP mutations (66).

These data suggest a model, wherein SF-1 is maintained in an inactive conformation by sphingosine under basal conditions (65; 67) and is activated by the binding of PA, which is generated subsequent to ACTH/cAMP signaling (66). The two different lipid species have opposing effects on the activity of SF-1, suggesting a regulatory mechanism in which the levels of these two lipids control the expression of genes linked to SF-1.

PIP2 versus PIP3

While no structures of a SF-1–PI or SF-1–PIP complex have been reported, modeling studies showed that phosphorylated PIs may be stabilized by several histidine residues around the mouth of the SF-1 LBP (68). Mutations to these residues greatly impaired exchange of bacterial PG with PIP2 and PIP3 and diminished SF-1 transcriptional activity, suggesting that the binding of PIPs to SF-1 is a biologically relevant interaction (68). Indeed, IPMK phosphorylates PIP2 only when bound to SF-1, increasing downstream gene transcription; likewise, PTEN cleaves PIP3 only when complexed with SF-1, attenuating downstream activity (69). Thus, the PIP–SF-1 interaction appears to introduce a regulatory mechanism not previously seen in NRs, in which the phosphorylation status of a bound ligand dictates the activity of its receptor.

PPARs

The peroxisome proliferator-activated receptors (PPARs α , β/δ , and γ) are members of the NR1C subfamily of NRs and play integral roles in the regulation of lipid metabolism and inflammation (70-72). PPARs form heterodimers with the retinoid X receptor (RXR) (73), and recognize an array of ligands, including fatty acids, eicosinoids, and oxidized lipid products (72).

PPAR α and PC 16:0/18:1

PPAR α is expressed in the heart, liver, kidney, muscle, and brown adipose tissue(74). As a fatty acid binding protein, PPAR α regulates the expression of many proteins involved in cellular fatty acid homeostasis (75-77) and systemic lipid balance (78). It has been implicated in atherosclerosis and dyslipidemia, and prolonged activation has been linked to oxidative damage and liver cancer (79). As such, PPAR α is an important pharmacological target. Fibrates, a class of drugs used to treat dyslipidemia, are pharmacological agonists of PPAR α , and exert their therapeutic effects by lowering triglyceride levels (80).

PPAR α is known to bind to many natural free fatty acids (FFAs) and while these are likely physiological ligands, proving that these are *bona fide* endogenous activators is technically

challenging. Like PLs, FFAs are typically insoluble, partitioning into droplets, membranes and soluble lipid binding proteins making direct correlations between binding affinity and activation difficult. It is clear, however, that μM levels of exogenous FFAs (1 – 50 μM) activate PPARs *in vivo* and in animals (81). This is on par with PL-dependent transactivation among NR5A receptors, which display EC50 values ranging from 30 – 100 μM for activating PC and PE isoforms (38; 42). This affinity for FFAs and PLs among nuclear receptors is likely a result of their “generous” lipid binding pockets which allow binding to an array of lipid metabolites.

In 2009, mass spectrometry experiments identified PC 16:0/18:1 as one of several lipids bound to PPAR α isolated from murine liver tissue, and the only one whose presence was dependent on fatty acid synthase (FAS)(81). Binding of this PC species was selective for PPAR α over PPAR δ and PPAR γ , and could be enhanced *in vivo* by FAS induction, and inhibited by treatment with a PPAR α agonist (81). Additionally, PC 16:0/18:1 treatment stimulated PPAR α -dependent gene expression and decreased fatty liver symptoms in mice, lending further credence to its suggested role as an endogenous PPAR α agonist (81).

PPAR γ and tail-oxidized PLs

PPAR γ , which regulates glucose and fatty acid metabolism, is an important target in the treatment of type II diabetes, and is the receptor upon which the thiazolidinedione class of drugs acts (82). In addition to metabolic regulation, PPAR γ is known to be an important player in anti-inflammatory pathways (83). Recently, 15-KETE- and 15-HETE PE, two oxidized PE species, were shown to activate PPAR γ *in vitro*. Reporter gene assays showed a dose dependent activation in HEK293 cells cotransfected with PPAR γ and a PPRE-luciferase construct, and in macrophages harvested from PPRE-EGFP transgenic mice. Furthermore, these oxidized PEs induce the PPAR γ -dependent expression of CD36 in human monocytes (84). Unoxidized PE showed no PPAR γ activation, suggesting that PPAR γ may specifically recognize oxidized PLs. While the formation of oxidized PEs is not dependent on lipases, it remains possible that

phospholipase A (PLA) isoforms may liberate oxidized fatty acids, which are also known PPAR activators. Earlier work showed that oxidized PLs bind directly to the LBP, and PPAR γ protects these oxidized PLs from phospholipase A1 mediated cleavage; however, this same work showed that PLA1 treated oxidized PLs had a similar ability to stimulate PPAR γ transactivation relative to untreated oxidized PLs (85). For PPAR α , however, PLA2 appears to be required for activation by oxidized PLs (86).

USP

Ultraspiracle protein (USP) was identified as the *Drosophila* homolog of mammalian RXR in 1990 (87; 88). Its major function is to serve as a binding partner for the ecdysone receptor (EcR); this heterodimer is a vital regulator of molting and metamorphosis, which is triggered by the binding of 20-hydroxyecdysone (20E) to the EcR subunit (89). However, USP itself can bind to several farnesoid insect juvenile hormones (90), and it is hypothesized to be a ligand-activated NR in its own right (91).

E. Coli PLs

Crystal structures of USP consistently show bacterially-derived PL bound in the LBP(92-95), stabilizing the receptor in an antagonist conformation (93). While most data implicate farnesoid derivatives as the endogenous USP ligand, it is conceivable that insect PLs may play a role in USP-mediated gene regulation, given the emerging role of PLs in other NR pathways. Insects have coopted PLs in the regulation of SREBP processing and nuclear translocation and may have independently evolved PL sensitive NRs. A comparison of the USP-PL crystal structures reveals a nearly identical mode of PL binding versus LRH-1 and SF-1.

PL transport and PL dependent coactivation

PPAR and PC-TP

In addition to direct NR-mediated gene expression, PLs have been shown to indirectly affect gene regulation through lipid shuttling proteins such as phosphatidylcholine transfer protein (PC-TP). PC-TP is a member of the steroidogenic acute regulatory protein (StAR)-related lipid transfer (START) domain superfamily that shares a common fold for lipid binding (96; 97). PC-TP is exquisitely selective for PCs (98), and was originally shown to catalyze both one-for-one PC exchange, and net PC transfer between membranes (99-101). PC-TP has since been identified as an important metabolic regulator, participating in hepatobiliary cholesterol, lipoprotein, glucose and fatty acid metabolism as well as brown fat-mediated thermogenesis (102).

Consistent with PC-TP's participation in metabolic processes, it has been identified as a binding partner for multiple metabolic proteins (103). Arguably, the most interesting of these interactions is with PPAR- α (104). In addition to PPAR- α regulating the expression of PC-TP, PC-TP was shown to up regulate the transcriptional activity of both PPAR- α and HNF-4 α (104). The mechanism of this effect on the transcriptional activity of NRs is not currently understood. Additionally, the context in which NRs bind to PL transporters is also unclear. There is a possibility that in addition to its role in the distribution of lipids in membranes, PC-TP may also deliver PL ligands to PL-sensitive receptors.

Structural Analysis of PL binding proteins

What does it take to bind to PLs as a ligand?

With a large aliphatic surface and significant conformational freedom for the bulk of the molecular structure, PLs certainly do not look like traditional NR ligands (Figure 1.1). Interaction with the hydrophobic tails, while energetically favorable, does not permit specificity by the usual suspects (e.g. H-bonds, salt bridges, cation- π interactions). Below, we discuss the distinction between soluble PL transporters and proteins that utilize the information contained in the PL headgroup to drive intermolecular signaling.

Shuttlers versus transcription factors

Structurally characterized soluble PL transport proteins such as PC-TP and PITP α , fully engulf PLs, interacting substantially with both the lipid tails and the headgroup (Figure 1.2 E-F) (98; 105). Headgroup specificity is generated via H-bonds, ionic interactions and cation- π interactions via residues located at the core of the protein. The lipid tails extend toward the protein surface but remain protected from bulk solvent. This binding mode is in stark contrast to PL-binding NRs, which bury PL tails and present the headgroup at the protein surface (Figure 1.2 A). The average LBP volume in PC-TP and PITP α is 2297 and 3000 \AA^3 , respectively; this is nearly twice as large as the LRH-1, SF-1 and USP LBPs. The molecular volumes of their bound lipids, however, are 874 and 552 \AA^3 , for PCTP and PITP α respectively. It is tempting to speculate the excess cavity volume and “tails out” PL conformation may be due to the requirement that transporters deliver their PL cargo to a target membrane or PL binding receptor prohibiting tight molecular interactions. Consistent with these observations, holo structures of PC-TP and PITP α show that atomic disorder increases distally from the headgroup suggesting less than optimal contacts are made with the PL tails which have vastly more potential energy to contribute to the protein-ligand interaction.

Parallels in the immune system

Both exogenous and endogenous PLs have been implicated as lipid antigens capable of activating natural killer T cells when presented by CD1 proteins localized on human antigen presenting cells (106; 107). CD1 proteins play a critical role in presenting both pathogen derived lipids and glycoproteins to initiate cell-mediated immunity (108). Like NRs, CD1 glycoproteins

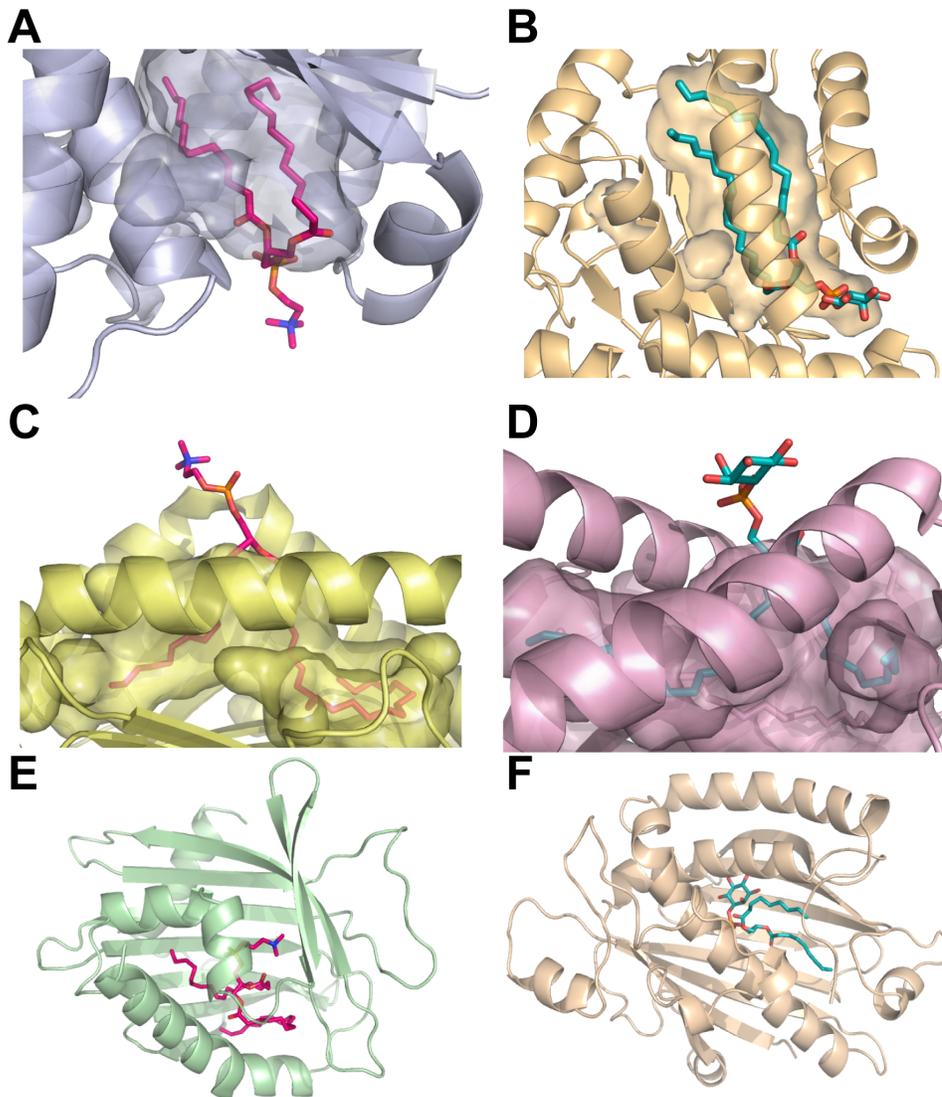


Figure 1.2 Crystal structures of soluble PL signaling proteins.

Proteins are depicted as ribbons with bound phospholipids represented as sticks (O, red; P, magenta; N, blue). Molecular surfaces are shown to highlight the ligand binding pockets. A. LRH-1 (slate) bound to DLPC (magenta) (41), B. SFH-1 (tan) bound to PI (cyan) (109) C. CD-1 (yellow) bound to PC (magenta) (106) D. CD-1 (pink) bound to PI (cyan) (107) showing the bound ligands with lipid head-groups exposed to solvent. In contrast, the lipid shuttling proteins e. PC-TP (light green) bound to PC (magenta) (98) and F. PITP (almond) bound to PI (cyan) (110) completely engulf their lipid ligands.

bind PLs in a “tails-first” orientation with the PL headgroup exposed to the protein surface. The binding and presentation of both PC and PI by CD1b and CD1d, respectively, is remarkably similar to the presentation of PLs by NRs (Figure 1.2 A and 1.2 C-D), whereby the lipid tails are buried and the headgroup is exposed to solvent. Thus, PL headgroup presentation may be a hallmark of PL dependent signaling.

Comparison to the PL PI/PC transporter Sec14

Sec14, originally defined by its ability to promote the movement of PC and PI between membranes, is now known as an integrator of PL signaling at the membrane (111). To accomplish this, Sec14 senses both PC and PI levels to stimulate PI4-K mediated PI phosphorylation – a process critical for vesicle biogenesis. Interestingly, Sec14 requires both PC binding and PI binding for activity (109), however, a PC/PI exchange model has been proposed whereby PC binding facilitates PI loading. While a direct interaction between Sec14 and PI4-K has not been observed, presentation of PI for decoration requires that the inositol moiety is accessible to protein surface (Figure 1.2 B). Indeed, while Sec14 completely buries the PC headgroup, the inositol ring of PI requires only the movement of few side chains to access the solvent. These observations parallel what we know for LRH-1/SF-1; they both are capable of binding PC and PI and presentation of the phosphorylated inositol headgroup is required for signaling (SF-1). Furthermore, since DLPC binding has not yet been tested *in vivo*, it is possible that the PC binding ability of LRH-1 and SF-1 may facilitate the loading of PI in a similar exchange reaction.

PL presentation as a model for PL dependent signaling.

Unlike widely prevalent PL binding domains such as PHD fingers that recognize PLs in the context of a membrane (112), NRs engulf PLs “tails first” making extensive hydrophobic contact with more than 15 residues and up to three hydrogen bonds near the surface of the receptor (113). It is clear that most of the binding energy is derived from interaction with the aliphatic tails,

which in all known structures, intertwine to fill large 1300-1750 Å³ binding pocket that starts at the core of the protein and terminates at the protein surface. Lipid tails occupy the very core of the receptor greatly enhancing protein stability (41). In this way, PLs act as folding nuclei much like the hormones in other NR family members (114). However, the vast diversity among PLs and the potential for lipid modifications suggests that PL dependent transcription factors may serve to integrate varying and complex signals to tune gene expression. This represents an added layer of complexity on the already complicated cistrome in which coregulators, DNA, chromatin modifying enzymes and accessory proteins orchestrate coordinated gene expression.

Closing Remarks

Evolution has generated a highly complex system to control energy homeostasis, including allosteric mechanisms within key metabolic enzymes, and the nutritional control of gene expression via transcription factors. Lipids are a major source of energy for the cell, and it is well known that the composition and availability of these lipids plays a central role in regulating glycolysis. Direct PL sensing by nuclear hormone receptors tie PL levels not only into glucose and lipid homeostasis but to steroid synthesis, reproduction, inflammation, development and cell differentiation (Figure 1.3).

Given the molecular properties of PLs, it is no surprise that PL-driven transcription factors have been largely recalcitrant to drug design. Proteins with large hydrophobic pockets typically require large ligands and the potential for specific interactions within core of the LBP are slim. While there have been a few successes in designing specific compounds targeting these receptors, improving these compounds and predicting their binding modes remain challenging. Clearly, modulating PL-driven transcriptional pathways remains an untapped therapeutic opportunity and advances in this area of research are desperately needed.

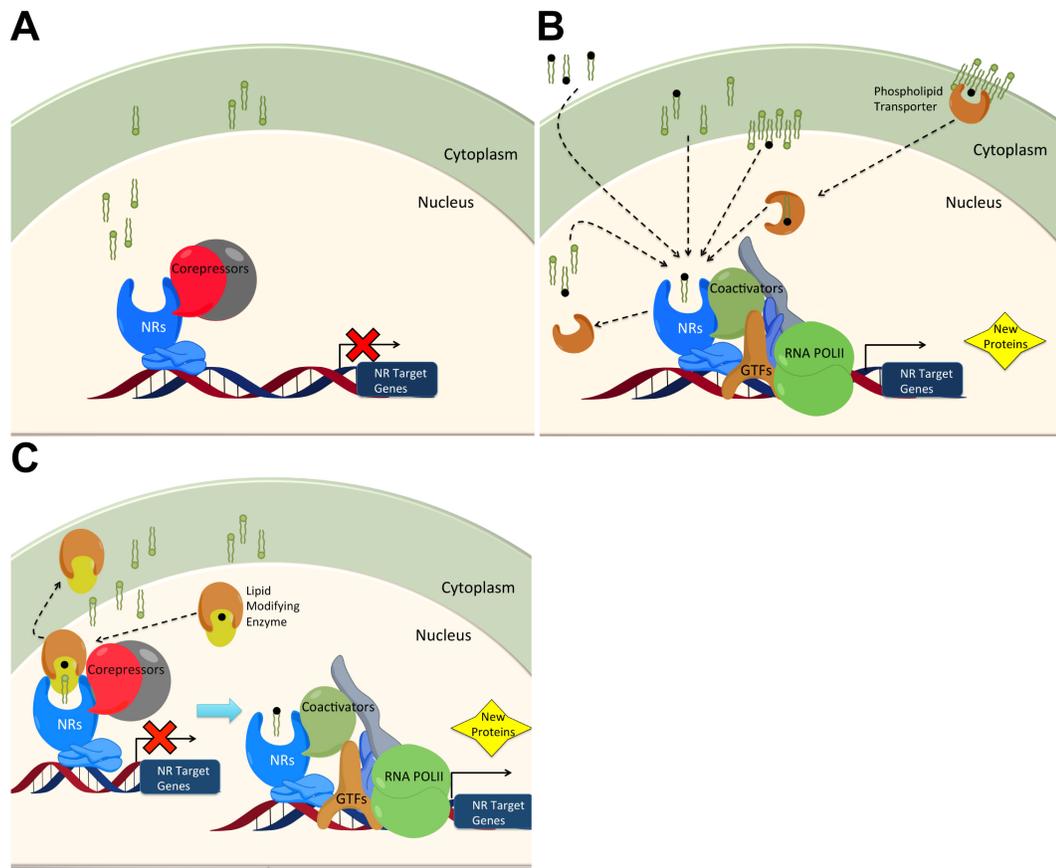


Figure 1.3 Phospholipid mediated transcription control.

A. In the absence of a PL agonist NRs are bound to corepressor proteins and block transcription.

B. Activating PLs from exogenous, membrane bound or cytoplasmic sources bind to NRs or are potentially delivered by PL transporter proteins. Once an activating PL is bound to the NR coactivator complexes along with other general transcription factors (GTFs) and RNA polymerase initiate the transcription of genes.

C. NRs can also be bound to non-activating lipids with lipid modifying enzymes altering the lipid in place to become an activating lipid.

LRH-1 Knowledge Gaps

While it is clear that LRH-1 is an important player in a number of neoplastic and metabolic disease states, an understanding of how ligands can alter the receptor's activation state is lacking. Previous work to structurally characterize the LRH-1 LBD has involved: human LRH-1 bound to *E. coli* PLs, and mouse LRH-1 in the absence of a bound PL. Prior to 2010, the receptor had been solved bound to a full length non canonical corepressor DAX (115), or peptides derived from either the non canonical corepressor SHP (40; 116; 117) or coactivator TIF-2 (118) (Fig. 1.4). These studies identified only that PL binding ability is important for receptor activation, but did not identify how a bound PL might act to promote activation. Additionally, work comparing mouse and human LRH-1 hypothesized that rodent LRH-1 may have lost its ability to bind to and be activated by PLs (40). The experiments henceforth focus on filling 3 major LRH-1 knowledge gaps: 1) identifying LRH-1's ligand/coregulator states, 2) identifying the structural mechanisms that may promote PL activation of the receptor, and 3) understanding how sequence variation throughout evolution has altered LRH-1's ability to respond to PL ligands.

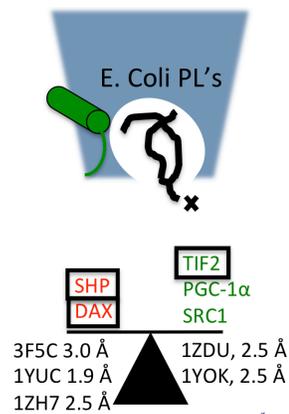


Figure 1.4 Published LRH-1 structures before 2010

Representation of the LRH-1 LBD shown in blue with the AF-H depicted as a green tube in the active orientation. The bound *E. coli* PL is drawn in black showing the intertwined acyl tails and an x denotes an unknown headgroup. Corepressors are printed in red and coactivators in green. Boxes denote a coregulator structure that has been solved. Published coregulator peptide complex structures are denoted by PDB ID and resolution.

Chapter 2: Antidiabetic phospholipid-nuclear receptor complex reveals the mechanism for phospholipid-driven gene regulation.²

²This chapter has been slightly modified from the published manuscript: Musille PM, Pathak MC, Lauer JL, Hudson WH, Griffin PR, Ortlund EA. (2012) Antidiabetic phospholipid-nuclear receptor complex reveals the mechanism for phospholipid-driven gene regulation. *Nat Struct Mol Biol.* 19(5):532-7, S1-2.

Introduction

The regulation of lipid and glucose homeostasis is of central importance to human physiology. Governing this process are both extra- and intracellular receptors that sense hormones and nutrients, namely fatty acids and glucose, to control behavior and nutrient homeostasis. Hepatic lipid metabolism and glucose regulation are intimately related and lipid accumulation can lead to metabolic diseases such as steatosis and diabetes. Recently, the dietary PL, DLPC (PC 12:0-12:0), was shown to lower serum lipid levels and reduce blood glucose levels in diabetic mice (42). This effect is completely dependent on LRH-1, a PL binding orphan nuclear receptor (NR) (42). This medium chain PC selectively activates LRH-1 mediated transcription in luciferase assays, increases the ability of LRH-1 to interact with coactivators, and increases the production of LRH-1 target genes. It has been shown using mass spectrometry that DLPC is able to competitively displace larger PLs from the ligand binding domain of LRH-1 *in vitro* while longer chain PLs such as DPPC (C16:0-C16:0) can not (42). Thus, while a range of PLs bind to LRH-1 (119-121), acyl chain length and head group composition dictate transcriptional activation (42). The mechanism governing this selective activation, however, is unclear since the binding mode of DLPC is unknown.

LRH-1 is a member of the NR5A class of NRs regulating the expression of genes central to embryonic development, reproduction, lipid homeostasis, and energy metabolism (51; 122; 123). LRH-1 is required to maintain Oct4 expression in undifferentiated embryonic stem cells to maintain pluripotency (124; 125). In breast cancer, LRH-1 regulates both estrogen synthesis and estrogen receptor (ER) expression (126-131). This, along with direct transcriptional regulation of LRH-1 expression by ER, makes LRH-1 a key element in the feed-forward loop driving sustained estrogen biosynthesis and signalling in ER⁺ breast cancer (126-131). In hepatic tissues, LRH-1 regulates genes central to lipid and bile acid homeostasis (123).

LRH-1, like most NRs, interacts with coactivators via a LXXLL motif (where X is any amino acid) at the interaction surface formed by an “active” AF-H packed against helices 3 and 4

(12). However, LRH-1 also uses this same “active” surface to interact with the atypical NRs SHP and Dax-1, which generally act to repress LRH-1 in hepatic tissues (132; 133). This, combined with the fact that all NR5A receptors have crystallized with the AF-H in the active orientation, has led to the belief that LRH-1 is not optimized to interact with typical corepressors such as SMRT and NCoR and the position of the AF-H is not altered by ligand binding (134; 135). Indeed, recombinant LRH-1 loaded with co-purified *E. coli* lipids is incapable of binding to SMRT, despite being specifically repressed by SMRT in a dose dependent manner *in vivo* (136).

To investigate the general mechanism of activation and PL binding capacity among LRH-1 family members, we generated structures of apo LRH-1 and LRH-1 in complex with the activating ligand DLPC. We use hydrogen-deuterium exchange, thermal stability, and functional assays to show that the lipid-free receptor undergoes structural fluctuations allowing it to interact with widely expressed corepressors. We show that binding of the antidiabetic ligand, DLPC, is a dynamic process that alters coregulator selectivity and we identify a previously unknown activation function surface integral for both ligand binding and receptor activation.

Experimental Procedures

Reagents:

Chemicals were purchased from Sigma, Fisher or Avanti Phospholipids. pMALCH10T and the vector for His tagged TEV was a gift from John Tesmer (UT Austin). pLIC_MBP and pLIC_HIS were gifts from John Sondek (UNC, Chapel Hill). Peptides were synthesized by RS Synthesis (Louisville, KY).

Protein Expression and Purification:

The human LRH-1 LBD (residues 291–541) was purified as described previously (119). For the LRH-1–DLPC complex, purified LRH-1 LBD was incubated with DLPC vesicles, prepared by sonication to optical clarity, at a 1:20 (protein:lipid) molar ratio for 24 hours at 22 °C. Receptor was purified away from unbound lipids by size exclusion chromatography, dialyzed against 60 mM NaCl, 100 mM ammonium acetate, pH7.4, 1 mM DTT, 1 mM EDTA, and 2 mM CHAPS

and concentrated to 5–7 mg ml⁻¹. For apo LRH-1 crystallization, purified LRH-1 LBD (residues 291–541) was incubated with 1,2-ditetracosanoyl-sn-glycero-3-phosphocholine (PC 24:0-24:0) (Avanti Polar Lipids) and (RJW101) at a final PC24:ligand:protein ratio of 20:3:1 (137). The receptor was purified away from unbound PC 24:0–24:0 and the weakly bound agonist by size exclusion chromatography, dialyzed against 60 mM NaCl, 100 mM ammonium acetate, pH7.4, 1 mM DTT, 1 mM EDTA, and 2 mM CHAPS and concentrated to 5-7 mg ml⁻¹. For apo LRH-1 crystallization, purified LRH-1 LBD was incubated with 1,2-ditetracosanoyl-sn-glycero-3-phosphocholine (PC 24:0-24:0) (Avanti Polar Lipids) and (RJW101) at a final PC24:ligand:protein ratio of 20:3:1 (137). The receptor was purified away from unbound PC 24:0–24:0 and the weakly bound agonist by size exclusion chromatography, dialyzed against 60 mM NaCl, 100 mM ammonium acetate, pH 7.4, 1 mM DTT, 1 mM EDTA, and 2 mM CHAPS and concentrated to 5-7 mg ml⁻¹.

Structure Determination:

Crystals of the LRH-1 LBD–DLPC–TIF2 complex were grown by hanging drop vapor diffusion at 20 °C from solutions containing 3 µL of protein at 6.5 mg ml⁻¹ protein complexed with a peptide derived from human TIF2 NR Box 3 (⁺H₃N- KENALLRYLLDKDD-CO₂⁻) at a 1:4 molar ratio and 1 µL of the following crystallant: 18%–24% PEG 400, 5% glycerol, 0.1 M lithium sulfate, and 0.1 M sodium acetate pH 5.2. Crystals were cryoprotected in crystallant containing 12% glycerol and 12% ethylene glycol and flash-frozen in liquid N₂. Crystals of apo LRH-1 LBD–SHP complex were grown by hanging drop vapor diffusion at 20 °C from solutions containing 1 µL of protein at 4 mg ml⁻¹ protein complexed with a peptide derived from human SHP NR box 1 (⁺H₃N-QGAASRPAILYALLSSSLK-CO₂⁻) at a 1:4 molar ratio and 1 µL of the following crystallant: 9.5%-15% PEG 3350, 5% glycerol, and 50 mM Bis-Tris, pH 6.4. Crystals were cryoprotected in crystallant containing 20% glycerol and flash-frozen in liquid N₂. Data to 1.8 Å (DLPC) and 1.9 Å (apo) resolution were collected at 100 K at the Southeast Regional Collaborative Access Team (SER-CAT) at the Advanced Photon Source (Argonne, IL), and were

processed and scaled with HKL2000 (Table 1) (138). Initial phases were determined using LRH-1 PDB coordinates, either 1YOK (120) or 1YUC (119), as a molecular replacement search model. The structure was refined using REFMAC5 within the CCP4 suite of programs (139; 140) and model building was performed in COOT (141; 142). The final model for the LRH-1–TIF2–DLPC complex contains LRH-1 residues 296–538, TIF2 residues 742–751, and displays good geometry, with 99.6%% and 0.4% of the residues in the favored and allowed regions of the Ramachandran plot, respectively, as analyzed by MolProbity (143). The final model for the apo LRH-1–SHP complex contains LRH-1 residues 300–397 and 421–538, SHP residues 16–27, and displays good geometry, with 98.5% and 1.5% of the residues in the favored and allowed regions of the Ramachandran plot, respectively, as analyzed by MolProbity (143). Data collection and refinement statistics can be found in Table 1. Coordinates and structure factors have been deposited with the Protein Data Bank under accession codes 4DOR and 4DOS.

Mass Spectrometry:

Samples were analyzed using electrospray mass injection-MS in the negative-ion mode to detect and identify phospholipids. Approximately 6 mg of wild-type or mutant forms of LRH-1 LBD and Ftz-F1 LBD were extracted with a 2:1 chloroform/methanol solution, diluted in 200 μ L of chloromethylene, and analyzed by negative ion ESI-MS on a Thermo LTQ FTMS using direct injection analysis with electrospray ionization (Thermo Finnigan, Somerset, NJ). All extractions were performed in duplicate. The major phospholipid species were identified by accurate mass measurements and MS-MS via collisionally induced dissociation (CID), which yields product ions characteristic of the head groups and attached fatty acids. Acquisition and analyses were performed using the instrument's Analyst QS software.

