Title

Molecular Mechanisms Underlying Repression Of Subfamily V Nuclear Receptors By A Ubiquitin-Like Protein SUMO

Permalink

https://escholarship.org/uc/item/7xv5t95k

Author

Campbell, Lioudmila

Publication Date

2008-01-03

Peer reviewed|Thesis/dissertation

MOLECULAR MECHANISMS UNDERLYING REPRESSION OF SUBFAMILY

V NUCLEAR RECEPTORS BY A UBIQUITIN-LIKE PROTEIN SUMO

by

LIOUDMILA ALEXANDROVNA CAMPBELL

DISSERTATION

Submitted in satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

BIOCHEMISTRY AND MOLECULAR BIOLOGY

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

DEDICATION

To my loving parents

My life companions Christopher and Snowshoes

ACKNOWLEDGMENTS

ABSTRACT

MOLECULAR MECHANISMS UNDERLYING REPRESSION OF SUBFAMILY V NUCLEAR RECEPTORS BY A UBIQUITIN-LIKE PROTEIN SUMO

LIOUDMILA ALEXANDROVNA CAMPBELL

Nuclear hormone receptors are transcription factors that are uniquely regulated by lipophilic ligands. Subfamily V contains two vertebrate proteins Steroidogenic Factor 1 (SF-1, NR5A1) and Liver Receptor Homolog 1 (LRH-1, NR5A2) and a *Drosophila* homolog Ftz-F1 involved in larval segmentation. SF-1 controls male sexual differentiation, endocrine organ formation, and adult adrenal function. LRH-1 regulates genes important in bile acid metabolism, cholesterol transport, and ovarian physiology. Because Subfamily V receptors govern such diverse and vital transcriptional programs, it is essential to understand how these receptors are regulated. In addition to a putative phospholipid ligand, numerous posttranslational modifications such as phosphorylation and acetylation are known to modulate SF-1 and LRH-1 activity. Sumoylation or posttranslational modification with a small ubiquitin-like protein (SUMO) is a novel mechanism that potently represses Subfamily V receptors. In collaboration with Dr. Martin Lee, I examined the regulation of SF-1 sumoylation by SUMO E3 ligases, PIAS proteins, and showed that transcriptional repression of SF-1 by SUMO1 involves a DEAD-box helicase DP103, uncovering a novel role for ATPases/RNA helicases in transcriptional control. Additionally, I dissected structural and biochemical consequences of SF-1 sumoylation using in vitro sumoylation system. Sumoylation reduces in vitro MAP kinase phosphorylation of SF-1, but does not affect SF-1 LBD structure, as

demonstrated by nuclear magnetic resonance spectroscopy. Moreover, I discovered a regulatory relationship between SF-1 DNA-binding and sumoylation. DNA-binding inhibits SF-1 sumoylation at the DBD, while SF-1 DBD sumoylation selectively inhibits SF-1 DNA- binding at low affinity sites. I propose that Subfamily V receptor sumoylation differentially represses target genes through combined action of the DEAD-box helicase DP103 and direct, but selective inhibition of DNA-binding. Collectively, these studies establish the importance of sumoylation in regulating transcription factor function.

TABLE OF CONTENTS

Page

LIST OF FIGURES AND TABLES

Page

CHAPTER II: Chemistry Of Nuclear Receptors

CHAPTER III: The DEAD-Box Protein DP103 (Ddx20, Gemin-3) Represses Orphan Nuclear Receptors Via SUMO-Modification

CHAPTER IV: Sumoylation Selectively Regulates Steroidogenic Factor 1 (NR5A1) Target Genes

CHAPTER I

Overview
Sumoylation

Overview

Chapter I consists of a brief overview of enzymology and the biological significance of posttranslational modification of proteins with a small ubiquitin-like modifier (SUMO) or sumoylation. Chapter II is a review of structure and function of nuclear hormone receptors with a special emphasis on the extent of ligand-dependence in each nuclear receptor subfamily. It also offers an overview of the latest research into novel approaches to manipulation of nuclear receptor activity outside the ligand-binding pocket. Chapter III is a collaboration with a talented scientist, Dr. Martin Lee. This chapter describes the original observation of Subfamily V nuclear hormone receptor sumoylation in cells and thoroughly examines how sumoylation of SF-1 leads to repression of SF-1 transcriptional activity through interaction with PIASy and a DEAD-box helicase DP103. In Chapter IV, I obtained the structure of the sumoylated SF-1 LBD and examined the molecular effects of sumoylation on SF-1 DNA-binding activity.