Phospholipid Quantification:

Six mgs of protein was subjected to chloroform:methanol extraction according to the two-step Bligh and Dyer method (144) to isolate phospholipid as described previously (119). Phospholipid quantification was performed according to the improved procedures for the determination of lipid

phosphorous by malachite green (145). Briefly, lipid extracts were dried completely and digested with perchloric acid (70%) at 140 °C until all color disappeared. To the cooled tubes, a solution of malachite green and ammonium molybdate was added and vortexed for 20 min. The interaction between phosphomoylbdenum and malachite green was monitored at 660 nm.

Generation of apo LRH-1:

Pure protein was subjected to chloroform:methanol extraction (2:1) to remove bound lipids according to the Bligh & Dyer method (144). The resulting pellet containing denatured protein was washed three times with chloroform to remove any trace lipids associated with the protein or vessel. The resulting white pellet was then dried by evaporation and resuspended in 6 M Guanidinium HCl. Empty protein was then refolded by fast dilution into a buffer containing 10 mM K₂HPO₄, 100 mM Tris (pH 7.4), 1 mM EDTA, and 500 uM CTAB at 4 °C. After ~ 20 hours, protein was then concentrated and purified by size exclusion chromatography to ensure a homogenous population of refolded receptors.

Transient Transfection Assay:

HeLa cells were cultured in DMEM supplemented with 10% (v v⁻¹) fetal bovine serum (HyClone). The cells were seeded overnight into 96-well cell-culture plates. The cells were then transfected with a Lipofectamine (Invitrogen)/DNA mixture containing 100 ng of SHP promoter firefly luciferase reporter construct, 25 ng of pCI_LRH-1 receptor (WT or mutant), and 1 ng renilla luciferase. Firefly luciferase activity was assayed using a BioTek Synergy 4 spectrophotometer and normalized to renilla luciferase activity 24 h after transfection.

Cofactor binding assays:

The polarization of fluorescein labeled peptides derived from SHP NR box 1 (³H₃N-QGAASRPAILYALLSSSLK-CO₂⁻), PGC-1α NR Box 2 (³H₃N-EEPSLLKLLLAPA-CO₂⁻), SRC-1 NR Box 2 (³H₃N-SPSSHSSLTERHKILHRLQLQEGSP-CO₂⁻), SMRT (³H₃N-TNMGLEAIIRKALMGKYDQW-CO₂⁻), TIF2 NR Box 3 (³H₃N-

PVSPKKKENALLRYLLDKDDT-CO₂) was monitored using a BioTek Synergy 4 spectrophotometer with polarizers (Winooski, VT) as a function of protein concentration. Experiments were conducted in 150 mM sodium chloride, 20 mM Tris-HCl (pH 7.4), and 5% (v v⁻¹) glycerol unless stated otherwise. All experiments were performed in triplicate and data were fit with Prism 5 (GraphPad, Inc.) by linear least-squares methods to a single site-binding model.

Circular Dichroism Spectroscopy and Thermal Unfolding Studies.

Circular dichroism (CD) studies were performed on a Jasco J-800 spectropolarimeter with a 1-mm cell. Proteins were dissolved at a concentration of 0.2 mg ml⁻¹ in 20 mM tris (pH 7.4), 0.1 M sodium chloride, and 10% glycerol. Wavelength scans were performed at 25 °C from 200 to 250 nm at a rate of 50 nm min⁻¹. For thermal unfolding studies, ellipticity was continuously monitored at 220 nm while the temperature was raised by use of a Jasco PFD-425S temperature control unit from 25 to 80 °C at a rate of 1 °C min⁻¹. The α -helix/ β -sheet ratio was calculated using the k2d3 server <http://www.ogic.ca/projects/k2d3> (146).

Proteolytic Protection Assay:

DLPC or apo LRH-1 (11.25 μ g) was digested with 80 ng of chymotrypsin (Protea Biosciences, Inc.) for 5 min at room temperature. The reaction was quenched with the addition of acetic acid and boiled for 5 minutes. The entire reaction was resolved by SDS-PAGE and stained by Coomassie blue. Gel regions below undigested intact protein were excised and were subjected to in-gel trypsin digestion. The digested peptides were analyzed by reverse-phase liquid chromatography coupled with tandem mass spectrometry (LC-MS/MS) as previously described (147). Briefly, peptide mixtures were loaded onto a C₁₈ column (75 μ m i.d., 30 cm long, 3 μ m resin from Michrom Bioresources, Inc., Auburn, CA) and eluted over a 12-35% gradient (Buffer A: 0.1% Formic Acid, 0.005% heptafluorobutyric acid, and 5% Acetonitrile; Buffer B: 0.1% formic acid, 0.005% heptafluorobutyric acid, and 95% Acetonitrile). Eluates were monitored in a MS survey scan followed by ten data-dependent MS/MS scans on an LTQ-Orbitrap ion trap mass spectrometer (Thermo Finnigan, San Jose, CA). The LTQ was used to acquire MS/MS spectra (2

m/z isolation width, 35% collision energy, 5,000 AGC target, 300 ms maximum ion time). The Orbitrap was used to collect MS scans (300-1600 m/z^{-1} , 1,000,000 AGC target, 750 ms maximum ion time, resolution 60,000). The acquired MS/MS spectra were searched against a concatenated target-decoy *E. coli* database (UNIPROT January 23, 2011) that included the LRH-1 sequence using the SEQUEST Sorcerer algorithm (version 2.0, SAGE-N) (148). Searching parameters included: partial tryptic restriction, parent ion mass tolerance (± 10 ppm), product ion tolerance ($\pm 0.5 m/z^{-1}$), and dynamic modifications for oxidized Met (+15.9949 Da). The peptides were classified by charge state and trypticity (fully and partial) and filtered dynamically by increasing XCorr and ΔC_n values to reduce the protein false discovery rate to less than 5%. The MS/MS spectra of matched LRH-1 peptides were manually inspected. Trypsin digest sites were removed manually and spectral counts per peptide were used to determine the relative amount of each chymotrypsin proteolysis fragment between DLPC and apo RLH-1. The protection factor reported in Fig. 2.5b and Table 2.3 is the result of subtracting spectral counts for chymotryptic proteolysis fragments observed in the LRH-1-DLPC complex from the same chymotryptic proteolysis fragments generated from the apo protein. A higher protection factor indicates less chymotrypsin cleavage events upon DLPC binding.

HDX:

Solution-phase amide HDX was performed with a fully automated system as described previously (149). Briefly, 4 μL of protein was diluted to 20 μL with D_2O -containing HDX buffer, and incubated at 25 $^\circ\text{C}$ for; 10 s, 30 s, 60 s, 900 s, or 3600 s. Following on-exchange, unwanted forward or back exchange was minimized and the protein is denatured by dilution to 50 μL with 0.1% TFA in 5 M urea (held at 1 $^\circ\text{C}$). Samples were then passed across an immobilized pepsin column (prepared in house) at 50 $\mu\text{L min}^{-1}$ (0.1% TFA, 15 $^\circ\text{C}$) and the resulting peptides were trapped on a C8 trap cartridge (Thermo Fisher, Hypersil Gold). Peptides were then gradient eluted (4% CH_3CN to 40% CH_3CN , 0.3% formic acid over 5 minutes, 2 $^\circ\text{C}$) across a 1 mm \times 50 mm C18 HPLC column (Hypersil Gold, Thermo Fisher) and electrosprayed directly into an Orbitrap

mass spectrometer (LTQ Orbitrap with ETD, Thermo Fisher). Data was processed with in-house software and visualized with PyMOL (Schrödinger, LLC.). To measure the difference in exchange rates, we calculated the average percentage deuterium uptake for native LRH-1 LBD following 10, 30, 60, 900 and 3600 seconds of on exchange. From this value, we subtract the average percent deuterium uptake measured for the DLPC bound LRH-1 LBD. Positive perturbation values indicate exchange rates are faster for these regions within LRH-1–DLPC.

Results

Structure of the activated LRH-1–DLPC–TIF2 NR box 3 complex

To visualize the molecular mechanism driving DLPC biology we determined the structure of the LRH-1 LBD in complex with DLPC and a fragment of the human coactivator Transcriptional Intermediary Factor-2 (TIF2) to a resolution of 1.8 Å (Fig. 2.1a; Table 2.1). Electron density within the ligand binding pocket (LBP) showed clear evidence for bound DLPC (Fig. 2.1b). Electron density was considerably weaker for the distal portion of the lipid tails, indicating that the terminal 2 atoms on the *sn-1* acyl tail and 6 atoms on the *sn-2* acyl tail are mobile. This is in stark contrast to the larger co-purified *E. coli* PLs reported previously (C16:1-C18:1) which show fully ordered lipid tails that intertwine to fully occupy the LBP (Fig. 2.1c) (39; 40; 118).

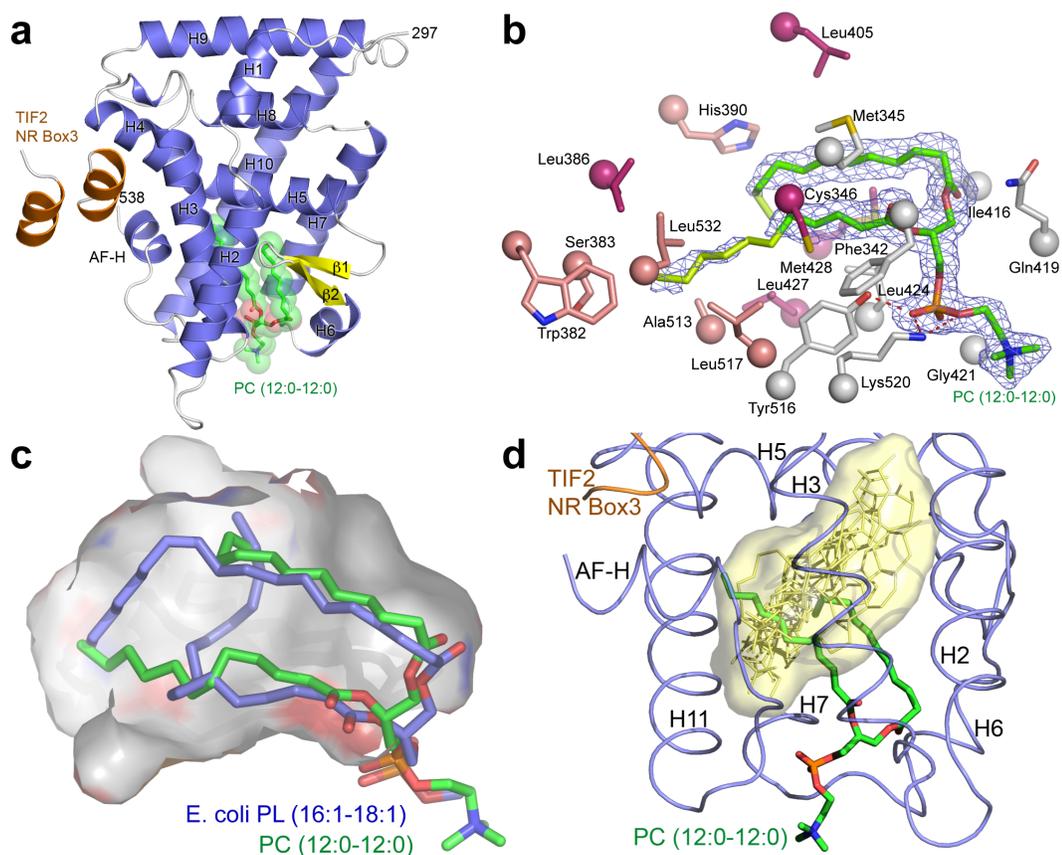


Figure 2.1 DLPC binds directly to LRH-1 and promotes activation through unique interactions.

(a) Ribbon diagram of DLPC bound LRH-LBD (α -helices, blue; β -strands, yellow) with the TIF2 (TIF2) NR box 3 peptide (orange). The bound DLPC is depicted as sticks (C, green; O, red; P, magenta; N, blue) surrounded by transparent spheres. (b) F_o-F_c omit electron density (contoured at 1σ) for the bound DLPC (12:0-12:0). Well-ordered lipid atoms are colored in green while poorly ordered atoms are colored in lime. Side chains lining the LBP of LRH-1 that contact DLPC atoms with strong electron density are colored grey and side chain contacting DLPC atoms with weak electron density are colored salmon. Side chains that contact bacterial PL (PDB 1YUC)(119) but not DLPC are shown in pink. (c) Superposition of the bacterial C16:1-C18:1 phospholipid from 1YUC (blue) on DLPC with the surface of the LBP outlined in grey. (d)

Superposition of LRH-1–DLPC on representative NR family members (RXR, VDR, ER α , PPAR α , FXR, and TR) with their endogenous ligands shown as yellow sticks and their combined overall molecular surface highlighted in yellow. Figures were generated in PyMOL.

	LRH-1 – DLPC – TIF2 NRbox3	LRH-1 – SHP NRbox2
Data collection		
Space group	P2 ₁ 2 ₁ 2 ₁	P2 ₁
Cell dimensions		
<i>a, b, c</i> (Å)	49.5, 71.3, 99.7	63.5, 59.4, 74.2
<i>a, b, c</i> (°)	90.0, 90.0, 90.0	90.0, 90.0, 100.1
Resolution (Å)	1.8 (1.86 – 1.80)*	1.9 (1.97 – 1.90)*
<i>R</i> _{sym} or <i>R</i> _{merge}	9.3 (30.6)	10.2 (49.7)
<i>I</i> / <i>σI</i>	13.0 (3.6)	19.3 (3.5)
Completeness (%)	93.9 (78.3)	99.9 (99.8)
Redundancy	3.9 (3.3)	4.1 (3.7)
Refinement		
Resolution (Å)	1.8	1.9
No. reflections	31603	83181
<i>R</i> _{work} / <i>R</i> _{free}	17.6 / 21.9	17.3 / 21.9
No. atoms		
Protein	2518	3886
Ligand/ion	45	47
Water	222	165
<i>B</i> -factors		
Protein	23.5	25.6
Ligand/ion	31.5	38.4
Water	30.7	28.9
R.m.s. deviations		
Bond lengths (Å)	0.006	0.014
Bond angles (°)	1.2	1.2
PDB	4DOS	4DOR

Table 2.1 Data collection and refinement statistics

* Data collected from a single crystal *Values in parentheses are for highest-resolution shell.

We hypothesized, that DLPC, with 10 fewer acyl chain carbons, would insert itself deeper in the LBP. Instead, the phosphoglycerol backbone of DLPC binds to LRH-1's $\sim 1300 \text{ \AA}^3$ LBP in a similar position to the phosphoglycerol backbone of *E. coli* lipids in previously reported structures (64; 119; 120). The $\sim 240 \text{ \AA}^3$ difference in molecular volume between the LBP occupying atoms of DLPC and *E. coli* PL translates directly to additional unoccupied space in the deepest regions of the LBP, increasing the unoccupied pocket volume to $\sim 870 \text{ \AA}^3$. Earlier LRH-1-*E. coli* phospholipid structures were unable to identify which portions of the LBP are important for coordinating receptor activation (64; 119; 120). The binding mode of DLPC is radically different from the endogenous ligands of other classes of NRs as the space used to coordinate ligands in nearly all other family members is left almost completely unfilled (Fig. 2.1d). Instead, the most ordered regions of the bound DLPC are near the "mouth" of the LBP (defined by loops between helices 2-3, 6-7 and 10-AF-H). Comparing differential lipid-LRH-1 residue contacts between the atoms of DLPC and bacterial PL in PDB 1YUC (119), reveals that DLPC maintains contacts at the mouth of the pocket and loses contacts with residues on H5, H7 and the AF-H (Fig. 2.2 and Fig. 2.1b). Taken together, these results suggest that a unique set of protein-ligand interactions outside of the canonical LBP and a lack of interactions in the deepest regions of the pocket governs LRH-1 transcriptional activation by DLPC.

Phospholipid tail composition drives differential receptor dynamics

The impact of these differential contacts on the conformational dynamics of LRH-1 is difficult to predict via the crystal structure alone; therefore, we used thermal unfolding and hydrogen deuterium exchange coupled to mass spectrometry (HDX) to assess DLPC's effect on LRH-1's LBD in solution. Consistent with the observation that DLPC occupied less space and contacted fewer amino acids than *E. coli* PLs, DLPC binding decreased LRH-1's ability to resist thermal denaturation (Fig. 2.3a). To identify which regions of the protein were being specifically destabilized by DLPC, we used high resolution HDX to compare the *E. coli* PL

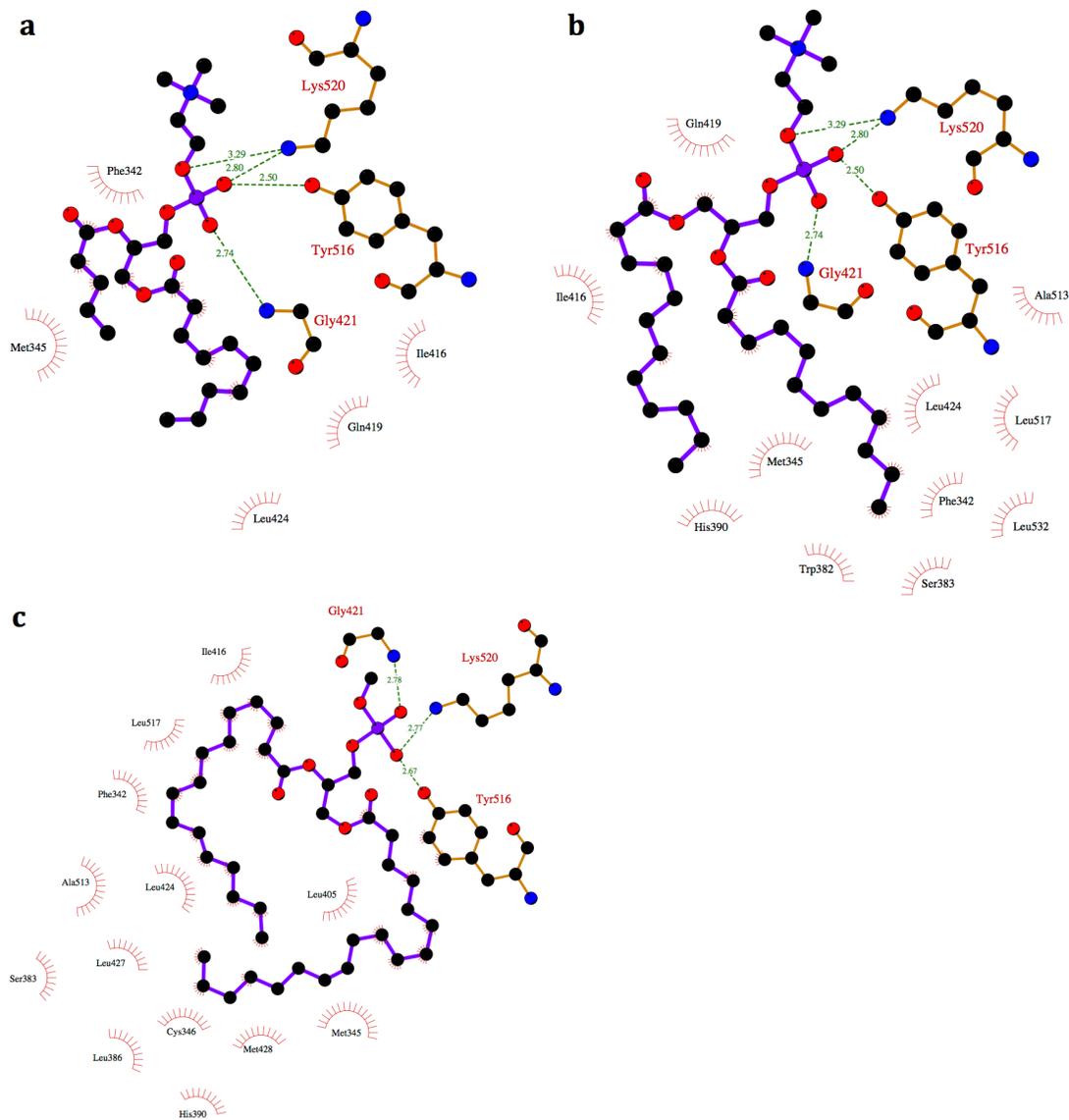


Figure 2.2 Ligplot representation of bound PLs.

Ligplot representation of contacts between (A) LRH-1-DLPC showing only well ordered ligand atoms, (B) LRH-1-DLPC shown all ligand atoms, and (C) LRH-1-*E. coli* PL (PDB code 1YUC) (116). The LRH-1-DLPC complex shows far fewer contacts between the acyl tails and the back for the LRH-1 ligand binding pocket vs. the LRH-1-*E. coli* PL complex. Figures were generated in Ligplot+ and LigEd (150).

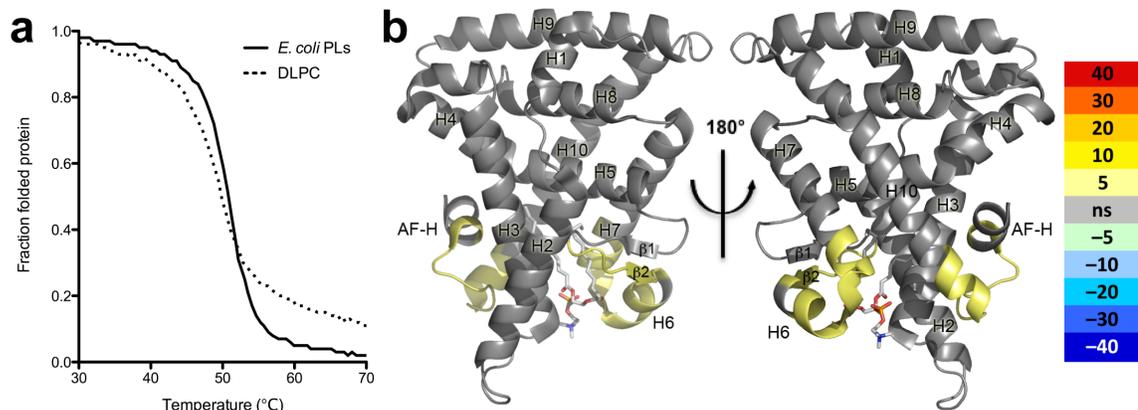


Figure 2.3 DLPC alters LRH-1 stability and structural dynamics.

(a) Thermal melting monitored by circular dichroism showing native recombinant LRH-1 LBD (solid line) $T_m = 50.90 \pm 0.021$ °C and DLPC bound LRH-1 (dashed line) $T_m = 49.12 \pm 0.0384$. Native and DLPC bound LRH-1 LBD are at identical concentrations. (b) Differential HDX between native LRH-1 and DLPC bound LRH-1 LBD mapped onto PDB ID code 1YUC (119). Percent deuterium incorporation difference is indicated by color scale bar. Figures were generated in PyMOL.

bound receptor and the LRH-1 DLPC complex. This solution based structure probing revealed that the β -sheet–helix 6 region and surprisingly helix 10 and the AF-H were more dynamic in the DLPC complex suggesting that differences in PL acyl tail length may affect the ability of LRH-1 to interact with co-regulators (Fig. 2.3b).

Generating apo LRH-1

To date LRH-1 has only been characterized in complex with contaminating *E. coli* PLs (119; 120). We therefore used organic solvents to denature and strip recombinantly expressed LRH-1 LBD of bound *E. coli* PLs followed by refolding to generate apo receptor. Lipid phosphorous assays confirmed that the receptor contains only trace amounts of PL (Fig. 2.4a, and Fig. 2.5a) (119) while circular dichroism detected similar secondary structure composition albeit with less overall secondary structure than native LRH-1 (Fig. 2.4b).

Ligand Binding Alters Co-regulator Preference

To assess the activity of apo LRH-1, we monitored the recruitment of co-activator peptides derived from SRC-1, TIF2, PGC-1 α , and a co-repressor peptide derived from SHP. Apo LRH-1 showed low μ M binding affinity for SRC-1, TIF2, and SHP and no detectable binding to the peptide derived from PGC-1 α . Upon binding various (C14–C20) *E. coli* lipid isoforms, the affinity for SRC-1 was not affected, the affinity for TIF2 increased 10-fold, and binding to SHP decreased 12-fold. Finally we tested the effect of DLCP binding on co-regulator peptide selectivity. DLPC enhanced the ability of LRH-1 to recruit SRC-1 and TIF2 approximately 4-fold, had no effect on PGC-1 α recruitment and abolished the interaction with SHP (Fig. 2.4c and Table 2.2). Our data is consistent with the observation that medium chain PLs are activating while long chain lipids are not (42; 64).

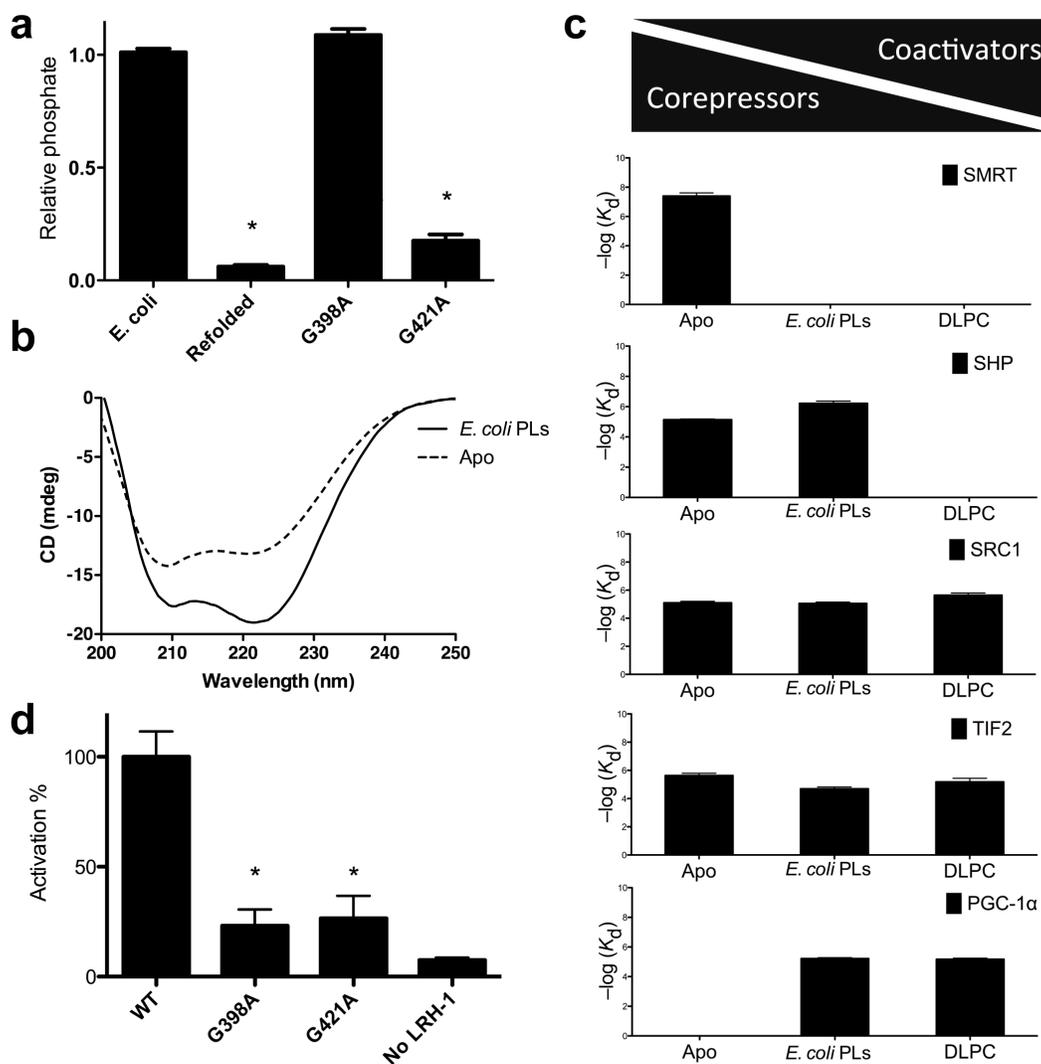


Figure 2.4 DLPC binding affects LRH-1 dynamics and co-regulator preference.

(a) Total PL quantification of LRH-1 variants following chloroform:methanol extraction. Each bar represents 3 experiments repeated in triplicate. (b) Circular dichroism chromatogram showing native LRH-1 (solid line) and refolded apo LRH-1 (dashed line). The percent α -helical/ β -sheet content for the native vs. refolded protein is 56:7 and 51:9 respectively. (c) Activity of the wild-type, G398A, and G421A LRH-1 variants in HeLa cells using a SHP luc reporter. Data are represented as mean \pm s.e.m.; * denotes $P < 0.001$ from 6 replicates in 2 independent experiments. (d) K_d s from Table 2.2 plotted in bar graph form as the inverse log of the $K_d \pm$ s.e.m.; * denotes $P < 0.001$ resulting from 3 experiments performed in triplicate. Apo

LRH-1 was chloroform:methanol extracted and refolded as detailed in the Experimental Procedures. *E. coli* PLs (x-axis) are the range of PLs bound to LRH-1 when purified from *E. coli* and DLPC (x-axis) represents LRH-1 that has been loaded with DLPC and re-purified as detailed in the Experimental Procedures. The wedges at the top represent LRH-1's preference for coactivators (PGC-1 α , SRC1, TIF2) vs. corepressors (SHP, SMRT) in varying liganded states.

Ligand	PGC-1α		SRC1		TIF2	
	K_d (μ M)	R ²	K_d (μ M)	R ²	K_d (μ M)	R ²
LRH-1 <i>E. coli</i> PLs	6.1 +/- 0.4	0.99	8.8 +/- 0.8	0.99	20.1 +/- 2.9	0.98
Refolded/Apo	n.b.	~	8.0 +/- 0.9	0.97	2.3 +/- 0.4	0.95
LRH-1 DLPC	6.8 +/- 0.5	0.99	2.3 +/- 0.4	0.96	6.5 +/- 2.0	0.91

Ligand	SHP		SMRT	
	K_d (μ M)	R ²	K_d (μ M)	R ²
LRH-1 <i>E. coli</i> PLs	0.6 +/- 0.1	0.99	n.b.	~
Refolded/Apo	7.36 +/- 0.4	0.99	0.050 +/- .003	0.98
LRH-1 DLPC	n.b.	~	n.b.	~

n.b. - no binding detected

Table 2.2 Coactivator Peptide Recruitment

Affinity of LRH-1 binding to various coactivator and corepressor peptides in various liganded states as determined by fluorescence polarization

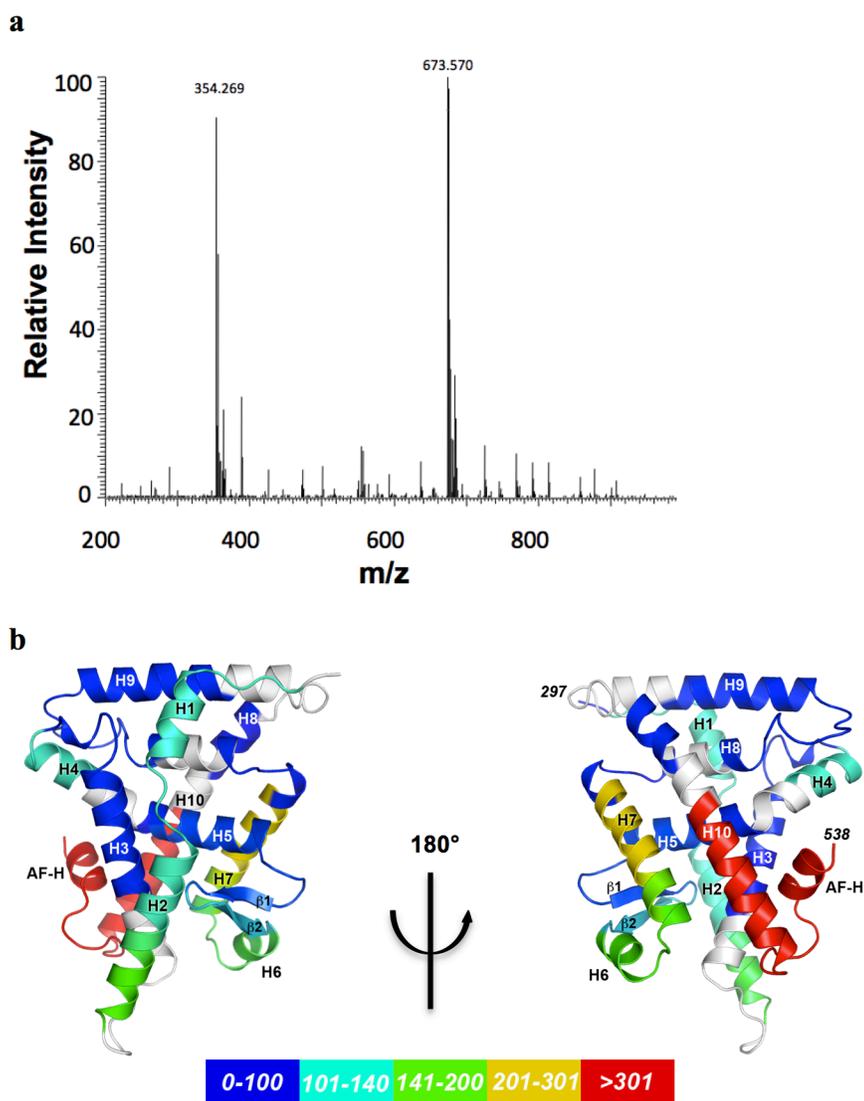


Figure 2.5 Characterization of apo LRH-1.

(a) Negative ESI MS of apo LRH-1 showing no evidence for bound phospholipid. (b) Degree of protection from chymotrypsin proteolysis conferred by DLPC binding (Table S2) mapped with color on DLPC-LRH-1 (a higher number is a higher degree of protection from proteolysis after DLPC binding). Residues that were not observed via MS are colored grey.

β -sheet–helix 6 mobility is essential for PL driven transactivation

To identify novel mobile regions of LRH-1 that are sensitive to ligand status, we performed a low-resolution proteolysis protection assay coupled to quantitative mass spectrometry. By mapping the chymotrypsin proteolysis patterns in DLPC bound vs. apo receptor we were able to identify regions of the protein that were highly mobile and protease sensitive that are stabilized upon DLPC binding (Fig. 2.5b and Table 2.3). Contrary to our expectations we found that the most mobile portion of apo LRH-1 is helix 10 and the AF-H. This observation directly refutes the idea that the AF-H of LRH-1 is rigid and insensitive to ligand status (119-121; 134). To quantify the relative conformational mobility of the receptor in the both the apo and ligand bound states we again employed high resolution HDX, revealing that the β -sheet–helix 6 region is even more dynamic than the AF-H (Fig. 2.6a–b and Fig. 2.7). We were also surprised to find regions of the receptor, such as helix 9 and 10, specifically stabilized upon lipid binding. These helices have been recently reported to be the interaction site for β -catenin, which serves as a coactivator for LRH-1 (151). These data suggest that β -catenin's coactivation of LRH-1 may be ligand regulated.

Structure of the apo LRH-1–SHP NR Box1 complex

To visualize the structural perturbations introduced in the absence of ligand, we determined a structure of apo LRH-1, stabilized by a fragment of the atypical corepressor SHP, to a resolution of 1.9 Å representing an inactivated form of the receptor to juxtapose with the active LRH-1–DLPC complex (Fig. 2.6c and Table 2.1). Unexpectedly, we found that in the absence of PL, residues 397-421 are completely disordered including both β -strands and helix 6 which form one wall of the LBP and one half of the mouth of the receptor (Fig. 2.6c–d and Fig. 2.8). Since this region forms extensive contacts with DLPC, and was identified as having an altered HDX profile when complexed with different length PLs, we hypothesized that this mobile element may be important for sensing and transmitting ligand status. To test whether conformational flexibility

AA number	AA sequence	Protection Factor
299-323	SSPASIPHLILELLKCEPDEPQVQAKIM	119
324-325	AY	137
326	L	147
327-333	QQEQANR	155
334-347	SKHEKLSTFGLMCK	ND
348-354	MADQTLF	14
355-369	SIVEWARSSIFFREL	33
370-377	KVDDQMKL	115
378-382	LQNCW	ND
383-385	SEL	1
386-405	LILDHIYRQVVHGKEGSIF	47
406	L	63
407-413	VTGQQVDY	89
414-424	SIASQAGATL	153
425-427	NNL	211
428	M	239
429-438	SHAQELVAKL	282
439-446	RSLQFDQR	8
447-452	EFVCLK	ND
453-481	FLVLFSLDVKNLENFQLVEGVQEQVNAAL	34
482-495	LDYTMCNYPQQTEK	ND
496-502	FGQLLR	15
503-507	LPEIR	ND
508-539	AISMQAEEYLYYKHLNGDVPYNNLLIEML HAK	439
540-541	RA	ND

Table 2.3 Proteolysis protection data showing spectral counts for trypsin fragments.

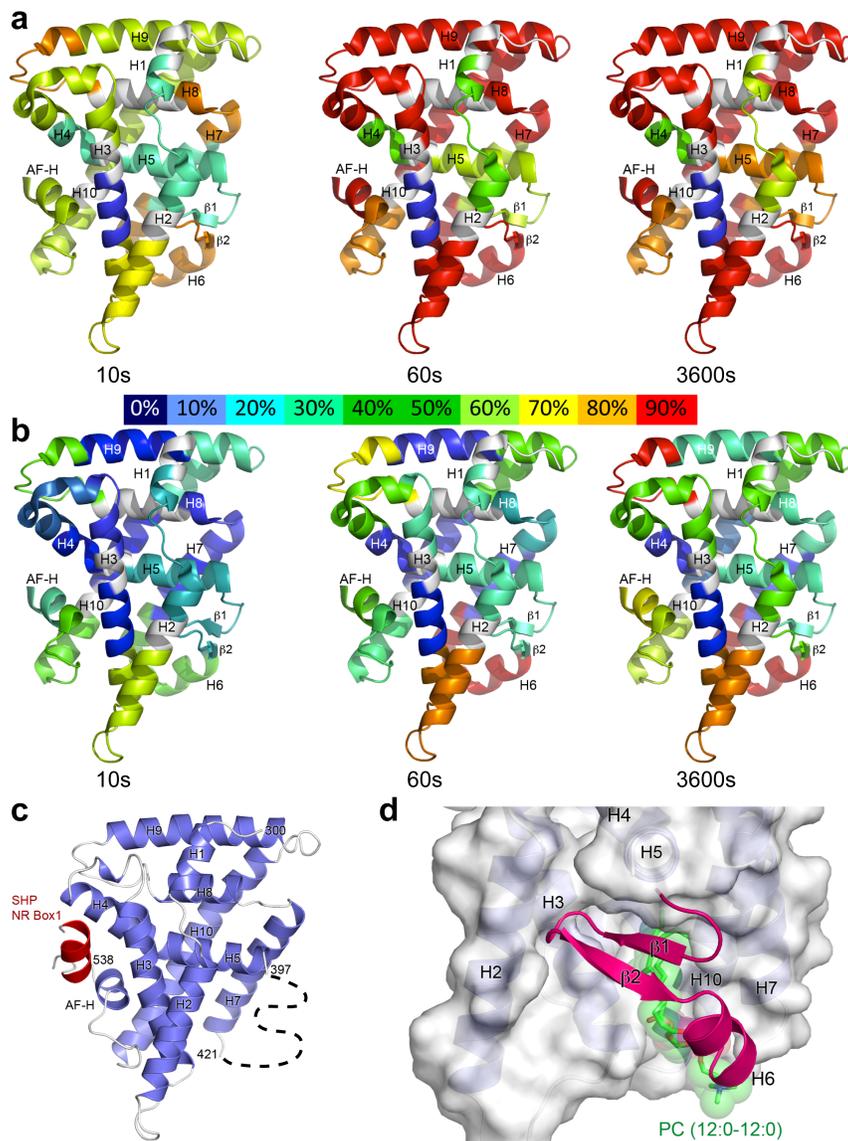


Figure 2.6 Structure of apo LRH-1 identifies a novel mobile activation function region.

(a) Percent deuterium incorporation over time for apo LRH-1 LBD and (b) native LRH-1 LBD mapped to PDB ID code 1YUC(116). (c) Ribbon diagram of apo LRH-LBD (blue) with the SHP NR box 1 peptide (red). Residues 398–420 lack traceable main chain density and have been omitted from the structure (dashed line). (d) Molecular surface of apo LRH-1 (grey) with residues 398–420 (pink) and DLPC (green) superposed from the LRH-1–DLPC complex. Figures were generated in PyMOL (Schrödinger, LLC).

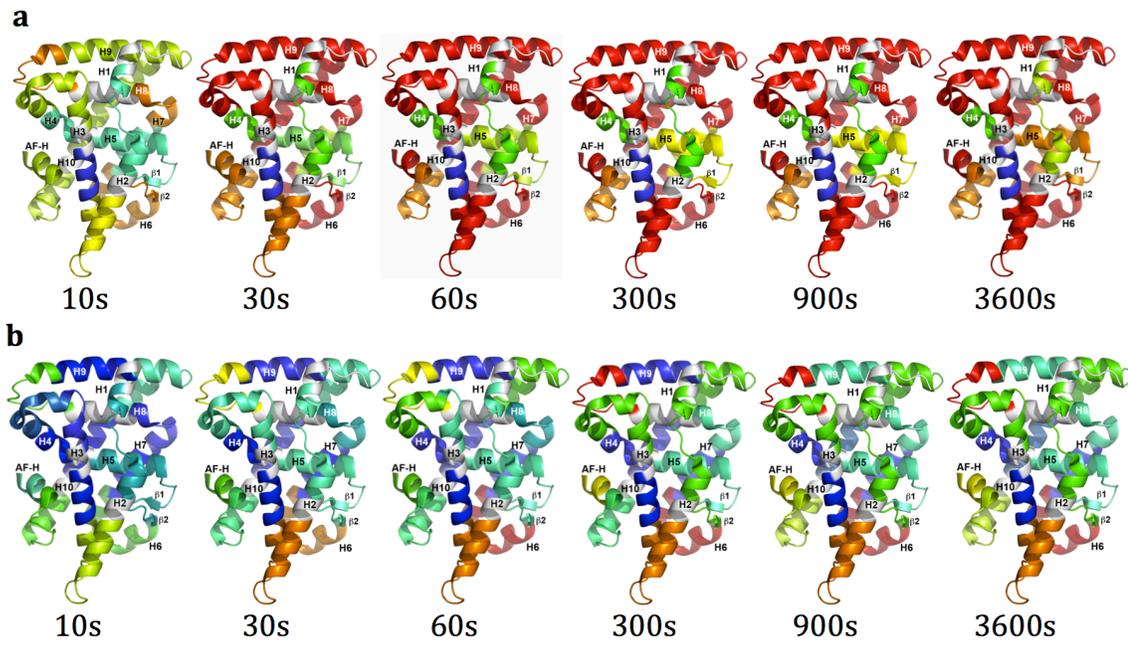


Figure 2.7 Percent deuterium incorporation over time.

Percent deuterium incorporation over time for **(a)** apo LRH-1 LBD and **(b)** native LRH-1 LBD mapped to PDB ID code 1YUC (116).

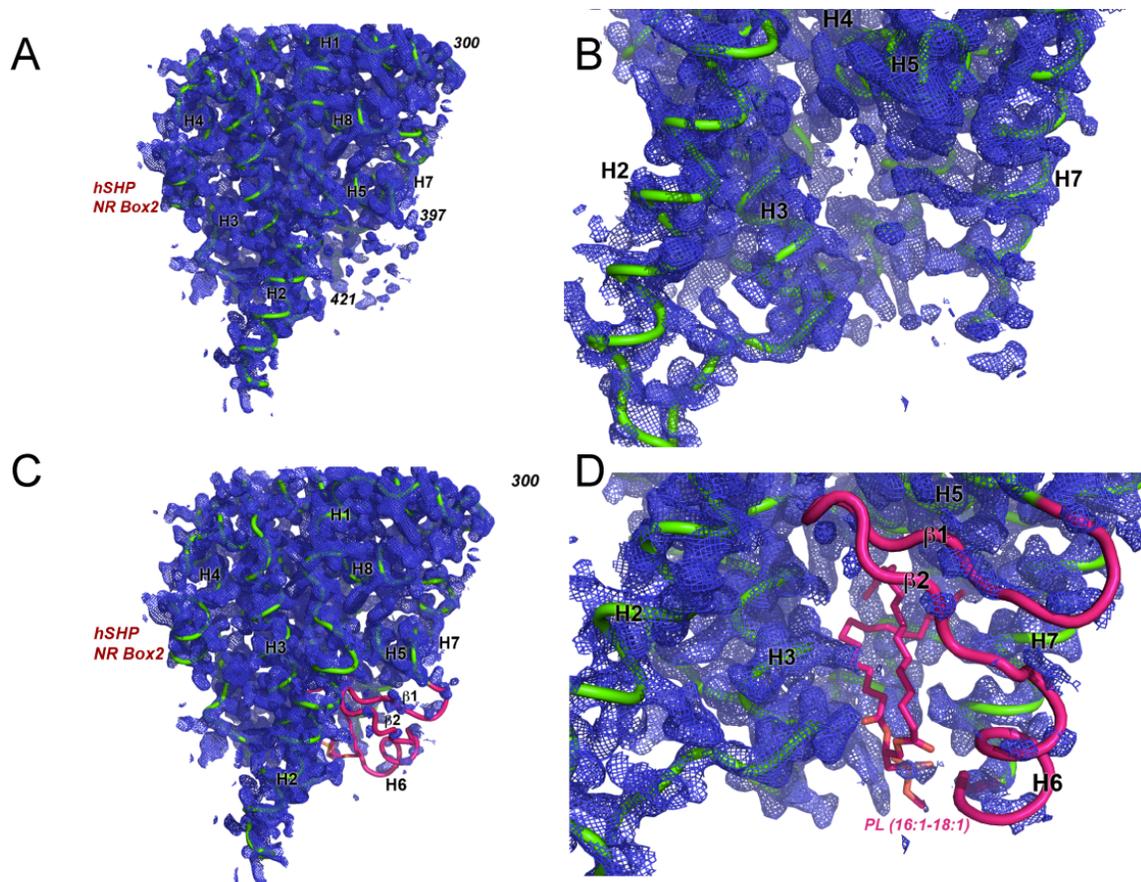


Figure 2.8 Structure of apo LRH-1 Displays Disorder around LBP.

$2F_o - F_c$ electron density contoured at 1σ showing no evidence for either bound PL or for amino acids residues 397-421 comprising the $\beta 1$ -H7 region of the ligand binding pocket. apo LRH-1 is rendered as a cartoon loop with observed residues in green (**a-b**). (**c-d**) Identical views of apo LRH-1 showing the location of missing residues (magenta) that are well ordered in the LRH-1-DLPC complex. Attempts to model either PL or the missing apo LRH-1 residues resulted in clear negative $F_o - F_c$ electron density (not shown).

in this region is required for efficient PL binding and transcriptional activation we individually mutated two conserved flanking glycines at positions 398 and 421 to alanine (Fig. 2.6c). In support of our hypothesis we found that the both mutations significantly reduced LRH-1's ability to activate transcription (Fig. 2.4d). Only the G421A mutation significantly reduced the ability of the protein to bind PLs when purified from *E. coli* (Fig. 2.4a); however, G421 participates in a backbone amide H-bond with the PL phosphate. While a corresponding alanine mutation should preserve this interaction, based on the allowed backbone torsion angles, the greater conformational mobility of G241 clearly plays an important role in recognizing PLs. Thus, we have identified the β -sheet-helix 6 region as a novel activation function of LRH-1, critical for sensing ligand status and for driving receptor activation.

Empty receptor binds with high affinity to SMRT

Since our data show that the AF-H is mobile in the absence of ligand, we investigated whether apo LRH-1 is capable of binding to a traditional corepressor such as SMRT. Contrary to studies using receptor co-purified with contaminating *E. coli* PLs (136), we found that a peptide derived from SMRT containing the corepressor motif (LxxxIxxxI/L) binds to apo LRH-1 with 49.6 nM affinity (Fig. 2.4c and Table 2.2). This is by far the highest affinity interaction between LRH-1 and a coregulator derived peptide that we have tested, supporting our observation that in the absence of ligand, the AF-H is free to rotate away from helix 3 and helix 4 to allow traditional corepressor binding.

Discussion

The traditional model explaining NR transcriptional activation describes a system where high affinity ligand binding drives a conformational switch from an inactive to an active state involving a repositioning of the AF-H. Appropriate NR activation is the result of a fine balance between receptor stability and ligand affinity. In the absence of ligand, NRs populate a partially unfolded or “molten globule” state where ligand binding catalyzes proper receptor folding and

activation (114; 152). The binding energy gained from contacts with the ligand is not enough to explain the high affinity; rather, ligand binding allows for additional intramolecular contacts at the mouth of the LBP between helices 2-3, 6-7, and 10-AF-H which ultimately supports receptor activation(114). In contrast, it is clear from the LRH-1–DLPC structure that activating PLs do not facilitate equivalent direct interactions. Rather, LRH-1 relies on PLs to bridge these critical intramolecular interactions with intermolecular interactions to achieve receptor activation. In this way, LRH-1 has tinkered with the canonical molecular switch, adapting it to respond to diverse PL ligands by using the phosphoglycerol backbone to transmit a signal from helix 6 to the AF-H while relying on deep pocket interactions with the lipid tails to fine tune receptor dynamics and thus co-regulator specificity.

In the absence of ligand, we show by HDX that LRH-1 undergoes rapid folding-to-unfolding transitions for several key regions within the ligand binding domain, including the helices 2, 5, 6 and the AF-H, with the majority of amide protein exchange occurring within 30–60 s. This unexpected structural plasticity likely confers SMRT interaction, which requires displacement of the AF-H from the active orientation to bind. This data is in line with previous reports showing that SMRT is capable of repressing LRH-1 *in vivo* in a dose dependent manner and provides the first direct evidence that LRH-1 may be sensitive to repression by direct interactions with traditional corepressors (134; 136). Since, SMRT does not bind a LRH-1–PL complex, it is likely that at least a portion of LRH-1 remains in the apo form in cells. In addition, SHP is able to bind both the apo receptor and the receptor loaded with bacterial PLs; however, SHP binding is completely lost upon DLPC addition. Since LRH-1's repression in the liver dictates much of LRH-1's effects on gene expression (123; 153), the effects of DLPC in the liver may be due in part to relieved SHP and SMRT repression rather than enhanced coactivator interaction.

Like a true agonist, DLPC simultaneously enhanced coactivator peptide recruitment while disfavoring both typical and atypical corepressor peptide interaction (Fig. 2.4c and Table

2.2). However, the coactivators SRC-1 and TIF2 retain their ability to interact with LRH-1 in the absence of ligand explaining its low basal activity. These results clearly show that LRH-1 undergoes profound structural changes upon ligand binding and definitively confirms LRH-1's viability as a therapeutic target for both agonist and antagonist design. Our results also suggest that dietary PLs may have signaling effects outside hepatic tissues. For example, if DLPC is trafficked out of the liver it may exacerbate LRH-1's malicious role in breast cancer by selectively recruiting PGC-1 α to Aromatase and Estrogen Receptor promoters, driving increased estrogen synthesis to fuel local tumor growth. Finally, these findings may facilitate the development of tissue selective LRH-1 modulators by revealing discrete regions of the LBP required for the recruitment of distinct coregulators.

Chapter 3: Divergent sequence tunes ligand sensitivity in phospholipid-regulated hormone receptors³

³This chapter has been slightly modified from the published manuscript: Musille PM, Pathak M, Lauer JL, Griffin PR, Ortlund EA. (2013) Divergent sequence tunes ligand sensitivity in phospholipid-regulated hormone receptors. *J Biol Chem.* 288(28):20702-12.

Introduction

The human liver receptor homologue-1 (LRH-1; NR5A2) is a member of the NR5A class of NRs that regulate the expression of genes central to embryonic development, cell cycle progression, reproduction, and lipid homeostasis (122) in response to activating PLs (154). This family includes steroidogenic factor-1 (SF1; NR5A1) and *Drosophila melanogaster* fushi tarazu factor 1 (Ftz-F1; NR5A3). SF-1 plays a role in steroidogenesis and the proper development of the testes and adrenal glands (51; 155) and the founding member of the family, Ftz-F1, controls segmentation in flies (156).

LRH-1 plays a crucial role in early embryonic development as it is required to maintain Oct4 expression in undifferentiated embryonic stem cells (124; 157-160), which renders LRH-1 knockout mice unable to progress past embryonic day 6.5 (161). Beyond development, overexpression of both LRH-1 and SF-1 drives the reprogramming of murine somatic cells to pluripotent stem cells without requiring simultaneous overexpression of Oct4 (22). LRH-1 overexpression appears to drive the expression of Nanog and works synergistically with other well known factors such as Sox2 and Klf4, to mediate cellular reprogramming (22). In fact, LRH-1 has been implicated as a new stem cell factor since it is the only protein discovered to date that can replace Oct4, which until now was considered absolutely required to manipulate cells into a pluripotent state (157).

In adults, LRH-1 is expressed predominantly in the liver, small intestine, preadipocytes, ovary, placenta and brain (162). In the ovary, LRH-1 regulates ovarian steroidogenesis through control of CYP19 transcription (163). In hepatic tissues, LRH-1 regulates genes central to bile acid homeostasis, lipid and cholesterol absorption, and cholesterol reverse transport (35; 126; 132; 162). Identifying endogenous or synthetic small molecule modulators of LRH-1 activity may lead to promising therapies to treat conditions ranging from metabolic to neoplastic diseases.

LRH-1, like most NRs, interacts with coactivators through an LXXLL motif (where X is any amino acid). Recent studies showed that apoLRH-1 interacts with widely expressed corepressors

such as SMRT and NCoR (37; 41; 136) in addition to atypical NRs which have evolved to specifically and efficiently repress LRH-1 by mimicking coactivators (127; 132).

Although the endogenous ligand for hLRH-1 is currently unknown, Lee et al. recently showed that LRH-1 is specifically activated by the exogenous medium chain phosphatidylcholine isoforms, diundecanoyl (DUPC, PC 11:0/11:0) and dilauroyl (DLPC, PC 12:0/12:0) phosphatidylcholine (42). These medium chain PC agonists increase the ability of LRH-1 to interact with coactivators, and reduce blood lipid and glucose levels in diabetic mice in a LRH-1 dependent manner (42). We have shown that DLPC is able to bind to the LRH-1 ligand binding domain (LBD) and activate the receptor by altering receptor dynamics at both an alternate activation function surface and the canonical activation function helix (AF-H) (41). The alternate activation function region in hLRH-1 is comprised of residues 398-421 and makes direct contact with bound PLs. We showed that the dynamics of this region is coupled to ligand binding and restricting motion in this region ablates receptor activation (41). Orthologs of hLRH-1, such as mLRH-1 and dmFTZ-f1, have evolved divergent sequences in this region potentially altering ligand binding and response.

Indeed, mLRH-1 showed no evidence of PL binding in crystal structures as a direct result of this late evolutionary adaptation, which resulted in six amino acid substitutions within the alternate activation function that presumably stabilizes the ligand binding pocket (LBP) in the absence of ligand (37; 41). Surprisingly, the AF-H of apo mLRH-1 was in the active conformation despite the presence of a large empty LBP (22). Subsequent mass spectroscopy analysis showed that mLRH-1 is capable of binding to PCs; however, PL binding was reduced compared to human NR5A receptors (40). In line with these results, humanization of mLRH-1, by reversing a key sequence substitution in the alternate activation function region, increased sensitivity to PL regulation suggesting that the mechanism for PL-driven activation has diverged in rodents (40). Placement of the derived rodent sequence in hLRH-1, slightly reduced PL binding, minimally impacted transactivation in HeLa and MCF-7 cells yet reduced transactivation

when transiently overexpressed with SRC-2 and SRC-3 compared to wild-type hLRH-1 (39). Recent work however, has shown that mLRH-1 is as robustly activated by DLPC as the human receptor (42). These conflicting observations highlight the need to understand the structural mechanism allowing such a divergent sequence at the alternate AF to support PL driven regulation (39; 40; 64; 67; 68).

Mice serve as an important model system to study stem cell biology and both normal and aberrant hepatic biology including biliary cirrhosis, lipid dysregulation, and diabetes. Flies serve as powerful developmental models. Since both models are used to study LRH-1 biology it is critical to determine how LRH-1 orthologs differentially interact with PLs.

To address this, we use biochemical assays and mass spectrometry to show that mLRH-1 binds to PLs while FTZ-f1 does not – suggesting that mLRH-1 is PL regulated while FTZ-f1 is ligand independent. To isolate the effects of the rodent-specific sequence adaptations, we test the ability of a variant form of LRH-1, mouse-loop LRH-1 (mlLRH-1), to interact with co-regulator peptides both in the absence and presence of bound PLs *in vitro*. Further, we determine the structure of the apo mlLRH-1 variant to 2.75 Å resolution. Finally, we use hydrogen deuterium exchange coupled with mass spectrometry (HDX-MS) to show that the mouse-loop sequence stabilizes the alternate activation function surface and AF-H in the absence of PL, while only minimally impacting PL binding.

Experimental Procedures

Reagents:

Chemicals were purchased from Sigma, Fisher or Avanti PLs. pMALCH10T and the vector for His tagged TEV were a gift from John Tesmer (UT Austin). pLIC_MBP and pLIC_HIS were gifts from John Sondek (UNC, Chapel Hill). Peptides were synthesized by Synbiosci (Livermore, CA).

Protein Expression and Purification:

The hLRH-1 and mLRH-1 ligand binding domains (LBDs), residues 291-541, were purified as described previously (39). Pure mLRH-1 LBD was dialyzed against 100 mM ammonium acetate (pH 7.4), 1 mM DTT, 1 mM EDTA, and 2 mM CHAPS, and concentrated to 3-5 mg/mL prior to crystallization. The dmFtz-F1 LBD, residues 791-1025, and mLRH-1 LBD, residues 320-557, were cloned into the pLIC_MBP vector C-terminal to a cassette containing a 6xHis tag, maltose binding protein (MBP), and a TEV protease cleavage site. The fusion proteins were expressed in BL21(DE3) pLysS cells using standard methods and purified using affinity chromatography with TEV cleavage of the fusion partners.

Structure Determination:

Crystals of mLRH-1 LBD were grown by hanging drop vapor diffusion at 22 °C from solutions containing 0.75 mL of protein at 6.5 mg/mL protein and 0.75 mL of the following crystallant: 9.5%-15% PEG 3350, 5% glycerol, and 50 mM Bis-Tris, pH 6.4. Crystals were cryoprotected in crystallant containing 20% glycerol and flash-frozen in liquid N₂. Data to 2.75 Å resolution were collected at 100 K at the Southeast Regional Collaborative Access Team (SER-CAT) at the Advanced Photon Source (Argonne, IL), and were processed and scaled with HKL2000 (Table 1) (138). Initial phases were determined using the structure of the mLRH-1 LBD (1PK5) as a molecular replacement search model(37). The crystals were pseudo-merohedral twins with a twinning fraction of 45% and the data were de-twinning using the -h, l, k operator in Detwin(140). The structure was refined using the REFMAC5 within CCP4 suite of programs (139; 140) and

model building was performed in COOT (141; 142). The final model contains two mLRH-1 LBD monomers (residues 300-528) and exhibits good geometry (164). Two loops (H2-H3 residues 332-337 and H11-H12 residues 528-530) displayed poor density presumably due to high disorder and were omitted from the final model. The R_{factors} for the final model are 22.4% and 25.7% for R and R_{free} respectively. MolProbity was used for model validation, indicating that 95.3% of the residues fall in the most favored regions of the Ramachandran plot and none in disallowed regions. The overall MolProbity score was 2.51, placing mLRH-1 in the 91st percentile for overall geometric quality among protein crystal structures of comparable resolution (143).

Mass Spectrometry:

Samples were analyzed using electrospray ionization in the negative-ion mode to detect and identify PLs. Approximately 6 mg of wild-type or mutant forms of LRH-1 LBD and Ftz-F1 LBD were extracted with a 2:1 chloroform/methanol solution, diluted in 200 μL of chloromethylene, and analyzed by negative ion ESI/MS on a Thermo LTQ FTMS using direct injection analysis with electrospray ionization (Thermo Finnigan, Somerset, NJ). All extractions were performed in duplicate. The high-resolution analyses were performed in the FTMS at a resolution of 100000 at 400 m/z . The MS/MS experiments were done in the ion trap portion of the instrument with a mass selection of 3 atomic mass units (amu) and a normalized collision energy of 30 V. The major PL species were identified by accurate mass measurements and MS/MS via collisionally induced dissociation (CID), which yields product ions characteristic of the head groups and attached fatty acids. Acquisition and analyses were performed using the instrument's Analyst QS software.

Phospholipid Quantification:

Preceding PL quantification, 1 mg of protein was digested by 0.45 ml of 8.9 N sulfuric acid at > 200 $^{\circ}\text{C}$ for 25 minutes in glass tubes. Tubes were allowed to cool before the addition of 150 μL of hydrogen peroxide. Tubes were again heated to > 200 $^{\circ}\text{C}$ for 30 minutes. 3.9 ml deionized water and 0.5 ml 2.5% ammonium molybdate(VI) tetrahydrate was added and tubes were vortexed 5

times each followed by the addition of 0.5 ml of 10% ascorbic acid solution and vortexing. Tubes were capped and heated at 100 °C for 7 minutes and then allowed to cool before determining the absorbance of each of the samples at 820 nm. All experiments were performed in triplicate and scaled to hLRH-1.