Sumoylation

In recent years, a great number of small proteins that all share a ubiquitin-like fold and are attached at lysine residues in substrate proteins have been discovered [1]. Among them is a 100 amino acid protein called small ubiquitin-like modifier (SUMO) that shares 18% sequence identity with ubiquitin [2]. Sumoylation occurs on substrate protein canonical sites, ψKXE, where ψ is a hydrophobic amino acid, X is any amino acid, and K is the acceptor lysine. Consequently, the sites of SUMO attachment within the substrate protein can be predicted from the amino acid sequence [3]. The SUMO protein family

now consists of four functionally-distinct members, SUMO1 through 4. SUMO1 does not have a consensus sumoylation motif involved in SUMO-chain formation; therefore, each target lysine will be conjugated to only one SUMO1 molecule. Modification of proteins with SUMO1 has diverse consequences, with its most prominent role emerging in repression of transcription [2]. SUMO2 and SUMO3 form poly-SUMO chains, and their main function appears to be in responses to various cellular stresses [4]. SUMO-4 is 87% identical to SUMO2, and its cellular function is currently unknown, although it seems to be involved in susceptibility to type 1 diabetes [5].

Attachment of SUMO to substrates occurs by a mechanism similar to ubiquitination, and the sumoylation machinery is conserved from yeast to human [2]. SUMO1 is produced as a precursor protein, and undergoes C-terminal processing by SUMO isopeptidases. SUMO E1 is a heterodimer of two proteins, SAE1/SAE2 in human. In the presence of ATP, SUMO E1 activates the mature form of SUMO1 and transfers it to the catalytic site of SUMO E2, Ubc9. A thioester bond is formed between the C-terminal glycine in SUMO1 and the catalytic cysteine in Ubc9. Ubc9 directly recognizes the sumoylation consensus motif on the substrate protein and completes SUMO transfer through formation of an isopeptide bond between the C-terminal glycine in SUMO and the ε-amino group of the lysine. Intriguingly, Ubc9 is the sole E2 for sumoylation, and conjugates all SUMO paralogs. Molecular basis for selection of the proper SUMO paralog-substrate pair by Ubc9 is currently not clear and may involve SUMO E3 ligases and other accessory proteins.

The discovery of SUMO E3 ligases has been confounded by the fact that all proteins reported to have SUMO E3 activity are dispensable for in vitro sumoylation [2]. This is in contrast to the ubiquitination pathway where substrate-specific E3 ligases are essential for the reaction to occur in vitro. By definition, SUMO E3 ligase must bind Ubc9, interact with the substrate, and enhance the transfer of SUMO from Ubc9 to substrate. Several proteins satisfying these criteria have been identified including RanBP2, the polycomb group protein Pc2, and Protein Inhibitors of Activated STATs (PIAS) proteins, PIAS1, PIAS3, PIASy, PIASx [2]. PIAS proteins in particular seem to fit the role of SUMO E3 ligases. They contain the RING domain, also found in many ubiquitin E3 ligases, localize to the nucleus, and their function in regulation of transcription is well-established [6]. However, PIAS proteins do not seem to provide desired specificity to the sumoylation process as they have overlapping substrate selection and can regulate transcription factor activity by SUMO-independent mechanisms.

SUMO-modification is an easily reversible process. For many proteins, sumoylated species cannot be detected on a western blot without the use of SUMO isopeptidase inhibitors such as N-ethylmaleimide. There are two SUMO isopeptidases with non-redundant function in yeast, Ulp1 and Ulp2, and in mammals, the family has expanded to seven members (SENPs) [7]. Ulp/SENPs have diverged from deubiquitination enzymes in structure and catalytic mechanism, and are related to viral cysteine proteases [2]. Their functions encompass processing SUMO precursors into mature forms and cleavage of the isopeptide bond between SUMOs and substrates or in SUMO2/3 chains. Specificity is built into SUMO processing and deconjugation as SENP1 will only act on SUMO1, but

not on SUMO2/3 [7]. Moreover, isopeptidases display distinct expression and subcellular localization patterns suggesting roles in substrate-specific regulation of sumoylation. SENP1 and SENP2 are localized to the nuclear pore complex, while SUMO3 and SUMO5 are found largely in the nucleolus [7].