Generation of apoLRH-1:

Apo LRH-1 and mLHRH-1 were generated using published protocols (41). Briefly, pure protein was subjected to chloroform:methanol extraction (2:1) to remove bound lipids according to the Bligh & Dyer method (144). The resulting pellet containing denatured protein was washed 3 times with chloroform to remove any trace lipids associated with the protein or vessel. The resulting white pellet was then dried by evaporation and resuspended in 6 M Guanidinium HCl. Empty protein was then refolded by fast dilution into a buffer containing 10 mM K₂HPO₄, 100 mM Tris (pH 7.4), 1 mM EDTA, and 500 μM CTAB at 4 °C. After ~ 20 hours, protein was then concentrated and purified by size exclusion chromatography to ensure a homogenous population of refolded receptors.

Cofactor Binding Assays:

The polarization of fluorescein labeled peptides derived from SHP NR box 1 (³H₃N-QGAASRPAILYALLSSSLK-CO₂⁻), PGC-1α NR box 2 (³H₃N-EEPSLLKLLLAPA-CO₂⁻), SRC-1 NR box 2 (³H₃N-SPSSHSLTERHKILHRLQEGSP-CO₂⁻), SMRT (³H₃N-TNMGLEAIIRKALMGKYDQW-CO₂⁻), NCoR ID2 (³H₃N-DPASNLGLEDIIRKALMGSFDDK-CO₂⁻) or TIF2 NR box 3 (³H₃N-PVSPKKKE NALLRYLLDKDDT-CO₂⁻) was monitored with a BioTek Synergy 4 spectrophotometer with polarizers (Winooski, NJ) as a function of protein concentration. Experiments were conducted in 150 mM sodium chloride, 20 mM Tris-HCl (pH 7.4) and 5% (v/v) glycerol. All experiments were done in triplicate, and data were fit with Prism 5 (GraphPad) by the linear least-squares methods to a single site-binding model.

Hydrogen-Deuterium Exchange Mass Spectroscopy:

Solution-phase amide HDX was carried out with a fully automated system as described previously (149). Briefly, 4 μ l of protein was diluted to 20 μ l with D₂O-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 μ L in a low pH and low Temp buffer containing 0.1% (v/v) TFA in 5 M urea (held at 1 °C). Samples were then passed across an immobilized pepsin column (prepared in house) at 50 μ l 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) CH₃CN to 40% (w/v) CH₃CN, 0.3% (w/v) formic acid over 5 min, 2 °C) across a 1 mm \times 50 mm C18 HPLC column (Hypersil Gold, Thermo Fisher) and electrosprayed directly into an Orbitrap mass spectrometer (LTQ Orbitrap with ETD, Thermo Fisher). Data were processed with in-house software and visualized with PyMOL (Schrödinger, LLC). To measure the difference in exchange rates, we calculated the average percent deuterium uptake for apo mLRH-1 LBD 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 apo hLRH-1 LBD. Negative perturbation values indicate exchange rates are slower for these regions within apo mLRH-1.

Accession Numbers:

Coordinates and structure factors have been deposited in the Protein Data Bank with accession number 4IS8.

Results

Overall Structure

While nearly all the residues contacting bound PLs are conserved in the LRH-1/SF-1 family, residues 419-QAGATL-424 in hLRH-1 are replaced by 438-HTEVAF-443 in mLRH-1 as a result of a late evolutionary divergence in the rodent lineage (Figure 3.3D) (40). This six amino acid replacement creates a salt bridge between the H7 and H11 at the opening of the LBP, which presumably precludes the binding of PL as observed in hLRH-1 and SF-1 (37; 64). In line with this hypothesis, removal of the H7-H11 salt-bridge, via an E440G mutation in mLRH-1, enhanced sensitivity to PL levels (37; 64). To investigate the role of these rodent-specific residues in PL binding, we crystallized the LRH-1 LBD with this six-residue-replacement, termed mLRLH-1 (39), and determined its structure to 2.75 Å resolution (Table 3.1, Figure 3.1A). The mLRLH-1 LBD crystallized with two molecules in the asymmetric unit in the P2₁2₁2₁ space group. Previous studies showed that the mLRLH-1 LBD contains abundant phosphatidylethanolamine and phosphatidylglycerol species ranging in acyl tail length from C14 to C20, which fortuitously co-purify from *E. coli* (39). We again confirmed the presence of PLs using mass spectrometry, which showed strong peaks for several PL isoforms (Figure 3.1B). Despite the presence of bound *E. coli* PLs and coactivator peptide in the crystallization conditions, electron density for either bound PL or coactivator peptide was not evident. To confirm PL was absent in the crystal structure, we modeled a C16:1-C18:1 phosphatidylethanolamine in the LBP. Clear negative F_O-F_C electron density was observed encompassing the modeled PE backbone, confirming the mouse-loop structure is free of bound PLs (Figure 3.1C).

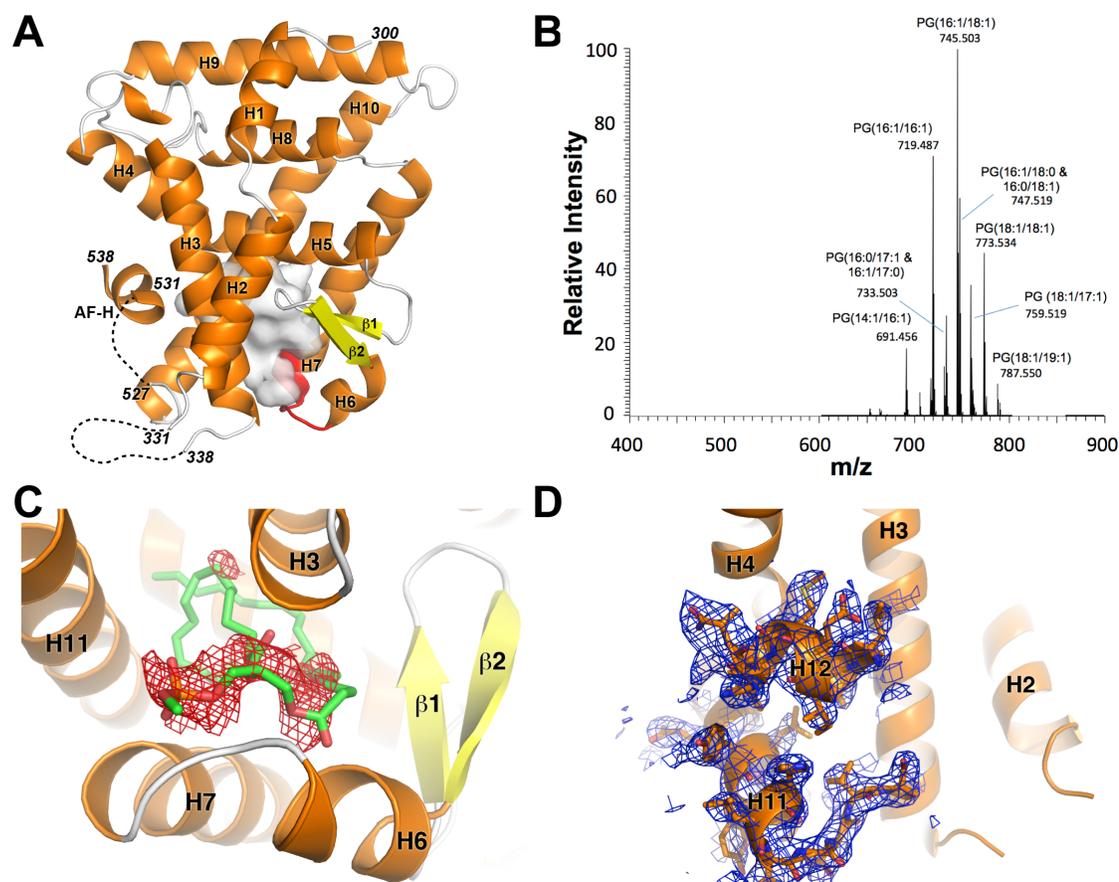


Figure 3.1 Overall Structure of apo mLRH-1.

Overall Structure of apo mLRH-1. A. Cartoon representation of mLRH-1 LBD with α -helices in orange, β -strands in yellow, and the six-residue rodent derived amino acid replacement in red. The empty LBP cavity is depicted as a white transparent surface. Disordered residues are represented by a black dashed line. B. Electrospray injection mass spectrometry (ESI-MS) of mLRH-1. Phospholipid peaks are labeled as phosphatidylglycerol (PG) with the acyl carbon tail lengths and unsaturation characterized by collision induced dissociation (CID). The Y-axis is scaled relative to the most abundant peak observed in the spectra. C. F_0 - F_c electron density (red) contoured at -3σ for phosphatidylethanolamine modeled into the LBP of mLRH-1. D. $2F_0$ - F_c electron density contoured at 1σ showing the partially active confirmation of mLRH-1 evidenced by the discontinuous electron density between helices 11 and 12.

Data collection and refinement statistics	
	mLLRH-1
Wavelength (Å)	1.00
Resolution (Å)	2.8 (2.85– 2.80)*
Space Group	P2 ₁ 2 ₁ 2 ₁
Unit Cell Dimensions	
a, b, c (Å)	36.3, 120.0, 123.8
α=β=δ (°)	90.0
No. of Reflections	14223
R ^a _{sym} (%)	7.6 (35.7)
Completeness (%)	98.2 (83.7)
Ave. Redundancy	4.4 (3.6)
I/σ	11.8 (3.2)
Monomers per asymmetric unit (AU)	2
No. of protein atoms/AU	1874
R ^b _{working} /R ^c _{free}	0.22/0.25
Ave. B-factors, Å ²	
Protein	56.5
r.m.s. deviations	
Bond lengths, Å	0.006
Bond angles, °	1.4

* Highest resolution shell is shown in parenthesis.

^a $R_{\text{sym}} = \frac{\sum |I - \langle I \rangle|}{\sum I}$, where I is the observed intensity and $\langle I \rangle$ is the average intensity of several symmetry-related observations.

^b $R_{\text{working}} = \frac{\sum ||F_o| - |F_c||}{\sum |F_o|}$, where F_o and F_c are the observed and calculated structure factors, respectively.

^c $R_{\text{free}} = \frac{\sum ||F_o| - |F_c||}{\sum |F_o|}$ for 7% of the data not used at any stage of the structural refinement.

Table 3.1 Data collection and refinement statistics for the mLLRH-1 crystal structure.

As was observed for the mLRH-1 structure (1PK5), the mLRH-1 AF-H adopts the “active” orientation and the opening of the LBP appears “closed” with a 2.9 Å ionic interaction between residues Glu421 and Lys520 blocking access to a well defined internal cavity with a volume of ~941 Å³. The side chain of Glu421 occupies the equivalent position of the lipid phosphate moiety in the hLRH-1–PL complex and the phenyl ring of Phe424 occupies the same space as the phosphoglycerol backbone (Figure 3.2A-B). A similar closed conformation was observed in a recent structure of the hLRH-1 LBD bound to a small molecule agonist (165), though by a completely different mechanism. Here, Gln419 H-bonds with the backbone amide of Phe342 to bridge H6 and H3 rather than the H6–H11 interaction facilitated by the mouse-loop substitutions (Figure 3.2C). Thus, the rodent-specific substitutions appear to stabilize this closed conformation of the LRH-1 LBD without the requirement for a small molecule ligand.

Apo mLRH-1 adopts a destabilized active orientation

The residues located in the loop preceding the AF-H are known to play a critical role in NR activation and mutations in this region dramatically alter co-regulator recruitment (37; 39; 166). In the mLRH-1 structure, there is no interpretable electron density for the H11-H12 loop preceding the AF-H (Figure 3.1D). This is mirrored in the structure of mLRH-1 which shows discontinuous electron density in the pre AF-H loop and roughly 2-fold higher B-factors (~80 Å² versus 45 Å² for all protein atoms) in this region (37). Thus, this region is destabilized in the absence of ligand. While this disorder does not displace the AF-H, it likely represents the structure of a “destabilized agonist” conformation previously observed in NRs when complexed with weak competitive antagonists (167-169).

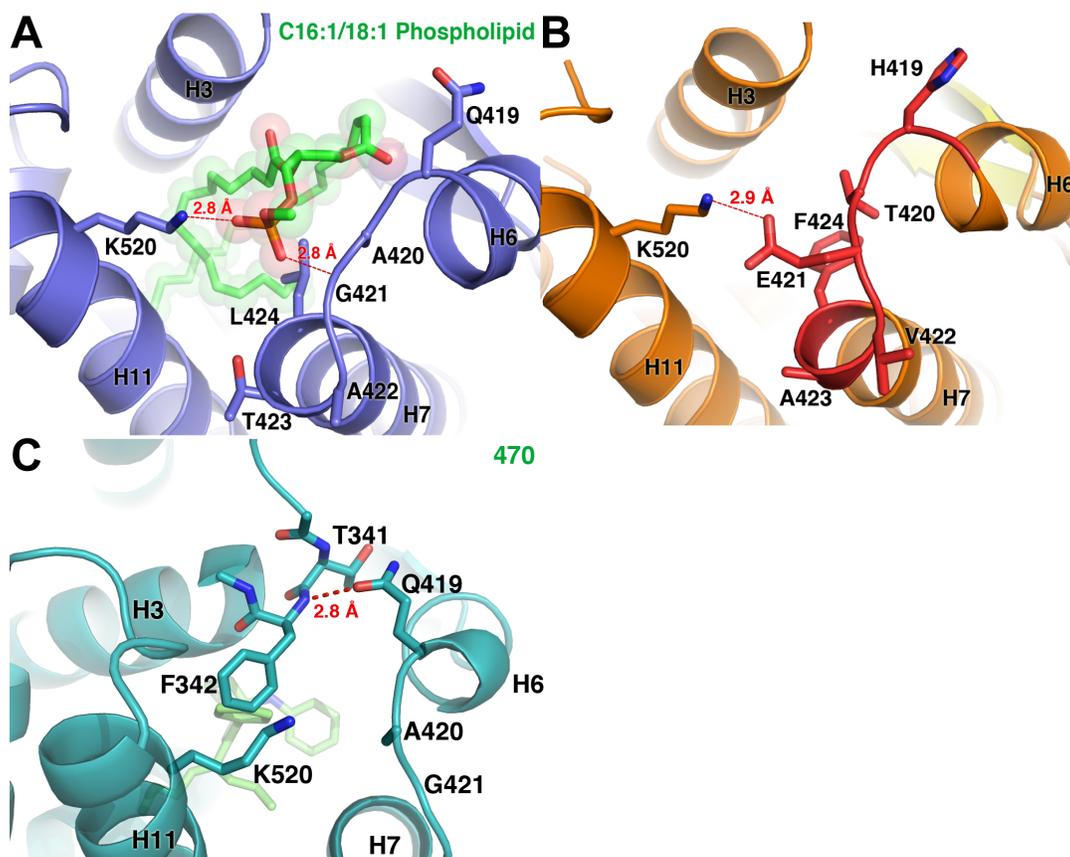


Figure 3.2 Opening to the LBP in the human and mLRH-1.

View of the opening of the LBP for A. hLRH-1 bound to an *E. coli* phospholipid (green) PDB ID: 1YUC (blue) B. MILRH-1 (orange) and C. hLRH-1 bound to a small molecule agonist (green) PDB ID: 3PLZ mLRH-1 (orange).

Sequence divergence allows to NR5A receptors adopt multiple conformations to achieve the active state

Since the *Drosophila* ortholog of LRH-1, Ftz-f1, also displays sequence divergence in the alternate activation function region (Figure 3.3D), we analyzed the dmFtz-F1 (NR5A3) crystal structure (PDB ID: 2XHS) (170). Interestingly, Ftz-F1 also crystallized in the “active” orientation without evidence for PLs in the LBP. Instead, residues in the β 1- β 2 and H6-H7 region adopt an unprecedented conformation that turns 90° inward towards the ligand pocket, filling the cavity occupied by PL in the human NR5A receptors (Figure 3.3A-C).

We have previously shown that flexibility in this region is important for hLRH-1 activation (41). Given the mobility of elements comprising this “wall” of the LBP among NR5A receptors and the fact that crystallization selected for an empty population of mLRH-1 receptors, it is possible crystallization also selected for apo Ftz-f1. We therefore asked whether mLRH-1 and Ftz-F1 are able to adjust their conformation to bind PLs when expressed and purified from *E. coli*.

Do mLRH-1 and Ftz-F1 bind PLs despite their ability to crystallize empty?

Using mass spectrometry, we discovered that indeed *E. coli* expressed recombinant mLRH-1 binds to PG with the most abundant isoforms being PG(18:1/18:1) and PG(18:1/16:1) (Figure 3.4A). Mass spectrometry also detects peaks in the 500 m/z range; however, these peaks

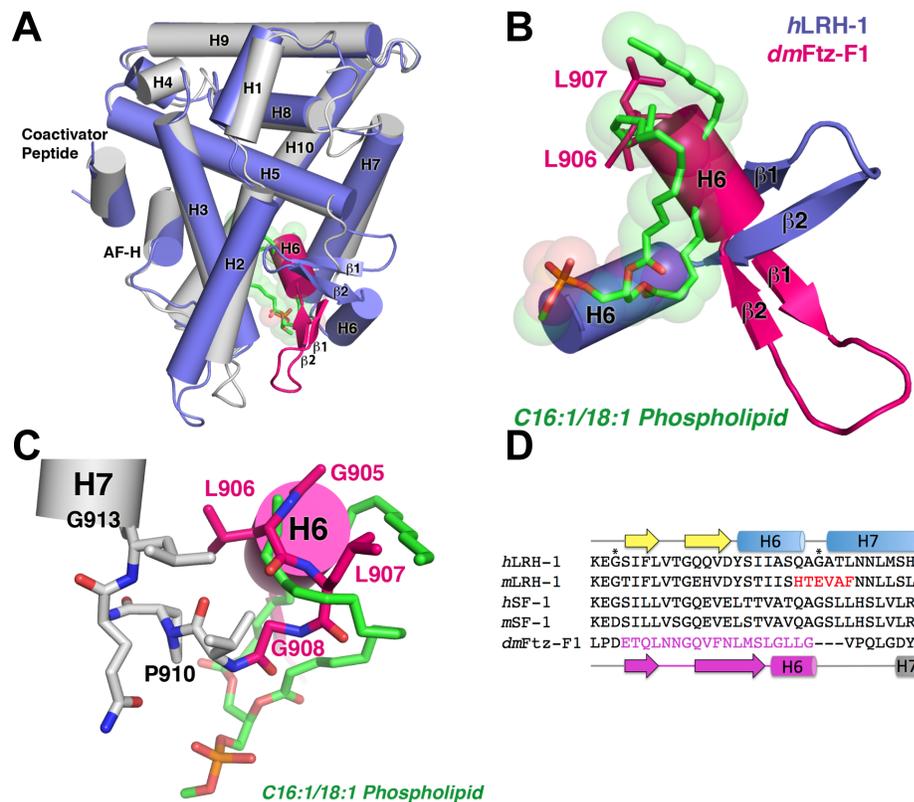


Figure 3.3 NR5A receptors have multiple modes for achieving the active conformation in the absence of ligand.

A. Superposition of the *hLRH-1*:phospholipid complex (1YUC; blue) on empty *Drosophila melanogaster* Ftz-F1 (2XHS; grey/pink). Ftz-F1 residues highlighted in pink are radically repositioned with respect to all other NR5A family members and disrupt phospholipid binding. B. Close up view of the β -sheet-H6 region inside the LBP highlighting the $\sim 90^\circ$ rotation of β -sheet-H6 region in the empty Ftz-F1. C. Close of view of the Ftz-F1 depicting the residues critical for stabilizing the empty LBP. The secondary structure is indicated and key residues are shown as sticks. D. Sequence alignment showing the β 1-H7 region of the LBP in NR5A receptors. The amino acids in pink correspond with the repositioned residues in Ftz-F1. Residues in red are unique to rodent LRH-1 and were mutated to create *mLRH-1*. Residues marked by asterisks are the glycines previously found to border LRH-1's alternate activation function.

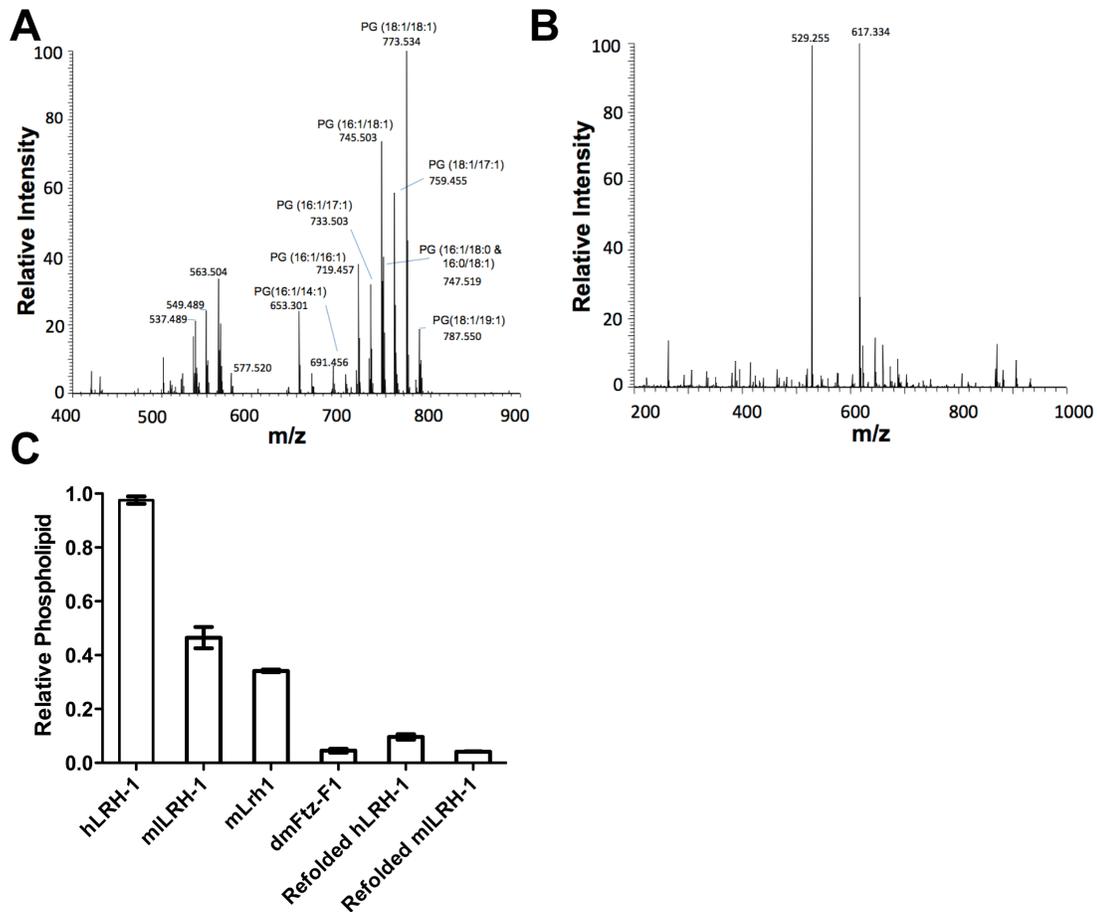


Figure 3.4 Phospholipid analysis of mLRH-1 and Ftz-F1.

Electrospray injection mass spectrometry (ESI-MS) of mLRH-1 A. and Ftz-F1 B. Phospholipid peaks are labeled as either phosphatidylglycerol (PG) or phosphatidylethanolamine (PE). The Y-axis is scaled relative to the most abundant peak observed in the spectra. C. Total phospholipid quantification of LRH-1 variants following chloroform-methanol extraction. Data are represented as mean \pm s.e.m.; * $P < 0.001$ from three experiments repeated in triplicate.

do not represent PLs and the collision induced decomposition analysis reveals that these peaks likely represent an unknown cyclic compound. Attempts to identify this compound were unsuccessful. Recombinant Ftz-F1, however, showed no detectible PL binding (Figure 3.4B). Two low abundance peaks with a m/z of 529 and 617 were present in the Ftz-F1 MS data at ~0.5% of the signal observed for the major peaks in the mLRH-1 spectra. These peaks do not correspond to known compounds and given their low abundance, attempts to identify these compounds were unsuccessful. To quantify the level of PL bound to Ftz-F1, mLRH-1 and mLHRH-1 we performed an assay to detect lipid phosphorous (Figure 3.4C). In line with the mass spectrometry analysis, Ftz-F1 showed almost no PL binding. mLRH-1 shows a diminished ability to bind *E. coli* PLs *versus* the human receptors but is on par with the mouse-loop variant of the receptor (39).

Testing ligand regulation

Previous cell-based studies showed that the transcriptional activity of mLHRH-1 ranged from nearly wild-type activity to half of that observed for the human receptor depending on cell type (39). Co-transfection with coactivators exaggerated this difference (39). To fully characterize the link between PL binding and receptor activation in this LRH-1 variant, we would ideally compare the coactivator preference and transactivation potential of the apo- versus PL-bound form of mLHRH-1 versus hLRH-1. However, we are currently unable to determine or control the levels of apoLRH-1 in mammalian cells. We are also unable to measure the levels or identity of bound endogenous PLs or whether the receptor is fully occupied with exogenously supplied DLPC when treated with this lipid. Therefore, we tested the impact of the H6-H7 substitutions on receptor activity *in vitro* by measuring the ability of either apo or DLPC-bound human and mLHRH-1 LBD to interact with coactivator and corepressor derived peptides. Refolded receptor was verified PL free following chloroform:menthol extraction (Figure 3.4C). Specifically, we monitored the ability of protein variants to recruit coactivator peptides derived

from SRC-1, TIF2, PGC-1 α , and corepressor peptides derived from SHP, SMRT, and NCoR ID2 using fluorescence polarization (Figure 3.5 and Table 3.2).

As was observed for hLRH-1, apo mLRH-1 bound to all corepressor peptides tested. The addition of DLPC prevented association between mLRH-1 and canonical LXXXIXXXI/L containing corepressor-derived peptides. mLRH-1-DLPC binds much more tightly to TIF than hLRH-1. Surprisingly, while DLPC relieves SHP binding for hLRH-1, the mLRH-1-DLPC complex binds to SHP *in vitro*, which may explain why mLRH-1 was observed to have lower basal activation in cells (39). These results are inline with the hypothesis that the rodent-specific divergence alters the dynamics of the apo LBD and suggests potential differences in SHP interaction and regulation between mice and humans. These data show that DLPC binding alters coregulator preference for mLRH and suggests that rodent divergence may have altered, but not ablated ligand regulation.

mLRH-1 is specifically stabilized versus hLRH-1

The coregulator peptide recruitment data show that apo mLRH-1 binds with higher affinity than apo hLRH-1 to all coactivators suggesting that it is better stabilized in the active orientation. To discover the differences in protein dynamics that underlie this enhanced ability to bind coactivators, we performed HDX-MS (Figure 3.5C), comparing apo mLRH-1 and apo hLRH-1. As hypothesized, the rodent-specific sequence replacement confers increased protection in the helix 6-7 region as evidenced by nearly 40% less deuterium exchange in this region. Surprisingly, no protection was observed for the AF-H, rather, the effect of the mouse-loop replacement is to stabilize the alternate activation function region in LRH-1. Previous work showed that mobility in the alternate activation region of LRH-1 is critical to support normal activation (41).

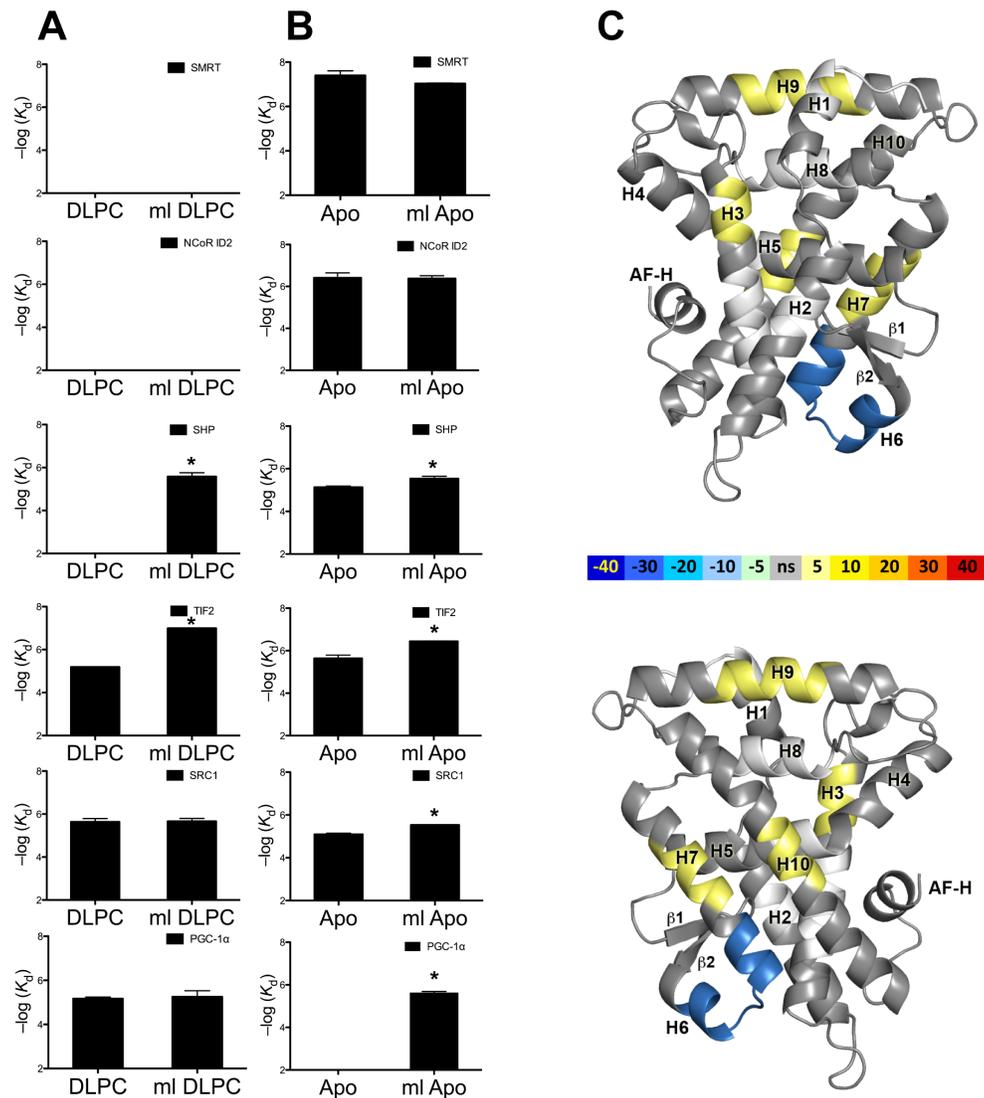


Figure 3.5 Apo mLRH-1 is more active than wt LRH-1 by enhancing stability in the alternate activation region.

A. Binding affinities for various peptides derived from coactivator and corepressor proteins expressed as $K_d \pm$ s.e.m. A. Human and mLRH-1 LBD bound to DLPC. B. Apo human and mLRH-1 LBD. C. Differential HDX-MS between apo hLRH-1 and apo ml LRH-1 LBD mapped onto PDB 1YUC. The difference in percentage of deuterium incorporation is indicated by the colored scale bar. Figures were generated in PyMOL.