Genetic studies demonstrate that sumoylation is essential for viability. In yeast, gene disruption of any of the sumoylation pathway components leads to cell cycle arrest and growth defects [8-11]. Ubc9 is also critical for mammalian embryonic development as mouse knockout of Ubc9 is early post-implantation lethal [12]. Additionally, genetic disruption of a mammalian SUMO isopeptidase, SENP1, leads to specific defects in Hypoxia-Inducible Factor 1 alpha function demonstrating that isopeptidases strongly contribute to substrate specificity in the sumoylation pathway [7, 13]. One study used a transgenic mouse complementation rescue approach to examine the role of transcription factor MafG sumoylation in megakaryocyte differentiation [14]. Transgenic mice expressing wild type or SUMO-mutant MafG protein in the *mafG*-null background were created, and sumoylation was shown to be essential for MafG to function as a repressor [14]. To date, no studies employing knock-in mouse technology to examine the role of sumoylation in regulating the activity of specific transcription factors under native conditions have been completed.

Sumoylation largely occurs in the nucleus where it is involved in a number of diverse processes such as chromosome structure maintenance and segregation, DNA mismatch repair, and nuclear transport [2]. Moreover, numerous transcription factors and

transcriptional regulators are modified with SUMO1 [15, 16]. Sumoylation for most transcription factors, including all nuclear hormone receptors, leads to repression of their transcriptional activity [15, 17, 18]. In cell culture reporter assays, mutating SUMO consensus lysines to arginines greatly increases transcriptional activity of both ligand-dependent and constitutively-active nuclear hormone receptors [18-21]. Intriguingly, this increase in activity displays cell-type and promoter specificity with greatest effects observed at promoters with multiple binding sites for steroid receptors [19, 22].

Diverse mechanisms have been proposed to explain the repressive effects of SUMO1 on transcription factor activity, and the two most researched mechanisms are described below. Sumoylation can intrinsically repress chromatin by recruiting histone deacetylases (HDACs). Promoter-targeting Ubc9 as a fusion to Gal4 DNA-binding domain results in sumoylation of histones, deacetylation and condensation of surrounding chromatin [23]. Similarly, SUMO1-modification of transcription factor Elk-1 enhances its interaction with HDAC-2 leading to deacetylation of Elk-1 target genes [24]. Additionally, sumoylation also inhibits transcription by sequestering transcription factors in subnuclear domains away from active chromatin. Following sumoylation, transcription factors p53, Lef-1, and Sp100 become concentrated in promyelocytic leukemia protein (PML) nuclear bodies or at the nuclear periphery [25-27]. PML protein non-covalently interacts with SUMO1, in addition to being sumoylated at two consensus sites [28, 29]. Thus, a model for formation of PML bodies has been proposed where PML protein can cluster sumoylated transcription factors through binding to SUMO1 [28, 29].

References

1. Herrmann, J., L.O. Lerman, and A. Lerman, *Ubiquitin and ubiquitin-like proteins in protein regulation.* Circ Res, 2007. **100**(9): p. 1276-91.
2. Johnson, E.S., *Protein modification by SUMO.* Annu Rev Biochem, 2004. **73**: p. 355-82.
3. Abgent-SUMOplot.
4. Saitoh, H. and J. Hinchey, *Functional heterogeneity of small ubiquitin-related protein modifiers SUMO-1 versus SUMO-2/3.* J Biol Chem, 2000. **275**(9): p. 6252-8.
5. Guo, D., et al., *A functional variant of SUMO4, a new I kappa B alpha modifier, is associated with type 1 diabetes.* Nat Genet, 2004. **36**(8): p. 837-41.
6. Sharrocks, A.D., *PIAS proteins and transcriptional regulation--more than just SUMO E3 ligases?* Genes Dev, 2006. **20**(7): p. 754-8.
7. Mukhopadhyay, D. and M. Dasso, *Modification in reverse: the SUMO proteases.* Trends Biochem Sci, 2007. **32**(6): p. 286-95.
8. Johnson, E.S. and G. Blobel, *Ubc9p is the conjugating enzyme for the ubiquitin-like protein Smt3p.* J Biol Chem, 1997. **272**(43): p. 26799-802.
9. Li, S.J. and M. Hochstrasser, *A new protease required for cell-cycle progression in yeast.* Nature, 1999. **398**(6724): p. 246-51.
10. Mossessova, E. and C.D. Lima, *Ulp1-SUMO crystal structure and genetic analysis reveal conserved interactions and a regulatory element essential for cell growth in yeast.* Mol Cell, 2000. **5**(5): p. 865-76.
11. Johnson, E.S., et al., *The ubiquitin-like protein Smt3p is activated for conjugation to other proteins by an Aos1p/Uba2p heterodimer.* Embo J, 1997. **16**(18): p. 5509-19.
12. Nacerddine, K., et al., *The SUMO pathway is essential for nuclear integrity and chromosome segregation in mice.* Dev Cell, 2005. **9**(6): p. 769-79.
13. Cheng, J., et al., *SUMO-specific protease 1 is essential for stabilization of HIF1alpha during hypoxia.* Cell, 2007. **131**(3): p. 584-95.
14. Motohashi, H., et al., *MafG sumoylation is required for active transcriptional repression.* Mol Cell Biol, 2006. **26**(12): p. 4652-63.
15. Gill, G., *Something about SUMO inhibits transcription.* Curr Opin Genet Dev, 2005. **15**(5): p. 536-41.
16. Verger, A., J. Perdomo, and M. Crossley, *Modification with SUMO. A role in transcriptional regulation.* EMBO Rep, 2003. **4**(2): p. 137-42.
17. Faus, H. and B. Haendler, *Post-translational modifications of steroid receptors.* Biomed Pharmacother, 2006. **60**(9): p. 520-8.
18. Pascual, G., et al., *A SUMOylation-dependent pathway mediates transrepression of inflammatory response genes by PPAR-gamma.* Nature, 2005. **437**(7059): p. 759-63.
19. Poukka, H., et al., *Covalent modification of the androgen receptor by small ubiquitin-like modifier 1 (SUMO-1).* Proc Natl Acad Sci U S A, 2000. **97**(26): p. 14145-50.
20. Tian, S., et al., *Small ubiquitin-related modifier-1 (SUMO-1) modification of the glucocorticoid receptor.* Biochem J, 2002. **367**(Pt 3): p. 907-11.