Coregulator Peptide Recruitment			
	SMRT	NCoR ID2	SHP
hLRH-1 APO	0.04 +/- 0.01	0.39 +/- 0.03	7.36 +/- 0.4
mlLRH-1 APO	0.09 +/- 0.002	0.42 +/- 0.02	2.9 +/- 0.3
<i>h</i> LRH-1 DLPC	n.b.	n.b.	n.b.
mlLRH-1 DLPC	n.b.	n.b.	2.6 +/- 0.5

	PGC-1α	SRC1	TIF2
hLRH-1 APO	n.b.	8.0 +/- 0.9	2.3 +/- 0.4
mlLRH-1 APO	2.4 +/- 0.1	2.9 +/- 0.1	0.36 +/- 0.004
<i>h</i> LRH-1 DLPC	6.8 +/- 0.5	2.3 +/- 0.4	6.5 +/- 2.0
mlLRH-1 DLPC	5.6 +/- 0.2	2.2 +/- 0.3	0.1 +/- 0.01

K_d's are μ M +/- S.E.M. *n.b.* - no binding detected

Table 3.2 Coactivator Peptide Recruitment.

Discussion

A small number of orphan NRs have acquired the ability to act in a ligand independent fashion by evolving diverse structural mechanisms to stabilize their overall fold in the absence of ligand (171). Still more NRs have been crystallized in the absence of ligand despite being ligand dependent (41; 172; 173). In fact, hLRH-1 was recently crystallized without ligand showing unexpected plasticity in the LBP (41). Despite the fact that crystallization selects for empty receptor, we show that mouse and the mouse-loop variants of LRH-1 bind PL's and that PL binding alters coregulator selectivity *in vitro*. The six-amino-acid mouse-loop replacement is within LRH-1 alternate activation function surface and serves to enhance its conformational stability as evidenced by HDX. Thus, rodent LRH-1 tunes its PL-sensitivity *in vitro* by altering receptor dynamics to slightly enhance coactivator and SHP interactions in the absence of ligand. This same sequence difference also confers binding to PGC-1 α derived peptides in the absence of ligand *in vitro* which is contrasted with the PL-dependent binding of PGC-1 α in hLRH-1. Importantly, the six-amino-acid replacement does not over stabilize the LBP since apo mLRH-1 retains the ability to interact with canonical corepressors motifs *in vitro*.

It remains unclear how mLRH-1 would coordinate activating PLs such as DLPC. Substitution of Gly420 to alanine (Glu440 in mLRH-1) abrogates both PL-binding and transactivation (41). Gly420 is strictly conserved in all nonrodent LRH-1 orthologs and coordinates the lipid phosphate moiety via H-bonds with its backbone amide (40). The conformational mobility of this glycine is thought to permit PL interaction suggesting that the additional rodent-specific substitutions within alternate activation function region were likely compensatory substitutions to tolerate this drastic G420E change (41).

It is clear that remarkable plasticity exists in the NR5A fold in regions outside the canonical activation function surface. This is evidenced by the ability of mLRH-1 and mLRH-1 to effectively close the entrance to the LBP and by the dramatic structural rearrangement observed in dmFtz-F1 to fill the LBP entirely. This structural plasticity is also supported by the

observation that binding of synthetic agonists drives a ~ 3 Å constriction of the opening of the LBP versus PL-bound receptor, which is on par with that of the mLRH-1 and mLHRH-1 crystal structures (Figure 3.6) (165). The ability of synthetic agonists to induce this constriction of the pocket may be a general feature of non-PL activators.

Taken together, these results show how sequence divergence in the NR5A alternate activation function region, has differentially tuned the sensitivity of NR5A receptors to PLs. It is possible that mLRH-1 evolved a tempered PL-response as a result of differential coregulator expression or limited access to activating PLs. Importantly, these data support PL-dependent regulation for mLRH-1 strengthening conclusions generated in previous studies (37; 64) and support the use of mice as viable models for studying PL-dependent LRH-1 signaling. Future studies focused on the identification and quantification of endogenous ligands and the mechanisms that govern LRH-1's spatial and temporal access to PLs are critically needed to understand how PL-sensing via NR5A receptors regulates metabolism, lipid flux, steroid synthesis and immunity.

Finally, we show by mass spectrometry and lipid phosphorous assays that dmFtz-F1 is devoid of PLs when expressed and purified from *E. coli*, supporting the claim that that dmFtz-F1 is a truly constitutive receptor (170). The selective advantage for PL-independent Ftz-F1 signaling in arthropods is unclear. The NR5A family of NRs arose before the divergence of Placozonas and Eumatozans from a ligand activated ancestor (171). The LRH-1 ortholog in nematodes, *nhr-25*, was recently shown to bind to long chain phosphatidylinositol phosphates (PIPs) and to directly participate in the control of enzymes required to maintain the cellular lipid pool (174). The arthropod-nematode split occurred nearly 1 billion years ago (175) suggesting that PL binding was gained deep in the metazoan lineage and that the constitutive activity of Ftz-F1 was a more recent adaptation in arthropods. Recent examples in marine invertebrates show that losing ligand regulation is possible even in complex systems such as estrogen signaling which controls diverse gene programs in vertebrate biology (42). It is possible that drosophila

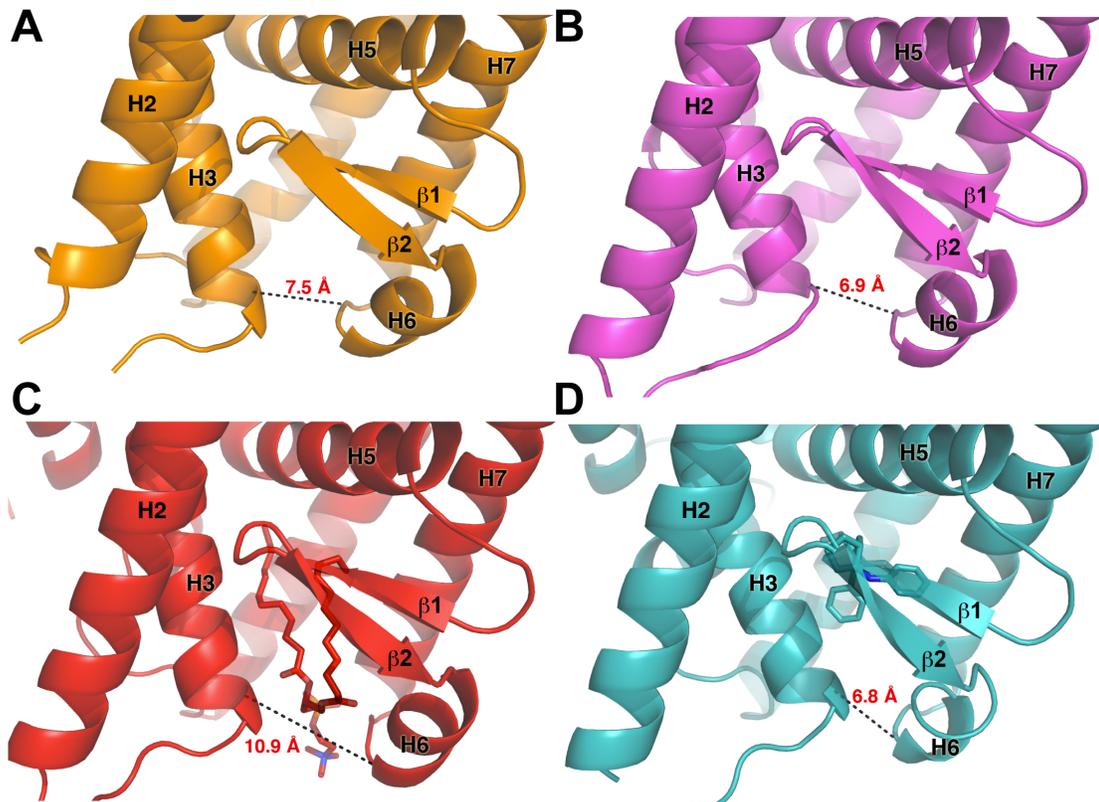


Figure 3.6 Coordination at the mouth of LRH-1's LBP is important for receptor activation.

A. mlLRH-1 B. mlLRH-1 (PDB ID: 3PK5) C. hLRH-1 bound to an activating phosphatidylcholine (PDB ID: 4DOS) D. hLRH-1 bound to a small molecule agonist (PDB ID: 3PLZ)

evolved ligand independence to escape the requirement for spatial/ temporal ligand presentation during critical developmental processes driven by this receptor.

While this *in vitro* study focuses on the isolated LRH-1 LBD, out of the context of post translational modifications and coregulator pools that often dictate activation, these results suggest that flexibility in key regions of the receptor, should be taken into account and may be capitalized upon to aid in further drug design efforts for LRH-1. For example, stabilization of the alternate activation function region may be a novel strategy for the design of LRH-1 agonists. Conversely, disruption of this region would represent a novel approach towards antagonist design.

Chapter 4: Unexpected Allosteric Network Drives Nuclear Receptor-Phospholipid Signaling

Introduction

Phospholipids (PLs) are best known for their structural role in membranes. However, PLs also play integral roles in a number of cellular signaling cascades. Although it has long been known that PLs can be used as synthesis material for some hormones, recent evidence suggests PLs are able to directly alter transcriptional patterns as ligands of some members of the nuclear receptor (NR) family ligand regulated transcription factors (176).

The human liver receptor homologue-1 (LRH-1; NR5A2) is a member of the NR5A class of nuclear receptors (NRs) that responds to PL ligands (42; 177). LRH-1 regulates the expression of genes central to embryonic development, cell cycle progression, reproduction, lipid homeostasis, local gut immune function and metabolism (43; 122; 178-182). LRH-1 is a pharmaceutical target for multiple disease states. In breast cancer, inhibition of LRH-1 selectively blocks aromatase production in breast tissue (183). LRH-1 is overexpressed in gastric tumors where it synergizes with β -catenin to drive aberrant cell cycle progression (21; 36). Oral treatment of diabetic mice with the exogenous lipid agonist dilauroylphosphatidylcholine (PC 12:0-12:0; DLPC) lowers serum lipid levels, reduces liver fat accumulation, and improves glucose tolerance in a LRH-1 dependent manner (42). As such, LRH-1 has been the subject of multiple attempts to identify small molecule modulators (165; 184-186). These attempts have been met with limited success due to our limited understanding of LRH-1's mechanism of activation.

We have shown that DLPC is able to bind directly to the LRH-1 ligand binding domain (LBD) and activate the receptor by affecting receptor dynamics in an alternate activation function region, encompassing the β -sheet-H6 region of the protein, to alter co-regulator binding preference (41). Importantly, it seems that DLPC may promote activation by relieving LRH-1 from repression by the non-canonical co-repressor NR SHP, which mimics a co-activator using the canonical LxxLL activation motif (25; 187). SHP is a robust corepressor of LRH-1's action and this interaction is central to the FXR-dependent negative feedback regulation of mouse

CYP7A1 expression (25; 188; 189). It is still unknown how LRH-1 discriminates between SHP and coactivators such as TIF2 that bind using a similar LxxLL NR box coactivator motif to recognize the active NR orientation. This is especially true when LRH-1 is bound to a non-ideal PL ligand or in its apo state. In both cases LRH-1 can bind to both SHP and coactivators. LRH-1 is unique in this manner, as other nuclear receptors are not competent to bind to coregulators in the absence of a ligand and can respond to coregulator concentration without addition of exogenous ligand (37). Additionally, it is unclear how PLs with tail lengths differing by as few as 2 carbons drive differential activation profiles.

This incomplete understanding of what dictates LRH-1's PL and coregulator selectivity limits our ability to design robust small molecule modulators for this intriguing pharmacological target. To address these questions we have generated crystal structures of the LRH-1–TIF2 complex in an apo state as well as a higher resolution structure of LRH-1 bound to *E. coli* PLs. These crystal structures when combined with previous structures of LRH-1 in various coregulator/ligand combinations have elucidated a fine tuning mechanism for receptor activation through engagement of a charge clamp. These studies have laid the groundwork for novel lipid binding assays and molecular dynamics simulations in order to determine the allosteric networks that drive phospholipid mediated NR signaling and coregulator selectivity.

Experimental Procedures

Reagents:

Chemicals were purchased from Sigma, Fisher or Avanti PLs. pMALCH10T and the vector for His tagged TEV were a gift from John Tesmer (UT Austin). pLIC_MBP and pLIC_HIS were gifts from John Sondek (UNC, Chapel Hill). Peptides were synthesized by RS Synthesis (Louisville, KY).

Protein expression and purification:

The human LRH-1 LBD (residues 291–541) was purified as described previously (116). Purified protein was dialyzed against 60 mM NaCl, 100 mM ammonium acetate (pH 7.4), 1 mM DTT, 1 mM EDTA and 2 mM CHAPS and concentrated using centrifugal filters with a 10-kDa cutoff to 5–7 mg ml⁻¹. For apo LRH-1 crystallization, purified LRH-1 LBD was incubated with 1,2-ditetracosanoyl-sn-glycero-3-phosphocholine (PC 24:0–24:0) (Avanti Polar Lipids) and (RJW101) at a final PC24:ligand:protein ratio of 20:3:1 (165). The receptor was purified away from unbound PC 24:0–24:0 and the weakly bound agonist by size exclusion chromatography, dialyzed against 60 mM NaCl, 100 mM ammonium acetate, pH 7.4, 1 mM DTT, 1 mM EDTA and 2 mM CHAPS and concentrated to 5–7 mg ml⁻¹.

Structure determination:

Both the apo LRH-1 LBD–TIF complex and the LRH-1 LBD–*E. coli* PL–TIF2 complex crystals were generated by hanging-drop vapor diffusion at 20 °C from solutions containing 1 µl of protein at 6.5 mg ml⁻¹ in complex with a peptide derived human TIF2 NR box 3 (+H3N-KENALLRYLLDKDD-CO2–) at a 1:4 molar ratio and 1 µl of the following crystal mixture: 0.7–1 M di-Sodium Malonate, 0.1 M HEPES pH 7.4, 0.5% Jeffamine ED-2001. Crystals were cryoprotected in crystallant containing 20% (v/v) glycerol and flash-frozen in liquid N₂. Data for the apo LRH-1 LBD–TIF complex were collected to 1.75 Å resolution at 100 K using a

wavelength of 0.9999 at 22-BM at the Southeast Regional Collaborative Access Team (SER-CAT) at the Advanced Photon Source and were processed and scaled with HKL2000 (190). Data for the LRH-1 LBD-*E. coli* PL-TIF2 complex were collected to 1.75 Å resolution at 100 K using a wavelength of 0.9999 at 22-ID at the Southeast Regional Collaborative Access Team (SER-CAT) at the Advanced Photon Source and were processed and scaled with HKL2000 (190). Initial phases for both structures were determined using LRH-1 PDB 1YOK as a molecular replacement search model. The structures were refined using the Phenix suite of programs, and model building was carried out in COOT (191; 192). The final model for the LRH-1-TIF2 complex contains LRH-1 residues 300–538 and TIF2 residues 742–752; it shows good geometry, with 98.4% and 1.6% of the residues in the favored and allowed regions of the Ramachandran plot, respectively. The final model for the LRH-1-*E. coli* PL-TIF2 NRbox3 complex contains LRH-1 residues 298–538 and TIF2 residues 743–750; it shows good geometry, with 98.7% and 1.3% of the residues in the favored and allowed regions of the Ramachandran plot, respectively. Data collection and refinement statistics are listed in Table 4.1.

	LRH-1 – TIF2 NRbox3	LRH-1 – <i>E. coli</i> PL – TIF2 NRbox3
Data collection		
Space group	P2 ₁ 2 ₁ 2 ₁	P2 ₁
Cell dimensions		
<i>a, b, c</i> (Å)	45.8, 65.7, 83.5	65.9, 76.9, 100.8
α, β, γ (°)	90.0, 90.0, 90.0	90.0, 95.5, 90.0
Resolution (Å)	1.75	1.75
<i>R</i> _{sym} or <i>R</i> _{merge}	6.6 (30.6)	6.6 (30.9)
<i>I</i> / σ <i>I</i>	18.99 (2.8)	12.8 (3.2)
Completeness (%)	99.4 (96.22)	92.6 (63.8)
Redundancy	3.9 (3.3)	3.6 (3.2)
Refinement		
Resolution (Å)	1.75 (1.81 – 1.75)*	1.75 (1.81 – 1.75)*
No. unique reflections	25933	6751
<i>R</i> _{work} / <i>R</i> _{free}	18.7 / 22.4	20.67 / 23.4
No. atoms		
Protein	2026	8117
Ligand/ion	42	493
Water	137	378
<i>B</i> -factors		
Protein	23.9	27.0
Ligand/ion	29.2	37.5
Water	29.4	32.7
R.m.s. deviations		
Bond lengths (Å)	0.022	0.025
Bond angles (°)	1.38	1.5

Table 4.1 Data collection and refinement statistics

*Values in parentheses are for highest-resolution shell.

Results

Structure of the apo LRH-1 LBD–TIF complex

LRH-1 is able to bind to both coactivators and corepressor proteins in the absence of a ligand. Previous studies on apo LRH-1 also showed that the receptor can access novel conformations without the added structural rigidity of ligand (41). In order to visualize the molecular perturbations necessary to bind to coactivators in its apo form, we crystallized LRH-1 LBD in the absence of a ligand bound to a fragment of the coactivator TIF2 and determined its structure to 1.75 Å (Fig. 4.1A). The receptor maintains the canonical nuclear receptor fold and, surprisingly, the LBP is constricted by 2 Å relative to the same lipid-bound structure. This is in stark contrast to the apo LRH-1 LBD- SHP NRBox 2 structure reported previously, which lacks electron density for the entire β -strands and helix 6 region (Fig 4.3 A vs E). This LBP construction which lacks the space to accommodate a PL combined with a lack of visible electron density did not support modeling of bound PL. Unlike the LBP of rodent LRH-1, the LBP constriction is not mediated by any intramolecular interactions (37), but rather the alternate activation function region is stabilized by crystallographic interactions between H6 and H9 from a symmetry mate. This structure represents a novel conformation of LRH-1 that positions its alternate activation function region in a manner that is distinct from that observed in other ligand/coregulator combinations.

Structure of the LRH-1 LBD–E. coli PL–TIF2 complex

As a control in our crystallization experiments we also included LRH-1 LBD protein that had not been stripped of its copurified *E. coli* PLs. We found that this protein, when bound to a fragment of the coactivator TIF2, also crystallized in conditions similar to the apo complex. We crystallized the LRH-1 LBD–*E. coli* PL–TIF2 complex and determined its structure to 1.75 Å (Fig. 4.2A). This represents an improved resolution as compared to the existing structures, which were both solved to 2.5 Å (40; 118), and will allow for an improved model for molecular

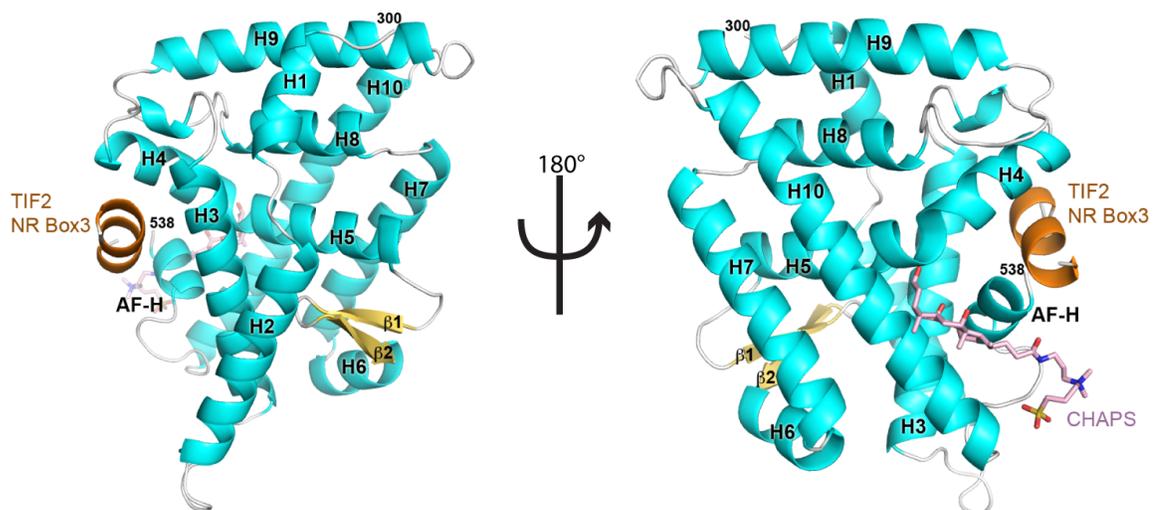


Figure 4.1 Structure of the apo LRH-1 LBD-TIF complex.

Ribbon diagram of apo LRH-LBD (α -helices, teal; β -strands, yellow) with the TIF2 NR box 3 peptide (orange). The surface bound CHAPS is depicted as sticks (C, pink; O, red; S, yellow; N, blue)

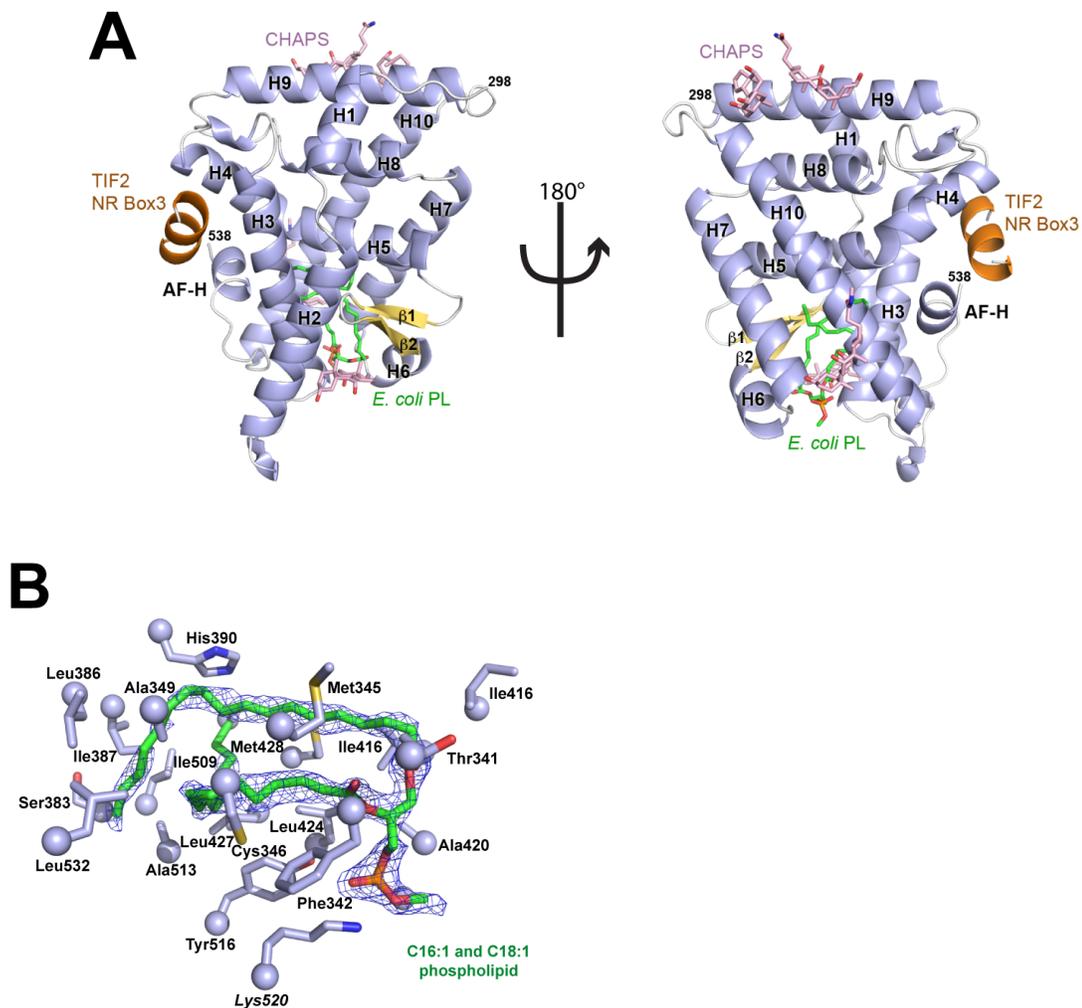


Figure 4.2 Structure of the LRH-1 LBD–*E. coli* PL–TIF2 complex.

(a) Ribbon diagram of *E. coli* PL bound LRH-1 LBD (α -helices, teal; β -strands, yellow) with the TIF2 NR box 3 peptide (orange). The bound *E. coli* PL is depicted as sticks (C, green; O, red; P, magenta). The surface bound CHAPS is depicted as sticks (C, pink; O, red; S, yellow; N, blue).

(b) $2F_o - F_c$ electron density (contoured at 1σ) for the bound C16:1 and C18:1 phospholipid observed in this structure, along with side chains lining the ligand-binding pocket of hLRH-1 that contact this ligand.

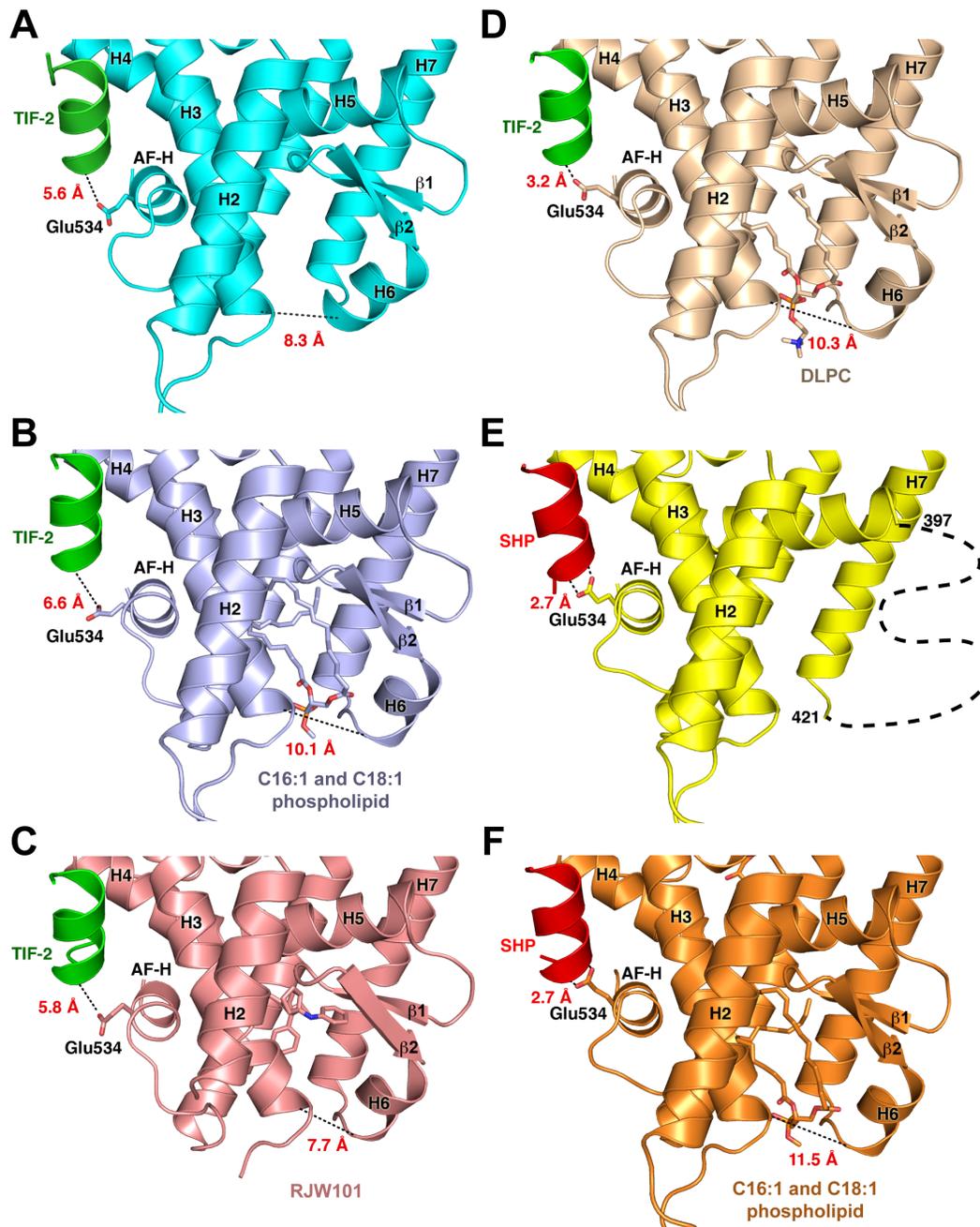


Figure 4.3 AF-2 charge clamp engagement is dictated by ligand coregulator combination.

Ligand binding pocket mouth measurements and analysis of Glu 534 charge clamp engagement for the (a) apo LRH-1–TIF2 complex, (b) LRH-1–*E. coli* PL–TIF2 complex, (c) LRH-1–small molecule agonist –TIF2 complex (PDB ID: 3PLZ), (d) LRH-1–DLPC–TIF2 complex (PDB ID: 4DOS), (e) apo LRH-1–SHP complex (PDB ID: 4DOR), and (f) LRH-1–*E. coli* PL–TIF2 complex (PDB ID: 1YUC)

dynamics and modeling studies. The structure is highly similar to the previous structures with an rmsd of 0.6 Å and maintains the previously observed *E. coli* PL binding pocket (Fig. 4.2B). Similar to previous structural work on LRH-1 in complex with the activating DLPC, the lipid acyl tails show a decrease in electron density near their terminus (41) (Fig. 4.2B). This again supports the hypothesis that LRH-1 specifically recognizes its PL ligands near the mouth of its LBP and the exact position of the acyl tails are less important than the amount of space they occupy.

Co-regulator binding interactions are altered by ligand status

The canonical model of NR activation revolves primarily around a mobile ligand sensing helix, called the activation function helix (AF-H). When a receptor is bound to an agonist the AF-H rotates to pack against the core helical bundle of the LBD. This repositioning forms a surface termed the activation function 2 (AF-2) surface that enables interaction with coactivator proteins that contain a LxxLL motif (12). This motif inserts its leucines into a pocket bounded by a charge clamp with Arg 361 on H3 and Glu 534 on the AF-H. This charge clamp is conserved across NRs and represents a general mechanism for activation (193). LRH-1 is unique because it is able to form a functional AF-2 in the absence of a ligand, which suggests that there is a finer tuned mechanism mediating LRH-1 activity between full activation and the large rearrangement necessary for the active repression by canonical corepressors. Having expanded the available ligand state/ coregulator combinations available to analyze, we were interested to compare how ligand status affects the interactions between the AF-2 and the bound coregulator peptides.