21. Chauchereau, A., et al., *Sumoylation of the progesterone receptor and of the steroid receptor coactivator SRC-1.* J Biol Chem, 2003. **278**(14): p. 12335-43.

22. Holmstrom, S., M.E. Van Antwerp, and J.A. Iniguez-Lluhi, *Direct and distinguishable inhibitory roles for SUMO isoforms in the control of transcriptional synergy.* Proc Natl Acad Sci U S A, 2003. **100**(26): p. 15758-63.

23. Shiio, Y. and R.N. Eisenman, *Histone sumoylation is associated with transcriptional repression.* Proc Natl Acad Sci U S A, 2003. **100**(23): p. 13225-30.

24. Yang, S.H. and A.D. Sharrocks, *SUMO promotes HDAC-mediated transcriptional repression.* Mol Cell, 2004. **13**(4): p. 611-7.

25. Sachdev, S., et al., *PIASy, a nuclear matrix-associated SUMO E3 ligase, represses LEF1 activity by sequestration into nuclear bodies.* Genes Dev, 2001. **15**(23): p. 3088-103.

26. Fogal, V., et al., *Regulation of p53 activity in nuclear bodies by a specific PML isoform.* Embo J, 2000. **19**(22): p. 6185-95.

27. Sternsdorf, T., K. Jensen, and H. Will, *Evidence for covalent modification of the nuclear dot-associated proteins PML and Sp100 by PIC1/SUMO-1.* J Cell Biol, 1997. **139**(7): p. 1621-34.

28. Duprez, E., et al., *SUMO-1 modification of the acute promyelocytic leukaemia protein PML: implications for nuclear localisation.* J Cell Sci, 1999. **112 (Pt 3)**: p. 381-93.

29. Shen, T.H., et al., *The mechanisms of PML-nuclear body formation.* Mol Cell, 2006. **24**(3): p. 331-9.

CHAPTER II

Chemistry of Nuclear Receptors

Lioudmila A. Campbell
Graduate Program in Biological Sciences, Department of Physiology,
University of California, San Francisco, California, USA.

Holly A. Ingraham
Departments of Cellular and Molecular Pharmacology, and Physiology
1550 4th Street, Rock Hall, Mission Bay Campus
University of California, San Francisco, California, USA 94143-261
holly.ingraham@ucsf.edu

Abstract:

Nuclear hormone receptors are integral players in endocrine networks, interfacing biology and chemistry. Unlike most other classes of transcription factors, these proteins are uniquely designed to bind small molecules, and thus affect gene expression in response to the cellular and organismal chemical environment. After several decades of research it is now appreciated that nuclear receptors bind very diverse lipophilic small molecules with a wide range of specificity and affinities. Recent nuclear receptor structures coupled with large-scale screening efforts challenge the dogma that all nuclear receptors, especially the large subset of constitutively active receptors, will have ligands and will represent tractable drug targets. As such, the "pharmacological future" for such orphan nuclear receptors may reside outside of the ligand-binding pocket.

Nuclear hormone receptors are classically defined as ligand-regulated transcription factors. The transcriptional programs affected by these proteins are linked to metabolic pathways, endocrine homeostasis, and organ development, and thus, both the loss and gain of function of these receptors are closely associated with a variety of human diseases including developmental and metabolic defects, cardiovascular disease, diabetes, reproductive failure, and cancer. Forty-eight nuclear receptors have been identified in the human genome and have been further classified into seven distinct subfamilies consisting of NR1, NR2, NR3, NR4, NR5, NR6 and NR0 based largely on sequence similarity in their two signature domains [1]. These two domains are present in almost all nuclear receptors and consist of the N-terminal DNA-binding domain (DBD) and the C-terminal

ligand-binding domain (LBD). The DBD interacts with specific DNA elements located in promoters of target genes, while the LBD binds hormones or other lipophilic molecules [2]. Additionally, receptors include two highly variable domains: the N-terminal domain preceding the DBD and the flexible hinge region between the DBD and the LBD. Currently, no pharmaceutical compound is directly targeted to the DBD or the flexible domains of any nuclear receptor.

Here we will focus on current progress in structural analyses of the nuclear hormone receptors, and how these proteins interact with their ligands, both natural and pharmaceutical. We will first provide a general overview of nuclear receptors and then using several nuclear receptors as examples, discuss the receptor-ligand specificity throughout the nuclear receptor superfamily and its implications for successful rational drug design to target the activity of these proteins. Additionally, we will review the emerging drug strategies that target regions outside of the ligand-binding pocket that might potentially provide new therapeutics aimed at this large family of receptors.

Overview

Nuclear receptors are sophisticated homeostatic sensors that function in the endocrine network of vertebrate organisms allowing for communication between or within different tissues and organs, often over large distances. These receptors are able to detect a constantly changing environment by binding small lipophilic hormones and metabolic intermediates. The ligand dependent feature of some nuclear receptors has been successfully exploited for therapeutic intervention against diseases such as breast cancer,

11

type 2 diabetes, and hypertension (Table 1) [3-5]. The use of nuclear receptors to mediate hormone signaling appears to have arisen late during metazoan evolution. Indeed, genome wide comparisons reveal that nuclear receptors are absent in some eukaryotic genomes. However, in those organisms other signaling pathways have been adapted to meet their endocrine needs and respond to small lipophilic molecules. For instance, no nuclear receptors have been identified in the yeast genome. Interestingly, a protein fold similar to the nuclear receptor LBD was identified by structural prediction in two transcription factors Oaf1 and Pip2 in the budding yeast *S. cerevisiae*. These transcription factors heterodimerize and bind the fatty acid oleate, reminiscent of the mammalian retinoic X receptor (RXR, NR2B)/ peroxisome proliferators-activated receptor (PPAR, NR1C) signaling pathway [6]. Similarly, hormone signaling in multi-cellular plants is not mediated by nuclear receptors despite the fact that sterols mediate many analogous functions in plant biology. Instead, plants appear to use other ligand binding motifs. For example, the growth promoting plant phytohormone brassinosteroid binds a cell surface receptor activating downstream kinases and ultimately Myc family transcription factors [7, 8]. Another large family of homeodomain-START (star-related lipid-transfer) proteins is hypothesized to directly affect gene expression after selectively binding sterols and lipids via the START domain [9, 10]. Collectively, these examples suggest a conserved signaling by lipophilic molecules using evolutionarily distinct binding proteins.

Ligand Activation of Nuclear Receptors

To carry out the transcriptional programs requiring both activation and repression of target genes, nuclear receptors interact with numerous coregulators, nucleating the assembly of macromolecular protein complexes that remodel chromatin and modulate transcription initiation or silencing [11, 12]. For ligand dependent receptors, the presence or absence of ligand determines the nature of the assembled protein complex. Given the importance of the LBD in binding ligand and interacting with coregulators, collective efforts of academia and industry have now elucidated LBD crystal structures for all seven subfamilies [13]. The nuclear receptor LBD structure is conserved and consists of an α-helical bundle ($\alpha1$-$\alpha12$), one to five β-strands, with three to four anti-parallel layers and a hydrophobic ligand-binding pocket occupying the core of the bundle [14]. The volume of this pocket varies greatly among the receptors, enabling these proteins to accommodate ligands of varying shapes and sizes (Figure 1) [13, 15]. In addition, the LBD contains a dimerization interface allowing receptors to bind DNA as homodimers or heterodimers [16]. And, for some receptors, it is also a key site for interaction with the heat-shock proteins [17].