We compared coregulator binding at the AF-2 binding pockets and found that regardless of the ligand state, Arg 361 on H3, formed the expected peptide interaction. In contrast, we were surprised to find that Glu 534, on the AF-H, did not make the expected hydrogen bond with coregulator peptide under all circumstances (Fig. 4.3). The position of Glu 534 corresponds to an agreement between the ligand state and the bound coregulator peptide. In Figures 4.3 A-C, there is either no bound ligand, or a poorly activating ligand, which does not correspond to the bound

peptide that is derived from the coactivator TIF2. In these cases Glu 534 has rotated out of hydrogen binding distance of the peptide. In contrast, when LRH-1 is bound to a strong agonist such as DLPC Glu 534 does make the expected hydrogen bond with a backbone phosphate of the TIF2 peptide (Fig. 4.3D). A similar agreement is also observed in structures bound to a peptide derived from the corepressor SHP. When there is no ligand present, or a non ideal PL is bound, SHP is able to efficiently repress LRH-1 and in these cases Glu 534 rotates into a position to allow for a strong interaction with the bound peptide (Fig 4.3E-F). These observations suggest that LRH-1 has an extensive allosteric network that allows the receptor to exquisitely tune receptor activation in between full activation and full repression. This graded agonism, and the ability to form a functional AF-2 in the absence of a ligand may explain LRH-1's observed low basal activity.

Discussion

LRH-1 is an untapped target for pharmaceutical intervention metabolic and neoplastic disease states (43; 122; 178). An incomplete understanding of what dictates LRH-1's PL and coregulator selectivity limits our ability to design robust small molecule modulators for this intriguing pharmacological target.

Previous structural studies of LRH-1 revealed that PL binding promotes activation in a way that is distinct from other NRs. These studies showed that activation of LRH-1 relied on PLs to provide additional intramolecular contacts at the mouth of the LBP between helices 2 and 3, 6 and 7 and 10 and AF-H (41) rather than making specific ligand contacts deep in the LBP. This work along with other studies on LRH-1 did little to explain how closely related PL species could coordinate with opposing coregulators to elicit differing activation responses from the protein. In this study we aimed to understand how receptor dynamics, coordinated motions, and allosteric networks mediate LRH-1's different responses to PLs and coregulators.

Similar to previous observations, LRH-1 is able to access a novel conformation when not bound to a ligand. While bound to the coactivator TIF2, LRH-1 positions its alternate activation

function in a manner that is distinct from other ligand/coregulator combinations (Fig. 4.3). This alternate activation function position that is neither propped open by a PL ligand or artificially constricted by a bound small molecule may be beneficial for *in silico* small molecule screening.

LRH-1 is known to have basal activity in the absence of external stimuli (37). This observation suggests that the receptor has a mechanism for fine tuning its activity in between full activation and full repression. Indeed, LRH-1 is competent to bind both coactivators and corepressors while apo (41). Comparison of the AF-2 region multiple coregulator/ligand combinations of LRH-1 shows that LRH-1 is able to form additional contacts with the coregulator through completion of the canonical NR charge clamp when both the ligand state and ligand state correspond to the same activation/repression state. This data suggests that complex allosteric networks relay ligand status to the AF-2 to fine tune receptor activation. These studies have laid the foundation for ongoing lipid binding assays and molecular dynamics simulations. When combined, these experiments have the potential to reveal the range of lipids LRH-1 is able to bind to and may also determine the allosteric networks that allow LRH-1 to fine tune its coregulator preferences that dictate transactivation versus transrepression.

Chapter 5: Discussion

Phospholipids as signaling molecules

The concept of PLs acting as more than structural elements is a relatively new concept. However there are now multiple example of PLs acting as signaling molecules that are able to alter complex gene programs by acting as ligands for NRs (42; 81; 91; 194; 195). Modulating PL-driven transcriptional pathways remains an untapped therapeutic opportunity. PLs, much like the hormones of other NRs, appear to act as a folding nucleus by occupying a very large hydrophobic pocket at the core of the receptor. The vast diversity of lipids and the potential for lipid modification complicate the molecular mechanism of PL activation. Understanding the binding modes of PLs to PL responsive transcription factors may assist in the design of small molecules targeting these receptors.

LRH-1 is an untapped pharmacological target

LRH-1 is a member of the NR superfamily that is known to directly bind to and be activated by PLs. LRH-1 regulates the expression of genes central to embryonic development, cell cycle progression, reproduction, lipid homeostasis, local gut immune function and metabolism (43; 122; 178-182). As a pharmacological target for metabolic and neoplastic disease, LRH-1 remains an untapped target due to limited success in finding potent small molecule modulators. This is due in part to a limited understanding of LRH-1's activation.

LRH-1's mode of activation is unique among NRs

Structural and biochemical experiments performed with LRH-1 bound to DLPC help to elucidate DLPC's mode of activation. Similar to observations in other NRs, DLPC promotes receptor activation by simultaneously enhancing coactivator binding and disfavoring corepressor binding. When compared to *E. coli* PLs, DLPC causes a slight decrease in LRH-1 stability. It is possible that when bound to *E. coli* PLs, considered poor agonists, the receptor is artificially over stabilized and DLPC may promote activation by providing stability at the LBP mouth while its

shorter acyl tails promote flexibility at the top of the LBP and allow the required flexibility deep in the LBP.

LRH-1 was previously thought to be a rigid protein that is unresponsive to ligands, making it a poor pharmacological target. The data presented in Chapter 2 show that, in contrast, LRH-1 is a dynamic protein that is able to alter its structure to accommodate multiple PLs and a wide range of comodulator protein partners. Unlike other NRs, which recognize their ligands deep in the core of the NR protein fold, LRH-1 binds to PLs far from this canonical NR ligand recognition site. LRH-1 instead relies interactions between a bound PL and the H2-3, H6-7 and H10 regions that specifically stabilize the mouth of the LBD. This stabilization includes the β -sheet-helix 6 region of LRH-1. The data unexpectedly showed that flexibility in this region is critical for receptor activation. This surface appears to act as a novel alternate activation function region of the receptor.

Sequence Divergence throughout evolution has altered the structural mechanisms driving LRH-1 activation

Sequence analysis coupled with functional and structural studies of LRH-1 have shown that sequence variation among LRH-1 orthologs may effect PL binding and regulation (37; 40; 170). While mLRH-1 was first hypothesized to be ligand independent (37), it is now known that both human hLRH-1 and mouse LRH-1 respond to DLPC to modulate lipid and glucose homeostasis (42). Mouse and Drosophila orthologs contain divergent sequences in the alternate activation function region and these alterations may alter PL-driven activation (37; 40; 170). Structural evidence suggests that these sequence differences in mLRH-1 and Drosophila FTZ-f1 (dmFTZ-f1) confer at least partial ligand independence, making them poor models for hLRH-1 studies; however, the mechanisms of ligand independence remain untested (37; 40; 170). Our data show that the recent evolutionary divergence of the mLRH-1 stabilizes the active conformation in the absence of ligand, yet does not abrogate PL-dependent activation. We also show by mass spectrometry and biochemical assays that FTZ-f1 is incapable of PL binding. This

work provides a structural mechanism for the differential tuning of PL sensitivity in NR5A orthologs. Importantly, this work supports the use of mice as viable therapeutic models for LRH-1 dependent diseases.

LRH-1 structures as invaluable tools to understand the detailed molecular mechanism driving receptor transactivation and transrepression.

Chapters 2-3 present data showing the molecular determinants of LRH-1 activation and highlight the importance of the newly identified alternate activation function region. This, however does not address two very important questions: How does LRH-1 recognize the difference between very similar PLs that have different activation effects? And how does LRH-1 select between closely related coregulators with opposing functions, like SHP and TIF2. For example, a PC with saturated acyl tail lengths of 12 is activating while a PC with saturated acyl tail lengths of 16 is not (42). Our data show that an ideal PL such as DLPC can relieve LRH-1 from SHP repression *in vitro*, but a non ideal ligand does not (41). The newly generated crystal structures presented in Chapter 4 present a novel conformation of apo LRH-1 bound to a fragment of the coactivator TIF and an improved resolution structure of LRH-1 bound to a PL from *E. coli* and a fragment of the coactivator TIF. These structures, combined with previously generated crystal structures, provide the critical mass of structural data required to launch mechanistic modeling studies aimed at deeply understanding how PL status drives coregulator selectivity. These ongoing experiments have already generated exciting results, showing perfect agreement between PL status and coregulator preference (e.g. activating lipids show strong characteristic interaction networks with coactivators while these same networks are in place with corepressors only in the absence of activating lipids).

Future Directions

Despite progress identifying PLs as exogenous ligands for LRH-1, the endogenous ligand has yet to be identified. It may be the case that LRH-1 binds to a wide range of readily available PLs in the nucleus, but is only activated by a select few signaling PLs. Alternatively, LRH-1 may

respond to a bolus of PLs ingested in a meal. A comprehensive study of the PLs that LRH-1 is able to bind to may offer some clues about the types of PLs LRH-1 may sample in cells. The orphan status of LRH-1 has hampered both an understanding of LRH-1's biology and structural studies of LRH-1 in its endogenous activation state. Ongoing work in the lab suggests that LRH-1 selects its lipid ligand based solely on the hydrophobic tails. This suggests that other factors, such as soluble lipid shuttling proteins or enzymes may direct PLs to the LRH-1 LBP for activation. Future studies will test the role of PC shuttling proteins known to traffic to the nucleus on PL binding. For example, preliminary data suggests that phosphatidylcholine transport protein (PC-TP) plays a role in stimulating LRH-1 transactivation via direct interactions. Additional research focusing on identifying the PLs or ligands that endogenously activate LRH-1 may dramatically advance the ease of studying LRH-1 biology.

In addition to binding to canonical coregulators, LRH-1's activity can also be altered by a range of other proteins (21; 34; 151; 196-198). LRH-1 has been shown to interact with the orphan nuclear receptor DAX and β -catenin at novel interaction interfaces (115; 151). A deeper understanding of these interaction surfaces may develop LRH-1 biology and offer novel targets for selective LRH-1 modulating compounds. An intriguing feature of LRH-1's PL binding mode is that the PL headgroup is displayed on the surface of the receptor (41; 177). Recent research has indicated that this display maybe be used as an interaction surface with other proteins (177). A similar mechanism may be involved when DLPC is bound to the receptor, as a new positive charge is displayed on the surface of the receptor and may be responsible for coordinating a yet-to-be discovered protein – protein interaction. Using agonist PLs for differential proteomic pull down experiments may elucidate new ligand dependent binding partners.

Concluding Remarks

Taken together, the data presented here greatly expands our mechanistic understanding of LRH-1's ability to respond to PL ligands. Previous hypotheses suggested that LRH-1 nonspecifically binds and folds around range of PLs with its transactivation sensitive not to ligand

but solely the pool of available coregulator proteins. The data presented here suggest that, in contrast to previous hypotheses, LRH-1 is able to exist in multiple PL-bound and apo states that have differing binding preferences for coactivators vs. corepressors (Fig. 5.1). These states range from apo and specifically repressed by canonical corepressors such as SMRT and NCoR to full activation when bound to DLPC, which selects for binding to coactivators (Fig. 5.1). This dynamic response to PLs may permit LRH-1 to generate differing transcriptional responses to changing PL pools. Additionally, these experiments support the use of mice as viable therapeutic models for LRH-1 dependent diseases, but not *Drosophila*. Taken together, this suggests that small molecules, like activating PLs, maybe be able to specifically bind to LRH-1 and alter its activation state by manipulating receptor dynamics to favor either coactivator or corepressor binding.

LRH-1 remains a highly interesting topic of study. Biologically, LRH-1 remains a key player in development, metabolism and cancer. As a regulator of pluripotency, an agonist of LRH-1 has the potential to enhance cellular reprogramming. As a key regulator of bile acid, glucose and lipid homeostasis, an agonist of LRH-1 has the potential to be a non-insulin treatment strategy of diabetes and pre-diabetes. As a regulator of steroidogenesis in breast cancer, a LRH-1 antagonist has the potential to be a non-estrogenic treatment of estrogen receptor positive breast cancer. LRH-1 remains an untapped pharmaceutical target. A continued understanding of its normal and aberrant biological roles along with further development of LRH-1 modulating compounds offers the possibility of exciting new therapeutics and research tools.

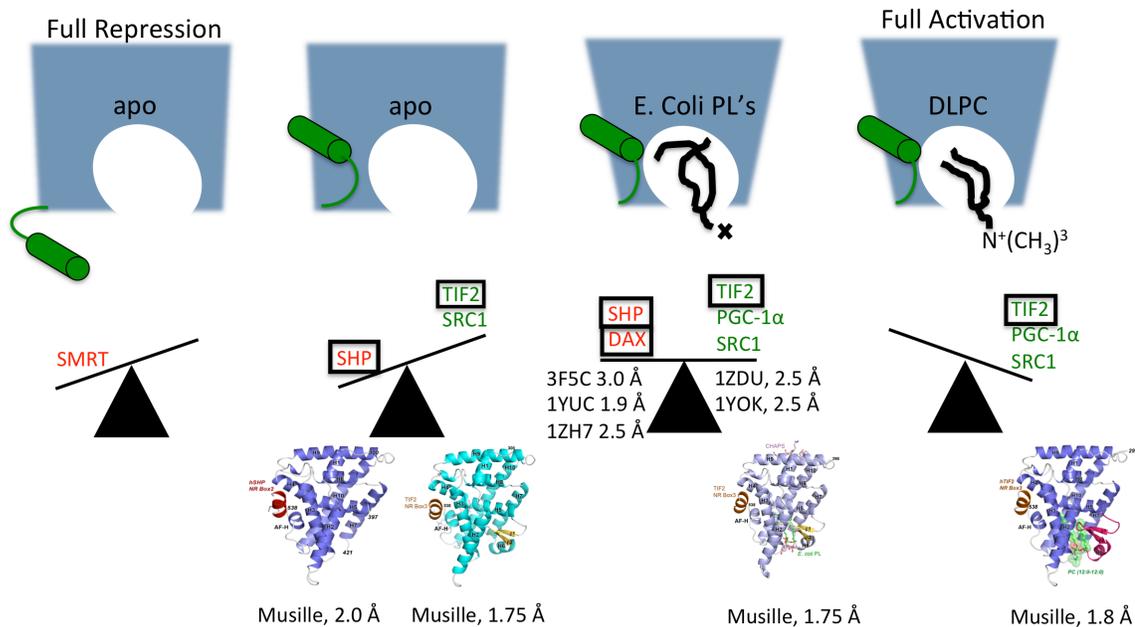


Figure 5.1 LRH-1 Activation States Model

Representations of the LRH-1 LBD shown in blue with the AF-H depicted as a green tube in either the inactive or active orientation. Bound PL's are drawn in black. Corepressors are printed in red and coactivators in green. Teeter-totter depicts the relative preference for coactivators vs. corepressors. Boxes denote a coregulator peptide complex structure that has been solved. Published structures are denoted by PDB ID and resolution. Structures presented in this thesis are pictured and labeled by resolution.

Appendix: Structure of Glycerol Dehydrogenase From *Serratia*⁴

⁴This chapter has been slightly modified from the published manuscript: Musille PM, Ortlund EA. (2013) Structure of glycerol dehydrogenase from *Serratia*. *Acta Crystallographica Section F: Structural Biology Communications* 70

Introduction

Structure determination of unidentified proteins can be a challenging endeavour. Often, a search of unit cell constants in the PDB will allow for identification of commonly crystallized proteins derived from the host expression organism (i.e. *E. coli*). More advanced searches have enabled the identification of unidentified protein crystals by performing molecular replacement on up to 100,000 protein domains (199). Often these methods may not be feasible due to the computing power needed for brute force molecular replacement. Experimental phasing techniques do not require a known sequence for phase determination, but in both cases knowing the sequence of the protein contained in the crystals can be of immense help for efficient structure solution.

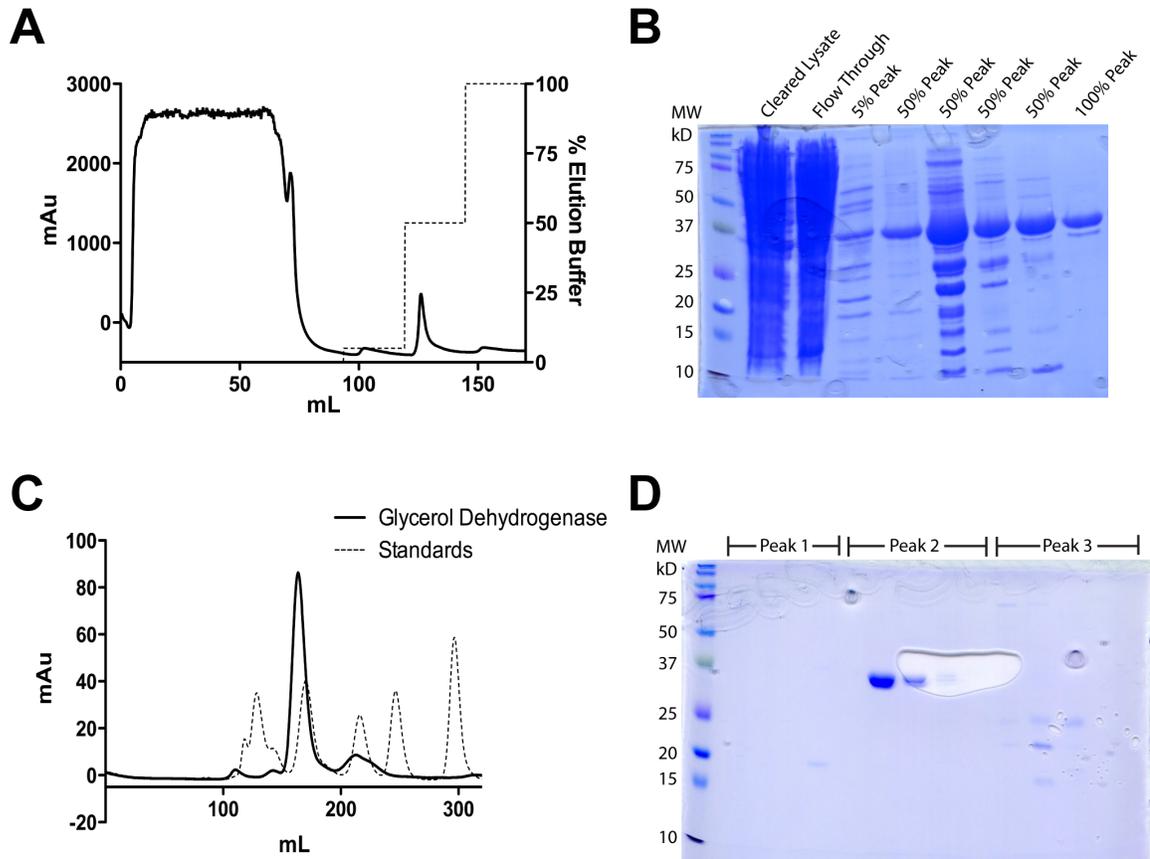
When structure solution does not progress as smoothly as planned it is often tempting, and wise, to double check the identity of the protein contained within the crystals that were exposed to X-rays. There are multiple examples of purification and crystallization of contaminants (200-203) and successful identification of the protein often is the key to successful structure solution.

In this study, we crystallized an off-target protein derived from an unknown source of contamination that diffracted to 1.9 Å. Protein directly from hanging drop vapour diffusion experiments was collected and subjected to sequencing by mass spectroscopy. This enabled the structure determination of the metabolic enzyme glycerol dehydrogenase (GDH) from the bacterial contaminate with a peptide sequence that most closely matches GDH within the genus *Serratia*.

Experimental Procedures

Protein Production and Purification:

A human target protein (predicted MW of ~40 kD) was cloned into the pLIC_His pET vector kindly provided by John Sondek (UNC-Chapel Hill, NC), which contains a N-terminal 6 x His tag followed by a TEV protease cleavage site preceding the target gene. The resulting plasmid was sequence confirmed showing insertion of the human target gene in frame with N-terminal tag and TEV cleavage site. In-house terrific broth (TB) media was prepared as a 10x concentrated stock, autoclaved, and diluted in filtered water. A flask containing 100 mL of TB supplemented with 100 µg/mL ampicillin was inoculated with a single colony containing *E. coli* BL21(DE3)pLysS cells transformed with the expression plasmid. This flask was maintained at 37 °C with vigorous agitation overnight. Flasks containing 1.3 L of TB supplemented with 100 µg/mL ampicillin were inoculated with 1 %v/v of this overnight starter culture and were maintained at 37 °C with vigorous agitation. At mid log phase, IPTG was added to a final concentration of 1 mM and the temperature was lowered to 23 °C for 24 hours. Cells were harvested by centrifugation and formed a visibly red tinted pellet. The bacteria were lysed by sonication and cleared by centrifugation. Cleared lysates were passed over a Ni²⁺ affinity column in a buffer containing 150 mM NaCl, 20 mM Tris pH 7.4, 5% glycerol and 20 mM imidazole. Trapped protein was eluted using a stepped protocol with the same buffer with the imidazole concentration increased to 250 mM (Appendix Figure 1a). The 50% and 100% peaks consisted primarily of a protein that migrated near the 37 kD standard by SDS-PAGE analysis (Appendix Figure 1b). These peaks were pooled and further purified by size exclusion chromatography (Appendix Figure 1c-d). Protein eluted at a volume suggesting a mass of ~160 kDa which was consistent with the expected mass of a tetramer of the target protein. The final protein yield was ~1 mg per liter of culture.

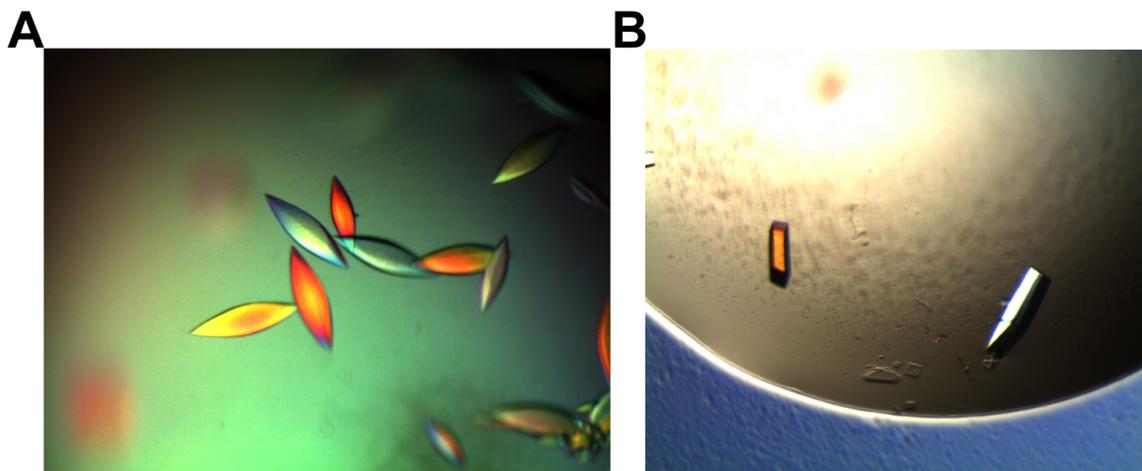


Appendix Figure 1 Protein Purification

(A) Ni²⁺ affinity chromatogram showing steps of 5%, 50%, and 100% of an elution buffer containing 250mM imidazole and (B) SDS-PAGE analysis of the resulting peaks. (C) Size exclusion chromatography and (D) SDS-PAGE analysis of the resulting peaks.

Crystallization:

Purified protein was concentrated to 6 mg/mL and dialyzed against 150 mM ammonium acetate, 50 mM sodium chloride, 20 mM tris pH 7.4, and 5% glycerol. This protein solution was used to set low volume (0.2 μ L protein solution, 0.2 μ L crystallant) sitting drop crystallization trials with a PHOENIX robot (Art Robbins Instruments, Sunnyvale, CA) using commercially available screens. Initial crystals were discovered in well G4 of The PEGs Suite (Qiagen, Germantown, MD). Larger crystals were obtained by mixing 1 μ L of protein solution with 1 μ L of crystallant containing 5-10% PEG 3350, 0.2 M calcium acetate, and 4% glycerol in a hanging drop vapour diffusion experiment. These crystals exhibited a prolate spheroid-like morphology that lacked defined edges and faces (Appendix Figure 2a). Manipulation of these fragile crystals proved difficult and diffraction limits between crystals varied. Crystal robustness and diffraction reproducibility were improved by chemical crosslinking with glutaraldehyde using 2 microliters of 25% glutaraldehyde that was exposed to the crystals by vapor diffusion for one hour (204). Diffraction of these crystals was limited to ~ 3 Å and were highly sensitive to radiation damage. To improve crystal morphology and diffraction limits we screened crystallization additives. Addition of 4% 2,2,2 trifluoroethanol significantly altered crystal morphology, generating crystals with a cuboid morphology with clearly defined edges and faces (Appendix Figure 2b). This crystal form withstood manipulation and initially diffracted to ~ 1.6 Å. Crystals were cryoprotected with crystallant solution plus 20% glycerol before flash cooling in liquid nitrogen.



Appendix Figure 2 Crystals of *Serratia* glycerol dehydrogenase

Crystals of *Serratia* glycerol dehydrogenase measuring approximately 75 x 75 x 200 microns before (A) and after optimization with 2,2,2-trifluoroethanol measuring approximately 50 x 50 x 150 microns (B).

Diffraction data collection:

One hundred and fifty degrees of data was collected in 0.5° oscillation frames from a single crystal. Data were collected at 100 K using a MARMOSAIC 300 mm CCD at Southeast Regional Collaborative Access Team (SER-CAT) 22-ID beamline at the Advanced Photon Source, Argonne National Laboratory. Diffraction data to 1.9 Å resolution were integrated and scaled using the program HKL-2000 (HKL Research, Inc.) (Appendix Table 1).

Structure Determination:

Molecular replacement using search models generated from homologous proteins of our target protein was unsuccessful. A search for similar unit cell parameters in the PDB also failed to provide a molecular replacement search model. Since we could not identify an appropriate model for molecular replacement, we performed extensive heavy atom and halide screening to facilitate *de novo* phasing. Osmium chloride preserved crystal quality and provided suitable signal for SAD experiments. This allowed us to calculate initial phases, which resulted in an interpretable map with visible density for side chains. Attempts to build the target human protein sequence into this map were unsuccessful suggesting that the target protein was not present in these crystals.

To expedite the identification of the crystallized protein, a drop from the hanging drop experiment was analysed by SDS-PAGE and the single visible band was excised and subjected to tryptic digestion. The resulting peptides were analysed by reverse-phase liquid chromatography coupled with LC-MS/MS on a LTQ Orbitrap Hybrid Mass Spectrometer (Thermo Scientific) as previously described (147). A search for peptides matching a human sequence database confirmed that the target protein was not present. To our surprise, a subsequent search of an *E. coli* sequence database also returned no clear matches. We then expanded our search to include all known

PDB code	4MCA
Data Collection	
Crystal system, space group	<i>I</i> 422
<i>a</i> , <i>b</i> , <i>c</i> (Å)	117.5, 117.5, 259.9
α , β , γ (°)	90, 90, 90
Beamline	APS beamline 22-ID, SER-CAT
Wavelength (Å)	1.0
Resolution range (Å)	50–1.90 (1.97–1.90)
No. of unique reflections	71425 (3706)
Completeness (%)	91.6 (52.3) ¹
Multiplicity	8.7 (4.0)
$\langle I/\sigma(I) \rangle$	20.2 (3)
R_{merge}	0.1 (0.41)
Refinement	
Refinement software	<i>REFMAC</i> 5.7.0032
Resolution range (Å)	39.6–1.90 (1.94–1.90)
Completeness (%)	91.5 (52.3)
No. of reflections, working set	65855 (2473)
No. of reflections, test set	3341 (110)
Final R_{cryst}	0.177 (0.258)
Final R_{free}	0.215 (0.298)
No. of non-H atoms	
Protein	5397
Ion	6
Ligand	24
Water	365
Total	5792
R.m.s. deviations	
Bonds (Å)	0.020
Angles (°)	1.863
Overall average <i>B</i> factor (Å ²)	30.0
Ramachandran plot analysis	
Most favoured regions (%)	98
Additionally allowed regions (%)	2
Disallowed regions (%)	0

Appendix Table 1 Crystallographic Statistics

Values for the outer shell are given in parentheses.

The data collected were anisotropic and resulted in low completeness in the 2 highest resolution shells. This data was include in spite of the low completeness due to an $I/\sigma(I)$ of 3 in the high

shell, a multiplicity of 4, and an R_{merge} of 0.41. Additionally, including this data had a positive influence on the resulting maps.

protein sequences, which provided peptide matches to glycerol dehydrogenase from the bacterial genus *Serratia* (Appendix Figure 3).

After protein identification, structure solution was readily achieved using PDBID: 1JPU (Crystal Structure of Glycerol Dehydrogenase) (205) as a search model for molecular replacement using the program Phaser (206). Subsequent cycles of model building with Coot (207) were alternated with crystallographic refinement using REFMAC 5.7.0032 (139; 191). The final model has no geometric outliers and a MolProbity score of 1.27 placing it in the 99th percentile among structures of comparable resolution.

Miscellaneous:

Figures were prepared with Geneious (Biomatters, Ltd.), PyMOL (Schrödinger, LLC) and LigPlot+ (150). Model validation was performed with MolProbity (143). Coordinates and structure factors have been deposited in the PDB (<http://www.pdb.org>) as entry 4MCA.

A	<i>Serratia proteamaculans</i>	B	<i>Serratia odorifera</i>
1	MLRIIQSPGK YIQGANALAA VGEYAKALAD HYFVIADDFV MKLTGDTLMG	1	MLRIIQSPGK YIQGANALVA VGOYAKALAD HYFVIADDFV MKLAGDTLMG
51	SLHQHGKHH ACVFNGECCH KEIDRLQDEL KAHGCRGVIG VGGKTLDTA	51	SLQHGKHH ACLFNGECCH KEIDRLGREL KAHGCRGVIG VGGKTLDTA
101	KAAAHYQHLF VVLIPTIAST DAPTSALSVI YTEQGEFDEY LIHTSNPDMV	101	KAAAHYQQLP VVLIPTIAST DAPTSALSVI YTEQGEFAEY LIYPRNPDV
151	IMDSSIATA PVRLVAGMG DALSTVFEAQ ACFDAQATSM AGGKSTLAAL	151	VMDSAIATA PVRLVAGMG DALSTVFEAQ ACFDAQATSM AGGKSTLAAL
201	SLARLCYETL LAEGVKAKLA VEAGVVTEAV ERRIEANTYL SGIGFESSGL	201	SLARLCYETL LAEGVKAKLA VEAGVVTEAV ERRIEANTYL SGIGFESSGL
251	AAAHAIHNGF TVLEBCHLY HGEKVAFTL SQLVLQSSM AQIETVLDFC	251	AAAHAIHNGF TVLEBCHLY HGEKVAFTL AQLVLQSSM AQIETVLDFC
301	QRLGLPITLA QMGVSGDPAE KIMAVABASC AAGETTHNMP FEVTPASVQA	301	HRIGLPITLA EMCVSGDAAE KIMAVAQASC AAGETIHNMP FKVTPASVQA
351	AILTADRLGN AWLQGHSS	351	AILTADRLGS AWLQQHQR

Appendix Figure 3 Peptide Identification

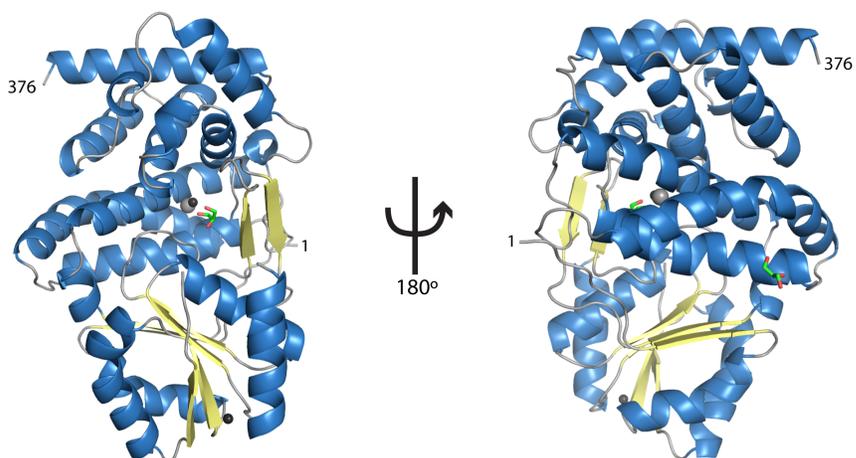
Peptides identified by LC-MS/MS, depicted as red letters within the sequences of glycerol dehydrogenase from (A) *Serratia proteamaculans* and (B) *Serratia odorifera*.

Results and Discussion

Like many labs interested in protein structure – function relationships, we set out to express, purify and structurally characterize a human protein of interest. We cloned the target gene to a popular expression vector containing a hexa-histidine tag for convenient *purification*. Initial purification utilizing Ni^{2+} -affinity chromatography resulted in a protein that was near the expected molecular weight as analysed by SDS-PAGE and formed what appeared to be the expected multimer when analysed by size exclusion chromatography. After purification, initial crystals were identified with limited screening efforts and both the morphology and diffraction limit of initial crystals were vastly improved by additive screening.

Initial failed attempts at molecular replacement were attributed to potentially poor search models due to a lack of closely related structures in the PDB. Our efforts then pivoted to *de novo* phasing techniques. Initial phases were calculated from a SAD dataset utilizing the signal from bound osmium ions and the resulting map suggested that our target protein's sequence was not present. Mass spectroscopy analysis of a crystallization drop ruled out that we had crystallized our target human protein: rather, it identified the protein as a probable glycerol dehydrogenase from the bacterial genus *Serratia*.

Identification of the protein contained in the crystals as probable member of the glycerol dehydrogenase family from *Serratia* allowed for identification of homologous crystal structures in the PDB suitable for molecular replacement search models. The final model contains 2 glycerol dehydrogenase monomers in the asymmetric unit with each monomer containing 2 glycerols, 2 zincs, and a sodium ion (Appendix Figure 4). While the protein does not contain more than two consecutive histidine residues in its primary sequence, it displays 11 surface exposed histidines, with H31, H59, H60, H83 forming a cluster on the surface of the protein. These histidine residues are within van der Waals contact distance and H31 and C85 coordinate a



Appendix Figure 4 Overall Structure

Overall structure of *Serratia* glycerol dehydrogenase represented by a ribbon diagram with α -helices colored blue and β -strands, yellow. Glycerol is depicted as sticks (C, green; O, red) and bound zinc is depicted as black sphere and bound sodium is depicted as a grey sphere.

zinc ion at this site. We hypothesize that this zinc cluster may have contributed to this glycerol dehydrogenase's affinity for Ni^{2+} media. The probable sequence of the protein, as determined by interpretation of electron density, has > 95% sequence identity to *Serratia proteamaculans*, *Serratia odorifera*, *Serratia plymuthica*, *Serratia marcescens*, and *Serratia liquefaciens*, but is not identical to any (Figure 5a). At only 3 positions does the electron density suggest an amino acid other than one found in close homologs: leucine 33, valine 154, and valine 319 (Appendix Figures 5b-d respectively). The next closest match identified in the non-redundant NCBI protein database is GDH from *Yersinia Intermedia*, which shows significantly less identity at 78 %. Thus, the protein we have purified is most likely either from one of the known *Serratia* species with naturally occurring variation or is from an unsequenced species within the genus.

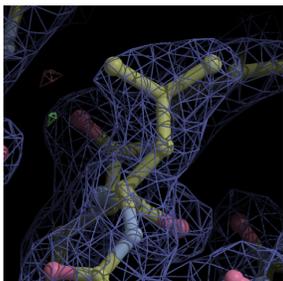
Glycerol dehydrogenase is the enzyme responsible for the oxidation of glycerol to dihydroxyacetone. This permits its entrance to the glycolytic pathway (208; 209). Thus, many organisms express glycerol dehydrogenase under anaerobic conditions to utilize glycerol as an energy source (209; 210). This oxidation requires concurrent reduction of NAD^+ to NADH (209) along with the presence of an active site zinc responsible for coordinating glycerol in the enzyme's active site (205; 211). It is plausible that in our culture conditions, glycerol dehydrogenase expression was increased to take advantage of the 10% glycerol supplementation in TB media.

Overall, the structure is highly similar to previously published structures of glycerol dehydrogenase with an R.M.S.D. of 1.1 Å vs. PDBID: 1JQA (*Bacillus stearothermophilus* Glycerol Dehydrogenase Complex with Glycerol) (205) (Appendix Figure 6a) and 0.96 Å vs. PDBID: 1KQ3 (Crystal Structure of a Glycerol Dehydrogenase (Tm0423) from *Thermotoga maritima* at 1.5 Å Resolution) (212) (Appendix Figure 6b) as calculated by matchmaker in the Chimera software package (213). The structure maintains a two-domain architecture separated

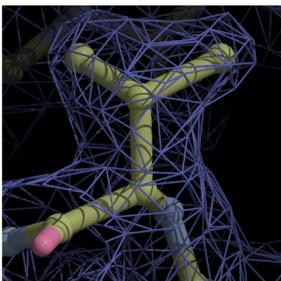
A

4MCA	1	10	20	30	40	50	60	70
4MCA	MLRITIQSPGKYIQGANALAAVGYAKSLADHYLVIADDFVMKLAGDTLMGSLQQHGVKHHAAALFNGECCHKEID							
<i>S. odorifera</i>V.....A.....F.....C.....							
<i>S. plymuthica</i>A.....F.....C.....							
<i>S. marcescens</i>S.....E.....F.....Q.....R.....R.....							
<i>S. liquefaciens</i>E.....N.....F.....CV.....							
<i>S. proteamaculans</i>E.....A.....F.....T.....H.....CV.....							
4MCA	80	90	100	110	120	130	140	
4MCA	RLGRRLKAHGCRGVIQVGGGKFLDTAKAIAHYQQLPVVLIPTIASTDAPTSALSVIYTEQGEFAEYLIYPRNPD							
<i>S. odorifera</i>							
<i>S. plymuthica</i>							
<i>S. marcescens</i>QT.....V.....R.....							
<i>S. liquefaciens</i>I.....Y.....H.....							
<i>S. proteamaculans</i>Q.....H.....D.....HTS.....							
4MCA	150	160	170	180	190	200	210	220
4MCA	MVMVDVAIIAKAPVRLVAGMGDALSTYFEAQAQCFDAQATSMAGGKSTIAALSLARLCYDTLLAEGVKAKLAVE							
<i>S. odorifera</i>S.....E.....							
<i>S. plymuthica</i>S.....E.....							
<i>S. marcescens</i>S.....E.....							
<i>S. liquefaciens</i>S.....E.....							
<i>S. proteamaculans</i>I.....SS.....E.....							
4MCA	230	240	250	260	270	280	290	
4MCA	AGVVTEAVERIIEANTYLSGIGFESSGLAAAHAIHNGFTVLEECHHLYHGEKVAFGTLAQLVLQNSPMAQIETV							
<i>S. odorifera</i>S.....S.....							
<i>S. plymuthica</i>S.....S.....							
<i>S. marcescens</i>G.....S.....S.....E.....							
<i>S. liquefaciens</i>S.....S.....S.....							
<i>S. proteamaculans</i>S.....S.....S.....							
4MCA	300	310	320	330	340	350	360	367
4MCA	LAFCHRIGLPI TLAE MGVSGDAVEKIMAVAQASCAAGETIHNMPFKVTFAGVQAATL TADRLGSAWLQQHQ							
<i>S. odorifera</i>	.D.....V.....A.....S.....R							
<i>S. plymuthica</i>	.D.....A.....V.....E.....T.....R							
<i>S. marcescens</i>	.D.....Q.....T.....A.....M.....E.....A.....R							
<i>S. liquefaciens</i>	.G.....QQQL.....Q.....A.....E.....S.....N.....R.....SS							
<i>S. proteamaculans</i>	.D.....Q.....L.....Q.....P.....A.....E.....T.....E.....S.....N.....SS							

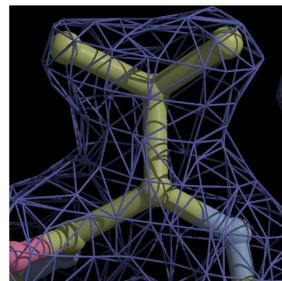
B



C

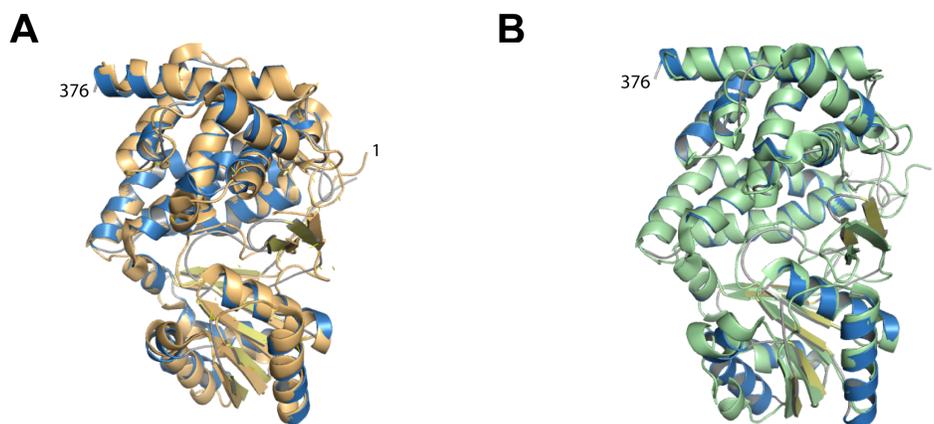


D



Appendix Figure 5 Sequence Confirmation

(A) Sequence of 4MCA compared by alignment to the 5 closest glycerol dehydrogenase sequence matches from various *Serratia* species. (B-D) $2F_o-F_c$ electron density contoured to 1σ (blue) and F_o-F_c contoured at 2 (green) and -2 (red) σ at amino acid (L33, V154, and V319 respectively) where electron density was used to interpret a sequence that differed from close homologs.



Appendix Figure 6 Superpositions of *Serratia* glycerol dehydrogenase

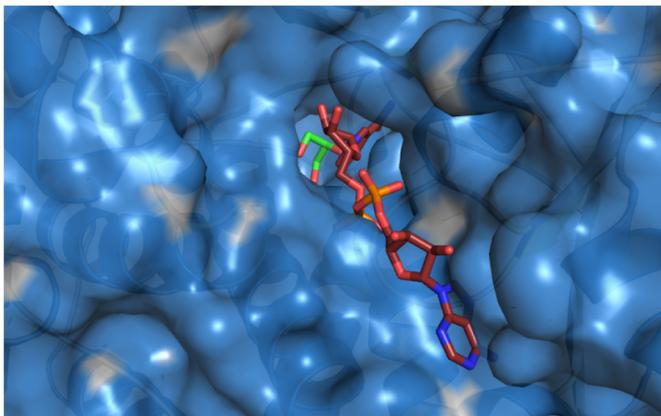
Superposition of *Serratia* glycerol dehydrogenase (blue) with (A) *Bacillus stearothermophilus* glycerol dehydrogenase (copper) with a R.M.S.D. of 1.14 Å and *Thermotoga maritime* glycerol dehydrogenase (green) with a R.M.S.D. of 0.96 Å as calculated by matchmaker in the Chimera software package (213).

by a deep cleft that has been observed to be the NAD binding site in other crystal structures (205). Modelling of a NAD molecule, by a superposition of PDBID: 1JQ5 (205) onto the *Serratia* structure, predicts that binding is conserved at this site maintaining all predicted H-bonding and hydrophobic interactions (Appendix Figure 7). The active site is highly conserved and contains a zinc ion coordinated by two histidines and an aspartate. This zinc is responsible for coordinating the bound glycerol molecule (Appendix Figure 8).

The genus *Serratia* contains bacteria that are Gram-negative, rod-shaped, and facultatively anaerobic (214). *Serratia* can often be visually identified as some strains have a characteristic red pigment and is often found as a red biofilm in bathrooms and in nature (214). It is considered an opportunistic pathogen and is known to infect respiratory and urinary tracts. Infections are most often a result of *Serratia marcescens* (214).

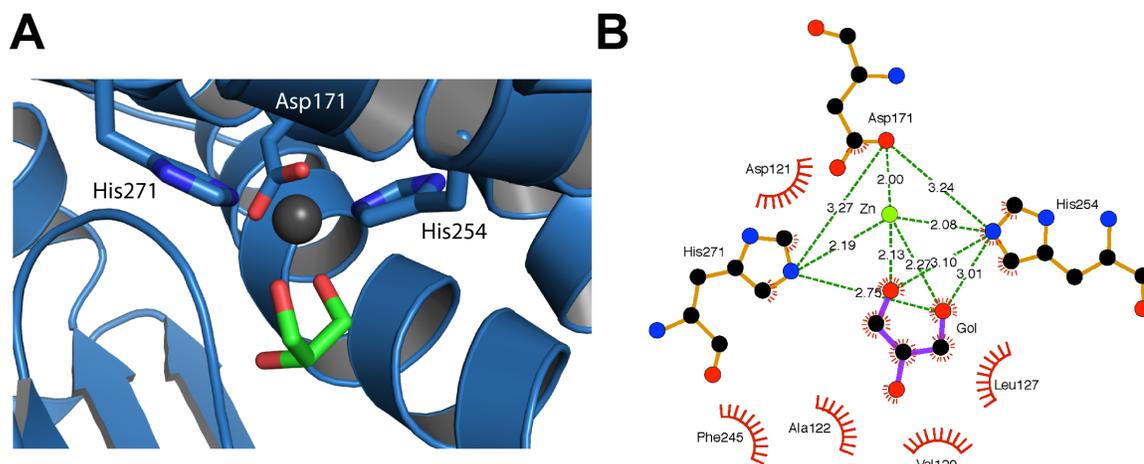
Before the pathogenicity of *Serratia* was appreciated, its red pigment was utilized as a tracer agent by the U.S. government in biological warfare and medical tests (214). It is now known that *Serratia* bacteria often contain widely ranging resistance to a variety of antibacterial agents (215). This antibiotic resistance presumably allowed the *Serratia* to escape ampicillin and chloramphenicol selection during protein expression. During routine cleaning and maintenance of our lab, we identified a small (~2 mm diameter) pink biofilm growing inside the in-house filtered water spigot that we hypothesize was the source of contamination in this experiment. The spigot was immediately sterilized and replaced preventing further investigation into the source of this contamination. It is apparent, however, that this Gram-negative facultatively anaerobic bacteria was able to out-compete, or at least co-exist, with *E. coli* in antibiotic-containing culture.

While it is possible to determine the structure of crystals without *a priori* knowledge of their sequence, this remains a difficult endeavour. In this study, we used LC-MS/MS to unambiguously identify the protein we crystallized a *Serratia* GDH without a reported DNA



Appendix Figure 7 Conserved NAD Binding Site

Molecular surface of *Serratia* glycerol dehydrogenase (blue) with bound glycerol depicted as sticks (C, green; O, red). NAD⁺, depicted as sticks (C, copper; O, red; P, magenta; N, blue), is placed from a superposition of *Bacillus stearothermophilus* glycerol dehydrogenase (PDBID: 1JQ5) which share a sequence identity of 46%.



Appendix Figure 8 Active Site of *Serratia* Glycerol Dehydrogenase

(A) Active site of *Serratia* glycerol dehydrogenase represented (blue ribbon, with close side chains colored: C, blue; O, red; N, blue) with bound glycerol depicted as sticks (C, green; O, red) and bound zinc is depicted as a black sphere. (B) Active site of *Serratia* glycerol dehydrogenase represented in 2D with non-covalent interactions highlighted by either green dashes or red combs.

Figure generated by Ligplot+ (150).

sequence. This highlights the importance of quality control measures, such as protein identification by mass spectroscopy, as part of a protein production protocol. In our case, attempting to express a potentially toxic human protein in *E. coli* may have allowed for the Gram-negative facultatively anaerobic *Serratia* bacteria to thrive rather than being outcompeted by a rapidly growing *E. coli* population.

References

1. Michell RH. 1975. Inositol phospholipids and cell surface receptor function. *Biochim Biophys Acta* 415:81-47
2. Berridge MJ, Irvine RF. 1984. Inositol trisphosphate, a novel second messenger in cellular signal transduction. *Nature* 312:315-21
3. Song G, Ouyang G, Bao S. 2005. The activation of Akt/PKB signaling pathway and cell survival. *Journal of cellular and molecular medicine* 9:59-71
4. Xiang SY, Dusaban SS, Brown JH. 2013. Lysophospholipid receptor activation of RhoA and lipid signaling pathways. *Biochim Biophys Acta* 1831:213-22
5. O'Donnell VB, Murphy RC. 2012. New families of bioactive oxidized phospholipids generated by immune cells: identification and signaling actions. *Blood* 120:1985-92
6. Irvine RF, Divecha N. 1992. Phospholipids in the nucleus--metabolism and possible functions. *Semin Cell Biol* 3:225-35
7. Irvine RF. 2003. Nuclear lipid signalling. *Nat Rev Mol Cell Biol* 4:349-60
8. Irvine RF. 2002. Nuclear lipid signaling. *Sci STKE* 2002:re13
9. Viiri K, Maki M, Lohi O. 2012. Phosphoinositides as regulators of protein-chromatin interactions. *Sci Signal* 5:pe19
10. Frascini A, Albi E, Gahan PB, Viola-Magni MP. 1992. TEM cytochemical study of the localization of phospholipids in interphase chromatin in rat hepatocytes. *Histochemistry* 97:225-35
11. McEwan IJ. 2009. Nuclear receptors: one big family. *Methods in molecular biology* 505:3-18
12. Nagy L, Schwabe JW. 2004. Mechanism of the nuclear receptor molecular switch. *Trends Biochem Sci* 29:317-24
13. Huang P, Chandra V, Rastinejad F. 2010. Structural overview of the nuclear receptor superfamily: insights into physiology and therapeutics. *Annu Rev Physiol* 72:247-72
14. Sladek FM. 2011. What are nuclear receptor ligands? *Mol Cell Endocrinol* 334:3-13
15. Kliewer SA, Lehmann JM, Willson TM. 1999. Orphan nuclear receptors: shifting endocrinology into reverse. *Science* 284:757-60
16. Schulman IG, Heyman RA. 2004. The flip side: Identifying small molecule regulators of nuclear receptors. *Chem Biol* 11:639-46
17. Moore JT, Collins JL, Pearce KH. 2006. The nuclear receptor superfamily and drug discovery. *ChemMedChem* 1:504-23
18. Fernandez-Marcos PJ, Auwerx J, Schoonjans K. 2011. Emerging actions of the nuclear receptor LRH-1 in the gut. *Biochim Biophys Acta* 1812:947-55
19. Lee Y-k, Moore DD. 2008. Liver receptor homolog-1, an emerging metabolic modulator. *Frontiers in bioscience : a journal and virtual library* 13:5950-8
20. Kellner S, Kikyo N. 2010. Transcriptional regulation of the Oct4 gene, a master gene for pluripotency. *Histology and histopathology* 25:405-12
21. Botrugno Oa, Fayard E, Annicotte J-S, Haby C, Brennan T, et al. 2004. Synergy between LRH-1 and beta-catenin induces G1 cyclin-mediated cell proliferation. *Molecular cell* 15:499-509
22. Heng JC, Feng B, Han J, Jiang J, Kraus P, et al. 2010. The nuclear receptor Nr5a2 can replace Oct4 in the reprogramming of murine somatic cells to pluripotent cells. *Cell stem cell* 6:167-74
23. Sung B, Do HJ, Park SW, Huh SH, Oh JH, et al. 2012. Regulation of OCT4 gene expression by liver receptor homolog-1 in human embryonic carcinoma cells. *Biochem Biophys Res Commun*
24. Clyne CD, Speed CJ, Zhou J, Simpson ER. 2002. Liver receptor homologue-1 (LRH-1) regulates expression of aromatase in preadipocytes. *The Journal of biological chemistry* 277:20591-7

25. Goodwin B, Jones Sa, Price RR, Watson Ma, McKee DD, et al. 2000. A regulatory cascade of the nuclear receptors FXR, SHP-1, and LRH-1 represses bile acid biosynthesis. *Molecular cell* 6:517-26
26. Chen F, Ma L, Dawson PA, Sinal CJ, Sehayek E, et al. 2003. Liver receptor homologue-1 mediates species- and cell line-specific bile acid-dependent negative feedback regulation of the apical sodium-dependent bile acid transporter. *J Biol Chem* 278:19909-16
27. del Castillo-Olivares A, Gil G. 2000. Role of FXR and FTF in bile acid-mediated suppression of cholesterol 7 α -hydroxylase transcription. *Nucleic Acids Res* 28:3587-93
28. Delerive P, Galardi CM, Bisi JE, Nicodeme E, Goodwin B. 2004. Identification of liver receptor homolog-1 as a novel regulator of apolipoprotein AI gene transcription. *Molecular endocrinology* 18:2378-87
29. Freeman LA, Kennedy A, Wu J, Bark S, Remaley AT, et al. 2004. The orphan nuclear receptor LRH-1 activates the ABCG5/ABCG8 intergenic promoter. *J Lipid Res* 45:1197-206
30. Inokuchi A, Hinoshita E, Iwamoto Y, Kohno K, Kuwano M, Uchiumi T. 2001. Enhanced expression of the human multidrug resistance protein 3 by bile salt in human enterocytes. A transcriptional control of a plausible bile acid transporter. *J Biol Chem* 276:46822-9
31. Lee YK, Schmidt DR, Cummins CL, Choi M, Peng L, et al. 2008. Liver receptor homolog-1 regulates bile acid homeostasis but is not essential for feedback regulation of bile acid synthesis. *Molecular endocrinology* 22:1345-56
32. Luo Y, Liang CP, Tall AR. 2001. The orphan nuclear receptor LRH-1 potentiates the sterol-mediated induction of the human CETP gene by liver X receptor. *J Biol Chem* 276:24767-73
33. Oosterveer MH, Matakci C, Yamamoto H, Harach T, Moullan N, et al. 2012. LRH-1-dependent glucose sensing determines intermediary metabolism in liver. *J Clin Invest* 122:2817-26
34. Bouchard MF, Taniguchi H, Viger RS. 2005. Protein kinase A-dependent synergism between GATA factors and the nuclear receptor, liver receptor homolog-1, regulates human aromatase (CYP19) PII promoter activity in breast cancer cells. *Endocrinology* 146:4905-16
35. Annicotte J-S, Chavey C, Servant N, Teyssier J, Bardin A, et al. 2005. The nuclear receptor liver receptor homolog-1 is an estrogen receptor target gene. *Oncogene* 24:8167-75
36. Wang S-L, Zheng D-Z, Lan F-H, Deng X-J, Zeng J, et al. 2008. Increased expression of hLRH-1 in human gastric cancer and its implication in tumorigenesis. *Molecular and cellular biochemistry* 308:93-100
37. Sablin EP, Krylova IN, Fletterick RJ, Ingraham HA. 2003. Structural basis for ligand-independent activation of the orphan nuclear receptor LRH-1. *Molecular cell* 11:1575-85
38. Wang W, Zhang C, Marimuthu A, Krupka HI, Tabrizizad M, et al. 2005. The crystal structures of human steroidogenic factor-1 and liver receptor homologue-1. *Proc Natl Acad Sci U S A* 102:7505-10
39. Ortlund EA, Lee Y, Solomon IH, Hager JM, Safi R, et al. 2005. Modulation of human nuclear receptor LRH-1 activity by phospholipids and SHP. *Nature structural & molecular biology* 12:357-63
40. Krylova IN, Sablin EP, Moore J, Xu RX, Waitt GM, et al. 2005. Structural analyses reveal phosphatidyl inositols as ligands for the NR5 orphan receptors SF-1 and LRH-1. *Cell* 120:343-55
41. Musille PM, Pathak MC, Lauer JL, Hudson WH, Griffin PR, Ortlund EA. 2012. Antidiabetic phospholipid-nuclear receptor complex reveals the mechanism for phospholipid-driven gene regulation. *Nat Struct Mol Biol* 19:532-7, S1-2

42. Lee JM, Lee YK, Mamrosh JL, Busby SA, Griffin PR, et al. 2011. A nuclear-receptor-dependent phosphatidylcholine pathway with antidiabetic effects. *Nature*
43. Moore D. Targeting nuclear receptors to treat type 2 diabetes. *Proc. 14th International Congress of Endocrinology, Kyoto, Japan, 2010:*
44. Hoivik EA, Lewis AE, Aumo L, Bakke M. 2010. Molecular aspects of steroidogenic factor 1 (SF-1). *Mol Cell Endocrinol* 315:27-39
45. Shinoda K, Lei H, Yoshii H, Nomura M, Nagano M, et al. 1995. Developmental defects of the ventromedial hypothalamic nucleus and pituitary gonadotroph in the Ftz-F1 disrupted mice. *Dev Dyn* 204:22-9
46. Ingraham HA, Lala DS, Ikeda Y, Luo X, Shen WH, et al. 1994. The nuclear receptor steroidogenic factor 1 acts at multiple levels of the reproductive axis. *Genes Dev* 8:2302-12
47. Mascaro C, Nadal A, Hegardt FG, Marrero PF, Haro D. 2000. Contribution of steroidogenic factor 1 to the regulation of cholesterol synthesis. *Biochem J* 350 Pt 3:785-90
48. Sugawara T, Holt JA, Kiriakidou M, Strauss JF, 3rd. 1996. Steroidogenic factor 1-dependent promoter activity of the human steroidogenic acute regulatory protein (StAR) gene. *Biochemistry* 35:9052-9
49. Cao G, Garcia CK, Wyne KL, Schultz RA, Parker KL, Hobbs HH. 1997. Structure and localization of the human gene encoding SR-BI/CLA-1. Evidence for transcriptional control by steroidogenic factor 1. *J Biol Chem* 272:33068-76
50. Lopez D, Shea-Eaton W, McLean MP. 2001. Characterization of a steroidogenic factor-1-binding site found in promoter of sterol carrier protein-2 gene. *Endocrine* 14:253-61
51. Parker KL, Schimmer BP. 1997. Steroidogenic factor 1: a key determinant of endocrine development and function. *Endocr Rev* 18:361-77
52. Schimmer BP, White PC. 2010. Minireview: steroidogenic factor 1: its roles in differentiation, development, and disease. *Mol Endocrinol* 24:1322-37
53. Ferraz-de-Souza B, Lin L, Achermann JC. 2011. Steroidogenic factor-1 (SF-1, NR5A1) and human disease. *Mol Cell Endocrinol* 336:198-205
54. Correa RV, Domenice S, Bingham NC, Billerbeck AE, Rainey WE, et al. 2004. A microdeletion in the ligand binding domain of human steroidogenic factor 1 causes XY sex reversal without adrenal insufficiency. *J Clin Endocrinol Metab* 89:1767-72
55. Camats N, Pandey AV, Fernandez-Cancio M, Andaluz P, Janner M, et al. 2012. Ten novel mutations in the NR5A1 gene cause disordered sex development in 46,XY and ovarian insufficiency in 46,XX individuals. *The Journal of clinical endocrinology and metabolism* 97:E1294-306
56. Achermann JC, Ito M, Hindmarsh PC, Jameson JL. 1999. A mutation in the gene encoding steroidogenic factor-1 causes XY sex reversal and adrenal failure in humans. *Nat Genet* 22:125-6
57. Lin L, Achermann JC. 2008. Steroidogenic factor-1 (SF-1, Ad4BP, NR5A1) and disorders of testis development. *Sex Dev* 2:200-9
58. Bulun SE, Utsunomiya H, Lin Z, Yin P, Cheng YH, et al. 2009. Steroidogenic factor-1 and endometriosis. *Mol Cell Endocrinol* 300:104-8
59. Pianovski MA, Cavalli LR, Figueiredo BC, Santos SC, Doghman M, et al. 2006. SF-1 overexpression in childhood adrenocortical tumours. *Eur J Cancer* 42:1040-3
60. Del Tredici AL, Andersen CB, Currier EA, Ohrmund SR, Fairbain LC, et al. 2008. Identification of the first synthetic steroidogenic factor 1 inverse agonists: pharmacological modulation of steroidogenic enzymes. *Mol Pharmacol* 73:900-8
61. Madoux F, Li X, Chase P, Zastrow G, Cameron MD, et al. 2008. Potent, selective and cell penetrant inhibitors of SF-1 by functional ultra-high-throughput screening. *Mol Pharmacol* 73:1776-84

62. Roth J, Madoux F, Hodder P, Roush WR. 2008. Synthesis of small molecule inhibitors of the orphan nuclear receptor steroidogenic factor-1 (NR5A1) based on isoquinolinone scaffolds. *Bioorg Med Chem Lett* 18:2628-32
63. Doghman M, Cazareth J, Douguet D, Madoux F, Hodder P, Lalli E. 2009. Inhibition of adrenocortical carcinoma cell proliferation by steroidogenic factor-1 inverse agonists. *J Clin Endocrinol Metab* 94:2178-83
64. Li Y, Choi M, Cavey G, Daugherty J, Suino K, 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
65. Urs AN, Dammer E, Sewer MB. 2006. Sphingosine regulates the transcription of CYP17 by binding to steroidogenic factor-1. *Endocrinology* 147:5249-58
66. Li D, Urs AN, Allegood J, Leon A, Merrill AH, Jr., Sewer MB. 2007. Cyclic AMP-stimulated interaction between steroidogenic factor 1 and diacylglycerol kinase theta facilitates induction of CYP17. *Mol Cell Biol* 27:6669-85
67. Urs AN, Dammer E, Kelly S, Wang E, Merrill AH, Jr., Sewer MB. 2007. Steroidogenic factor-1 is a sphingolipid binding protein. *Molecular and cellular endocrinology* 265-266:174-8
68. Sablin EP, Blind RD, Krylova IN, Ingraham JG, Cai F, et al. 2009. Structure of SF-1 bound by different phospholipids: evidence for regulatory ligands. *Mol Endocrinol* 23:25-34
69. Blind RD, Suzawa M, Ingraham HA. 2012. Direct modification and activation of a nuclear receptor-PIP(2) complex by the inositol lipid kinase IPMK. *Sci Signal* 5:ra44
70. Bensinger SJ, Tontonoz P. 2008. Integration of metabolism and inflammation by lipid-activated nuclear receptors. *Nature* 454:470-7
71. Desvergne B, Wahli W. 1999. Peroxisome proliferator-activated receptors: nuclear control of metabolism. *Endocr Rev* 20:649-88
72. Berger J, Moller DE. 2002. The mechanisms of action of PPARs. *Annu Rev Med* 53:409-35
73. Miyata KS, McCaw SE, Marcus SL, Rachubinski RA, Capone JP. 1994. The peroxisome proliferator-activated receptor interacts with the retinoid X receptor in vivo. *Gene* 148:327-30
74. Auboeuf D, Rieusset J, Fajas L, Vallier P, Frering V, et al. 1997. Tissue distribution and quantification of the expression of mRNAs of peroxisome proliferator-activated receptors and liver X receptor-alpha in humans: no alteration in adipose tissue of obese and NIDDM patients. *Diabetes* 46:1319-27
75. Martin G, Schoonjans K, Lefebvre AM, Staels B, Auwerx J. 1997. Coordinate regulation of the expression of the fatty acid transport protein and acyl-CoA synthetase genes by PPARalpha and PPARgamma activators. *J Biol Chem* 272:28210-7
76. Motojima K, Passilly P, Peters JM, Gonzalez FJ, Latruffe N. 1998. Expression of putative fatty acid transporter genes are regulated by peroxisome proliferator-activated receptor alpha and gamma activators in a tissue- and inducer-specific manner. *J Biol Chem* 273:16710-4
77. Dreyer C, Keller H, Mahfoudi A, Laudet V, Krey G, Wahli W. 1993. Positive regulation of the peroxisomal beta-oxidation pathway by fatty acids through activation of peroxisome proliferator-activated receptors (PPAR). *Biol Cell* 77:67-76
78. Schoonjans K, Peinado-Onsurbe J, Lefebvre AM, Heyman RA, Briggs M, et al. 1996. PPARalpha and PPARgamma activators direct a distinct tissue-specific transcriptional response via a PPRE in the lipoprotein lipase gene. *EMBO J* 15:5336-48
79. Pyper SR, Viswakarma N, Yu S, Reddy JK. 2010. PPARalpha: energy combustion, hypolipidemia, inflammation and cancer. *Nucl Recept Signal* 8:e002

80. Forman BM, Chen J, Evans RM. 1997. Hypolipidemic drugs, polyunsaturated fatty acids, and eicosanoids are ligands for peroxisome proliferator-activated receptors alpha and delta. *Proc Natl Acad Sci U S A* 94:4312-7
81. Chakravarthy MV, Lodhi IJ, Yin L, Malapaka RR, Xu HE, et al. 2009. Identification of a physiologically relevant endogenous ligand for PPARalpha in liver. *Cell* 138:476-88
82. Huang JV, Greyson CR, Schwartz GG. 2012. PPAR-gamma as a therapeutic target in cardiovascular disease: evidence and uncertainty. *J Lipid Res* 53:1738-54
83. Martin H. 2009. Role of PPAR-gamma in inflammation. Prospects for therapeutic intervention by food components. *Mutat Res* 669:1-7
84. Hammond VJ, Morgan AH, Lauder S, Thomas CP, Brown S, et al. 2012. Novel Keto-phospholipids Are Generated by Monocytes and Macrophages, Detected in Cystic Fibrosis, and Activate Peroxisome Proliferator-activated Receptor-gamma. *J Biol Chem* 287:41651-66
85. Davies SS, Pontsler AV, Marathe GK, Harrison KA, Murphy RC, et al. 2001. Oxidized alkyl phospholipids are specific, high affinity peroxisome proliferator-activated receptor gamma ligands and agonists. *J Biol Chem* 276:16015-23
86. Delerive P, Furman C, Teissier E, Fruchart J, Duriez P, Staels B. 2000. Oxidized phospholipids activate PPARalpha in a phospholipase A2-dependent manner. *FEBS Lett* 471:34-8
87. Oro AE, McKeown M, Evans RM. 1990. Relationship between the product of the Drosophila ultraspiracle locus and the vertebrate retinoid X receptor. *Nature* 347:298-301
88. Henrich VC, Sliter TJ, Lubahn DB, MacIntyre A, Gilbert LI. 1990. A steroid/thyroid hormone receptor superfamily member in Drosophila melanogaster that shares extensive sequence similarity with a mammalian homologue. *Nucleic Acids Res* 18:4143-8
89. Schwedes CC, Carney GE. 2012. Ecdysone signaling in adult Drosophila melanogaster. *J Insect Physiol* 58:293-302
90. Jones G, Sharp PA. 1997. Ultraspiracle: an invertebrate nuclear receptor for juvenile hormones. *Proc Natl Acad Sci U S A* 94:13499-503
91. Jones D, Jones G. 2007. Farnesoid secretions of dipteran ring glands: what we do know and what we can know. *Insect biochemistry and molecular biology* 37:771-98
92. Clayton GM, Peak-Chew SY, Evans RM, Schwabe JW. 2001. The structure of the ultraspiracle ligand-binding domain reveals a nuclear receptor locked in an inactive conformation. *Proc Natl Acad Sci U S A* 98:1549-54
93. Billas IM, Moulinier L, Rochel N, Moras D. 2001. Crystal structure of the ligand-binding domain of the ultraspiracle protein USP, the ortholog of retinoid X receptors in insects. *J Biol Chem* 276:7465-74
94. Billas IM, Iwema T, Garnier JM, Mitschler A, Rochel N, Moras D. 2003. Structural adaptability in the ligand-binding pocket of the ecdysone hormone receptor. *Nature* 426:91-6
95. Browning C, Martin E, Loch C, Wurtz JM, Moras D, et al. 2007. Critical role of desolvation in the binding of 20-hydroxyecdysone to the ecdysone receptor. *J Biol Chem* 282:32924-34
96. Iyer LM, Koonin EV, Aravind L. 2001. Adaptations of the helix-grip fold for ligand binding and catalysis in the START domain superfamily. *Proteins* 43:134-44
97. Ponting CP, Aravind L. 1999. START: a lipid-binding domain in StAR, HD-ZIP and signalling proteins. *Trends Biochem Sci* 24:130-2
98. Roderick SL, Chan WW, Agate DS, Olsen LR, Vetting MW, et al. 2002. Structure of human phosphatidylcholine transfer protein in complex with its ligand. *Nature structural biology* 9:507-11

99. Wirtz KW, Devaux PF, Bienvenue A. 1980. Phosphatidylcholine exchange protein catalyzes the net transfer of phosphatidylcholine to model membranes. *Biochemistry* 19:3395-9
100. Kamp HH, Wirtz WA, Baer PR, Slotboom AJ, Rosenthal AF, et al. 1977. Specificity of the phosphatidylcholine exchange protein from bovine liver. *Biochemistry* 16:1310-6
101. Johnson LW, Zilversmit DB. 1975. Catalytic properties of phospholipid exchange protein from bovine heart. *Biochim Biophys Acta* 375:165-75
102. Kang HW, Wei J, Cohen DE. 2010. PC-TP/StARD2: Of membranes and metabolism. *Trends Endocrinol Metab* 21:449-56
103. Kanno K, Wu MK, Agate DS, Fanelli BJ, Wagle N, et al. 2007. Interacting proteins dictate function of the minimal START domain phosphatidylcholine transfer protein/StarD2. *J Biol Chem* 282:30728-36
104. Kang HW, Kanno K, Scapa EF, Cohen DE. 2010. Regulatory role for phosphatidylcholine transfer protein/StarD2 in the metabolic response to peroxisome proliferator activated receptor alpha (PPARalpha). *Biochim Biophys Acta* 1801:496-502
105. Zheng J, Singh VK, Jia Z. 2005. Identification of an ITPase/XTPase in Escherichia coli by structural and biochemical analysis. *Structure* 13:1511-20
106. Giabbai B, Sidobre S, Crispin MD, Sanchez-Ruiz Y, Bachi A, et al. 2005. Crystal structure of mouse CD1d bound to the self ligand phosphatidylcholine: a molecular basis for NKT cell activation. *J Immunol* 175:977-84
107. Gadola SD, Zaccai NR, Harlos K, Shepherd D, Castro-Palomino JC, et al. 2002. Structure of human CD1b with bound ligands at 2.3 Å, a maze for alkyl chains. *Nature immunology* 3:721-6
108. Jullien D, Stenger S, Ernst WA, Modlin RL. 1997. CD1 presentation of microbial nonpeptide antigens to T cells. *The Journal of clinical investigation* 99:2071-4
109. Schaaf G, Ortlund EA, Tyeryar KR, Mousley CJ, Ile KE, et al. 2008. Functional anatomy of phospholipid binding and regulation of phosphoinositide homeostasis by proteins of the sec14 superfamily. *Mol Cell* 29:191-206
110. Tilley SJ, Skippen A, Murray-Rust J, Swigart PM, Stewart A, et al. 2004. Structure-function analysis of human [corrected] phosphatidylinositol transfer protein alpha bound to phosphatidylinositol. *Structure* 12:317-26
111. Bankaitis VA, Mousley CJ, Schaaf G. 2010. The Sec14 superfamily and mechanisms for crosstalk between lipid metabolism and lipid signaling. *Trends Biochem Sci* 35:150-60
112. Gozani O, Karuman P, Jones DR, Ivanov D, Cha J, et al. 2003. The PHD finger of the chromatin-associated protein ING2 functions as a nuclear phosphoinositide receptor. *Cell* 114:99-111
113. Ingraham Ha, Redinbo MR. 2005. Orphan nuclear receptors adopted by crystallography. *Current opinion in structural biology* 15:708-15
114. Gee AC, Katzenellenbogen JA. 2001. Probing conformational changes in the estrogen receptor: evidence for a partially unfolded intermediate facilitating ligand binding and release. *Mol Endocrinol* 15:421-8
115. Sablin EP, Woods A, Krylova IN, Hwang P, Ingraham HA, Fletterick RJ. 2008. The structure of corepressor Dax-1 bound to its target nuclear receptor LRH-1. *Proceedings of the National Academy of Sciences of the United States of America* 105:18390-5
116. Ortlund EA, Lee Y, Solomon IH, Hager JM, Safi R, et al. 2005. Modulation of human nuclear receptor LRH-1 activity by phospholipids and SHP. *Nat Struct Mol Biol* 12:357-63
117. Li Y, Choi M, Suino K, Kovach A, Daugherty J, et al. 2005. Structural and biochemical basis for selective repression of the orphan nuclear receptor liver receptor homolog 1 by small heterodimer partner. *Proceedings of the National Academy of Sciences of the United States of America* 102:9505-10

118. Wang W, Zhang C, Marimuthu A, Krupka HI, Tabrizizad M, et al. 2005. The crystal structures of human steroidogenic factor-1 and liver receptor homologue-1. *Proceedings of the National Academy of Sciences of the United States of America* 102:7505-10
119. Ortlund EA, Lee Y, Solomon IH, Hager JM, Safi R, et al. 2005. Modulation of human nuclear receptor LRH-1 activity by phospholipids and SHP. *Nat Struct Mol Biol*
120. Krylova IN, Sablin EP, Moore J, Xu RX, Waitt GM, et al. 2005. Structural analyses reveal phosphatidyl inositols as ligands for the NR5 orphan receptors SF-1 and LRH-1. *Cell* 120:343-55
121. Li Y, Choi M, Suino K, Kovach A, Daugherty J, et al. 2005. Structural and biochemical basis for selective repression of the orphan nuclear receptor liver receptor homolog 1 by small heterodimer partner. *P Natl Acad Sci USA* 102:9505-10
122. Fernandez-Marcos PJ, Auwerx J, Schoonjans K. 2010. Emerging actions of the nuclear receptor LRH-1 in the gut. *Biochim Biophys Acta*
123. Lee YK, Moore DD. 2008. Liver receptor homolog-1, an emerging metabolic modulator. *Front Biosci* 13:5950-8
124. Wagner RT, Xu X, Yi F, Merrill BJ, Cooney AJ. 2010. Canonical Wnt/beta-catenin regulation of liver receptor homolog-1 mediates pluripotency gene expression. *Stem cells* 28:1794-804
125. Gu P, Goodwin B, Chung AC, Xu X, Wheeler DA, et al. 2005. Orphan nuclear receptor LRH-1 is required to maintain Oct4 expression at the epiblast stage of embryonic development. *Mol Cell Biol* 25:3492-505
126. Clyne CD, Speed CJ, Zhou J, Simpson ER. 2002. Liver receptor homologue-1 (LRH-1) regulates expression of aromatase in preadipocytes. *J Biol Chem* 277:20591-7
127. Clyne CD, Kovacic A, Speed CJ, Zhou J, Pezzi V, Simpson ER. 2004. Regulation of aromatase expression by the nuclear receptor LRH-1 in adipose tissue. *Mol Cell Endocrinol* 215:39-44
128. Zhou J, Suzuki T, Kovacic A, Saito R, Miki Y, et al. 2005. Interactions between prostaglandin E-2, liver receptor homologue-1, and aromatase in breast cancer. *Cancer Res* 65:657-63
129. Chand AL, Herridge KA, Thompson EW, Clyne CD. 2010. The orphan nuclear receptor LRH-1 promotes breast cancer motility and invasion. *Endocr Relat Cancer* 17:965-75
130. Annicotte JS, Chavey C, Servant N, Teysier J, Bardin A, et al. 2005. The nuclear receptor liver receptor homolog-1 is an estrogen receptor target gene. *Oncogene* 24:8167-75
131. Thiruchelvam PT, Lai CF, Hua H, Thomas RS, Hurtado A, et al. 2010. The liver receptor homolog-1 regulates estrogen receptor expression in breast cancer cells. *Breast Cancer Res Treat*
132. Goodwin B, Jones SA, Price RR, Watson MA, McKee DD, et al. 2000. A regulatory cascade of the nuclear receptors FXR, SHP-1, and LRH-1 represses bile acid biosynthesis. *Mol Cell* 6:517-26
133. Sablin EP, Woods A, Krylova IN, Hwang P, Ingraham HA, Fletterick RJ. 2008. The structure of corepressor Dax-1 bound to its target nuclear receptor LRH-1. *Proc Natl Acad Sci U S A* 105:18390-5
134. Sablin EP, Krylova IN, Fletterick RJ, Ingraham HA. 2003. Structural basis for ligand-independent activation of the orphan nuclear receptor LRH-1. *Mol Cell* 11:1575-85
135. Ingraham HA, Redinbo MR. 2005. Orphan nuclear receptors adopted by crystallography. *Curr Opin Struct Biol* 15:708-15
136. Xu PL, Kong YY, Xie YH, Wang Y. 2003. Corepressor SMRT specifically represses the transcriptional activity of orphan nuclear receptor hB1F/hLRH-1. *Sheng Wu Hua Xue Yu Sheng Wu Wu Li Xue Bao (Shanghai)* 35:897-903

137. Whitby RJ, Stec J, Blind RD, Dixon S, Leesnitzer LM, et al. 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*
138. Otwinowski Z, Minor W. 1997. Processing of X-ray diffraction data collected in oscillation mode. *Method Enzymol* 276:307-26
139. Murshudov GN, Vagin AA, Dodson EJ. 1997. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* 53:240-55
140. Potterton E, Briggs P, Turkenburg M, Dodson E. 2003. A graphical user interface to the CCP4 program suite. *Acta Crystallogr D Biol Crystallogr* 59:1131-7
141. Emsley P, Cowtan K. 2004. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60:2126-32
142. Emsley P, Lohkamp B, Scott WG, Cowtan K. 2010. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66:486-501
143. Chen VB, Arendall WB, 3rd, Headd JJ, Keedy DA, Immormino RM, et al. 2010. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* 66:12-21
144. Bligh EG, Dyer WJ. 1959. A rapid method of total lipid extraction and purification. *Can J Biochem Physiol* 37:911-7
145. Zhou X, Arthur G. 1992. Improved procedures for the determination of lipid phosphorus by malachite green. *J Lipid Res* 33:1233-6
146. Louis-Jeune C, Andrade-Navarro MA, Perez-Iratxeta C. 2011. Prediction of protein secondary structure from circular dichroism using theoretically derived spectra. *Proteins*
147. Xu P, Duong DM, Peng J. 2009. Systematical optimization of reverse-phase chromatography for shotgun proteomics. *Journal of proteome research* 8:3944-50
148. Eng J, McCormack AL, Yates JR, 3rd. 1994. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J. Am. Soc. Mass Spectrom.* 5:976-89
149. Chalmers MJ, Busby Sa, Pascal BD, He Y, Hendrickson CL, et al. 2006. Probing protein ligand interactions by automated hydrogen/deuterium exchange mass spectrometry. *Analytical chemistry* 78:1005-14
150. Laskowski RA, Swindells MB. 2011. LigPlot+: Multiple Ligand-Protein Interaction Diagrams for Drug Discovery. *Journal of chemical information and modeling* 51:2778-86
151. Fumiaki Yumoto PN, Elena P. Sablin, John D. Baxter, Paul Webb, and Robert J. Fletterick 2011. Structural basis of coactivation of liver receptor homolog-1 by β -catenin. *Proc Natl Acad Sci U S A*
152. Jasuja R, Ulloor J, Yengo CM, Choong K, Istomin AY, et al. 2009. Kinetic and thermodynamic characterization of dihydrotestosterone-induced conformational perturbations in androgen receptor ligand-binding domain. *Mol Endocrinol* 23:1231-41
153. Venteclef N, Jakobsson T, Ehrlund A, Damdimopoulos A, Mikkonen L, et al. 2010. GPS2-dependent corepressor/SUMO pathways govern anti-inflammatory actions of LRH-1 and LXRbeta in the hepatic acute phase response. *Genes Dev* 24:381-95
154. Musille PM, Kohn JA, Ortlund EA. 2013. Phospholipid - Driven gene regulation. *FEBS Lett*
155. Hammer GD, Ingraham HA. 1999. Steroidogenic factor-1: its role in endocrine organ development and differentiation. *Front Neuroendocrinol* 20:199-223
156. Ueda H, Sonoda S, Brown JL, Scott MP, Wu C. 1990. A sequence-specific DNA-binding protein that activates fushi tarazu segmentation gene expression. *Genes Dev* 4:624-35
157. Takahashi K, Yamanaka S. 2006. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 126:663-76

158. Yu J, Vodyanik MA, Smuga-Otto K, Antosiewicz-Bourget J, Frane JL, et al. 2007. Induced pluripotent stem cell lines derived from human somatic cells. *Science* 318:1917-20
159. Kelly VR, Xu B, Kuick R, Koenig RJ, Hammer GD. 2010. Dax1 up-regulates Oct4 expression in mouse embryonic stem cells via LRH-1 and SRA. *Mol Endocrinol* 24:2281-91
160. Kelly VR, Hammer GD. 2011. LRH-1 and Nanog regulate Dax1 transcription in mouse embryonic stem cells. *Mol Cell Endocrinol* 332:116-24
161. Pare JF, Malenfant D, Courtemanche C, Jacob-Wagner M, Roy S, et al. 2004. The fetoprotein transcription factor (FTF) gene is essential to embryogenesis and cholesterol homeostasis and is regulated by a DR4 element. *J Biol Chem* 279:21206-16
162. Fayard E, Auwerx J, Schoonjans K. 2004. LRH-1: an orphan nuclear receptor involved in development, metabolism and steroidogenesis. *Trends Cell Biol* 14:250-60
163. Bookout AL, Jeong Y, Downes M, Yu RT, Evans RM, Mangelsdorf DJ. 2006. Anatomical profiling of nuclear receptor expression reveals a hierarchical transcriptional network. *Cell* 126:789-99
164. Laskowski RA, McArthur, M.W., Moss, D.S. & Thornton, J. M. 1993. Procheck: a program to check the stereochemical quality of protein structure. *J appl Cryst* 26:283-91
165. Whitby RJ, Stec J, Blind RD, Dixon S, Leesnitzer LM, et al. 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* 1
166. Ortlund EA, Bridgham JT, Redinbo MR, Thornton JW. 2007. Crystal structure of an ancient protein: Evolution by conformational epistasis. *Science* 317:1544-8
167. Schoch GA, D'Arcy B, Stihle M, Burger D, Bar D, et al. 2010. Molecular switch in the glucocorticoid receptor: active and passive antagonist conformations. *J Mol Biol* 395:568-77
168. Raaijmakers HC, Versteegh JE, Uitdehaag JC. 2009. The X-ray structure of RU486 bound to the progesterone receptor in a destabilized agonistic conformation. *Journal of Biological Chemistry* 284:19572-9
169. Bledsoe RK, Madauss KP, Holt JA, Apolito CJ, Lambert MH, et al. 2005. A ligand-mediated hydrogen bond network required for the activation of the mineralocorticoid receptor. *J Biol Chem* 280:31283-93
170. Yoo J, Ko S, Kim H, Sampson H, Yun JH, et al. 2011. Crystal structure of Fushi tarazu factor 1 ligand binding domain/Fushi tarazu peptide complex identifies new class of nuclear receptors. *J Biol Chem* 286:31225-31
171. Bridgham JT, Eick GN, Larroux C, Deshpande K, Harms MJ, et al. 2010. Protein evolution by molecular tinkering: diversification of the nuclear receptor superfamily from a ligand-dependent ancestor. *PLoS Biol* 8
172. Watkins RE, Wisely GB, Moore LB, Collins JL, Lambert MH, et al. 2001. The human nuclear xenobiotic receptor PXR: structural determinants of directed promiscuity. *Science* 292:2329-33
173. Nolte RT, Wisely GB, Westin S, Cobb JE, Lambert MH, et al. 1998. Ligand binding and co-activator assembly of the peroxisome proliferator-activated receptor-gamma. *Nature* 395:137-43
174. Mullaney BC, Blind RD, Lemieux GA, Perez CL, Elle IC, 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
175. Wang DY, Kumar S, Hedges SB. 1999. Divergence time estimates for the early history of animal phyla and the origin of plants, animals and fungi. *Proc Biol Sci* 266:163-71
176. Musille PM, Kohn JA, Ortlund EA. 2013. Phospholipid – Driven gene regulation. *FEBS Letters*

177. Blind RD, Suzawa M, Ingraham HA. 2012. Direct Modification and Activation of a Nuclear Receptor-PIP2 Complex by the Inositol Lipid Kinase IPMK. *Sci Signal* 5
178. Oosterveer MH, Schoonjans K. 2013. Hepatic glucose sensing and integrative pathways in the liver. *Cell Mol Life Sci*
179. Zhang C, Large MJ, Duggavathi R, DeMayo FJ, Lydon JP, et al. 2013. Liver receptor homolog-1 is essential for pregnancy. *Nature Medicine* 19:1061-6
180. Gerrits H, Parade MC, Koonen-Reemst AM, Bakker NE, Timmer-Hellings L, et al. 2014. Reversible infertility in a liver receptor homologue-1 (LRH-1)-knockdown mouse model. *Reproduction, fertility, and development* 26:293-306
181. Kelly VR, Xu B, Kuick R, Koenig RJ, Hammer GD. 2010. Dax1 Up-Regulates Oct4 Expression in Mouse Embryonic Stem Cells via LRH-1 and SRA. *Molecular endocrinology (Baltimore, Md.)* 24:1-11
182. Venticlef N, Jakobsson T, Steffensen KR, Treuter E. 2011. Metabolic nuclear receptor signaling and the inflammatory acute phase response. *Trends in endocrinology and metabolism: TEM*:1-11
183. Safi R, Kovacic A, Gaillard S, Murata Y, Simpson ER, et al. 2005. Coactivation of liver receptor homologue-1 by peroxisome proliferator-activated receptor gamma coactivator-1alpha on aromatase promoter II and its inhibition by activated retinoid X receptor suggest a novel target for breast-specific antiestrogen therapy. *Cancer research* 65:11762-70
184. Whitby RJ, Dixon S, Maloney PR, Delerive P, Goodwin BJ, et al. 2006. Identification of small molecule agonists of the orphan nuclear receptors liver receptor homolog-1 and steroidogenic factor-1. *Journal of medicinal chemistry* 49:6652-5
185. Busby S, Nuhant P, Cameron M, Mercer BA, Hodder P, et al. 2010. Discovery of Inverse Agonists for the Liver Receptor Homologue-1 (LRH1; NR5A2). In *Probe Reports from the NIH Molecular Libraries Program*. Bethesda (MD). Number of.
186. Benod C, Carlsson J, Uthayaruban R, Hwang P, Irwin JJ, et al. 2013. Structure-based discovery of antagonists of nuclear receptor LRH-1. *J Biol Chem* 288:19830-44
187. Zhi X, Zhou XE, He Y, Zechner C, Suino-Powell KM, et al. 2014. Structural insights into gene repression by the orphan nuclear receptor SHP. *Proc Natl Acad Sci U S A* 111:839-44
188. Goodwin B, Watson MA, Kim H, Miao J, Kemper JK, Kliewer SA. 2003. Differential regulation of rat and human CYP7A1 by the nuclear oxysterol receptor liver X receptor-alpha. *Molecular endocrinology* 17:386-94
189. Lu TT, Makishima M, Repa JJ, Schoonjans K, Kerr TA, et al. 2000. Molecular basis for feedback regulation of bile acid synthesis by nuclear receptors. *Molecular cell* 6:507-15
190. Otwinowski Z, Minor W. 1997. [20] Processing of X-ray diffraction data collected in oscillation mode. 276:307-26
191. Murshudov GN, Skubak P, Lebedev AA, Pannu NS, Steiner RA, et al. 2011. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr* 67:355-67
192. Echols N, Grosse-Kunstleve RW, Afonine PV, Bunkoczi G, Chen VB, et al. 2012. Graphical tools for macromolecular crystallography in PHENIX. *Journal of applied crystallography* 45:581-6
193. Li Y, Lambert MH, Xu HE. 2003. Activation of nuclear receptors: a perspective from structural genomics. *Structure (London, England : 1993)* 11:741-6
194. Li Y, Choi M, Cavey G, Daugherty J, Suino K, et al. 2005. Crystallographic identification and functional characterization of phospholipids as ligands for the orphan nuclear receptor steroidogenic factor-1. *Molecular cell* 17:491-502

195. Li D, Urs AN, Allegood J, Leon A, Merrill AH, Sewer MB. 2007. Cyclic AMP-stimulated interaction between steroidogenic factor 1 and diacylglycerol kinase theta facilitates induction of CYP17. *Molecular and cellular biology* 27:6669-85
196. Kanayama T, Arito M, So K, Hachimura S, Inoue J, Sato R. 2007. Interaction between sterol regulatory element-binding proteins and liver receptor homolog-1 reciprocally suppresses their transcriptional activities. *The Journal of biological chemistry* 282:10290-8
197. Qin J, Gao DM, Jiang QF, Zhou Q, Kong YY, et al. 2004. Prospero-related homeobox (Prox1) is a corepressor of human liver receptor homolog-1 and suppresses the transcription of the cholesterol 7-alpha-hydroxylase gene. *Molecular endocrinology* 18:2424-39
198. Ohno M, Komakine J, Suzuki E, Nishizuka M, Osada S, Imagawa M. 2010. Repression of the promoter activity mediated by liver receptor homolog-1 through interaction with ku proteins. *Biological & pharmaceutical bulletin* 33:784-91
199. Stokes-Rees I, Sliz P. 2010. Protein structure determination by exhaustive search of Protein Data Bank derived databases. *Proc Natl Acad Sci U S A* 107:21476-81
200. Bolanos-Garcia VM, Davies OR. 2006. Structural analysis and classification of native proteins from E. coli commonly co-purified by immobilised metal affinity chromatography. *Biochim Biophys Acta* 1760:1304-13
201. Psakis G, Polaczek J, Essen LO. 2009. AcrB et al.: Obstinate contaminants in a picogram scale. One more bottleneck in the membrane protein structure pipeline. *Journal of structural biology* 166:107-11
202. Kiser PD, Lodowski DT, Palczewski K. 2007. Purification, crystallization and structure determination of native GroEL from Escherichia coli lacking bound potassium ions. *Acta Crystallogr Sect F Struct Biol Cryst Commun* 63:457-61
203. Tiwari N, Woods L, Haley R, Kight A, Goforth R, et al. 2010. Identification and characterization of native proteins of Escherichia coli BL-21 that display affinity towards Immobilized Metal Affinity Chromatography and Hydrophobic Interaction Chromatography Matrices. *Protein Expr Purif* 70:191-5
204. Lusty C. 1999. A gentle vapor-diffusion technique for cross-linking of protein crystals for cryocrystallography. *Journal of applied crystallography* 32:106-12
205. Ruzheinikov SN, Burke J, Sedelnikova S, Baker PJ, Taylor R, et al. 2001. Glycerol dehydrogenase. structure, specificity, and mechanism of a family III polyol dehydrogenase. *Structure* 9:789-802
206. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. 2007. Phaser crystallographic software. *Journal of applied crystallography* 40:658-74
207. Emsley P, Lohkamp B, Scott WG, Cowtan K. 2010. Features and development of Coot. *Acta crystallographica. Section D, Biological crystallography* 66:486-501
208. Forage RG, Lin EC. 1982. DHA system mediating aerobic and anaerobic dissimilation of glycerol in Klebsiella pneumoniae NCIB 418. *Journal of bacteriology* 151:591-9
209. May JW, Sloan J. 1981. Glycerol Utilization by Schizosaccharomyces pombe: Dehydrogenation as the Initial Step. *Microbiology* 123:183-5
210. Lin EC. 1976. Glycerol dissimilation and its regulation in bacteria. *Annual review of microbiology* 30:535-78
211. Spencer P, Bown KJ, Scawen MD, Atkinson T, Gore MG. 1989. Isolation and characterisation of the glycerol dehydrogenase from Bacillus stearothermophilus. *Biochim Biophys Acta* 994:270-9
212. Brinen LS, Canaves JM, Dai X, Deacon AM, Elsliger MA, et al. 2003. Crystal structure of a zinc-containing glycerol dehydrogenase (TM0423) from Thermotoga maritima at 1.5 Å resolution. *Proteins* 50:371-4

213. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, et al. 2004. UCSF Chimera--a visualization system for exploratory research and analysis. *Journal of computational chemistry* 25:1605-12
214. Mahlen SD. 2011. Serratia infections: from military experiments to current practice. *Clinical microbiology reviews* 24:755-91
215. Stock I, Burak S, Sherwood KJ, Gruger T, Wiedemann B. 2003. Natural antimicrobial susceptibilities of strains of 'unusual' Serratia species: S. ficaria, S. fonticola, S. odorifera, S. plymuthica and S. rubidaea. *The Journal of antimicrobial chemotherapy* 51:865-85