Based on the first crystal structures of liganded nuclear receptors, the "mousetrap model" was proposed to account for ligand initiated activation [18]. Ligand was proposed to complete the hydrophobic core of the receptor, thus stabilizing the active conformation of the LBD. Concomitant with binding of the ligand, helix H12 containing the Activation Function 2 (AF2) undergoes a dramatic rearrangement, docking across the ligand-binding pocket and trapping the ligand inside [18]. This repositioning of helix H12 creates a new

hydrophobic surface [19, 20] that is bound by the LXXLL motif within coactivator proteins (Figure 2A) [21, 22]. Interestingly, corepressor proteins compete with coactivators for binding to the same hydrophobic groove but form a slightly extended surface eliminating the need for ligand [23, 24]. This mechanism allows the ligand to dictate nuclear receptor action by repositioning AF2 and thus shifting the equilibrium between coactivator and corepressor binding (Figure 2B) [25]. However, many nuclear receptors, especially orphan receptors, are constitutively active in the apparent absence of a ligand. Moreover, structural studies point to seemingly small receptor-specific differences within the LBDs that must underlie the diversity of receptor action in controlling distinct biological processes [13].

Thus far, about half of all nuclear receptors have been paired with physiological ligands and the other half remain orphaned, and either await identification of their native ligands or alternatively will never be bound by a ligand. For the most part, matching ligands with their cognate receptors has followed traditional drug discovery approaches using both cell based assays and biological clues. While nuclear receptors are readily found in tractable genetic model organisms, such as flies and worms [26], hunting for ligands by standard genetic screens has proven difficult and may reflect an overrepresentation of receptors belonging to the so called "orphan receptor" subfamilies in these invertebrate species. Exceptions include the discovery of the ecdysone hormone receptors and heme receptors in *Drosophila* [27] and 3-keto-sterols as ligands for the *C. elegans* nuclear hormone receptor DAF-12 involved in regulating lifespan [28, 29].

Despite the collective efforts of academia and pharmaceutical enterprises, ligands have remained elusive for a large number of receptors. A feature of these so called "orphan receptors" is their constitutive activity as evident by the robust, induced activity after overexpressing these receptors in a cellular reporter assay. In the absence of ligands, the best insights into the role of most of these receptors in vertebrate development and physiology comes from engineered mutants in mice or naturally occurring human mutations. Whether obligate ligands exist for the divergent but conserved ligand binding pockets of all nuclear receptors is highly debated. Moreover, there is no clear consensus on the evolution of ligand dependence, and two opposing hypotheses have been proposed. The first hypothesis suggests that primordial nuclear hormone receptors were ligand-independent, and regulation by specific, high affinity ligands evolved later, several times during the evolution of the nuclear receptor superfamily [30]. Consistent with this notion, many orphan receptor LBD structures reveal an active conformation with the AF2 containing helix H12 packed against the LBD with ligand-binding pockets that are either small or absent due to the presence of bulky hydrophobic amino acids (Figure 1C) [31]. These active but empty orphan receptors may represent an intermediate state as receptors were transitioning between ligand-independence to ligand-dependence [30].

The second hypothesis suggests that the ancestral nuclear receptors were ligand dependent, and throughout evolution particular receptors lost the need for ligand activation. The idea that primordial nuclear receptors were responsive to estrogens is consistent with this notion [32]. Moreover, the rodent lineage of the NR5A subfamily that includes Steroidogenic factor 1 and Liver receptor homolog 1 (SF-1, NR5A1 and

LRH-1, NR5A2) exhibits specific features that diminish ligand binding, suggesting that loss of ligand dependency occurred late in evolution. Structures of human and mouse NR5A subfamily LBDs revealed bacterial phospholipid ligands (Figure 1B) [33-36], except for the rodent LRH-1 where a key glycine residue has been replaced with a glutamate; the resulting salt bridge at the mouth of the ligand-binding pocket stabilizes the rodent LBD without a need for ligand. Taken together it suggests that the ancestral NR5A receptor was regulated by a ligand. Clearly the debate on whether nuclear receptors evolved to bind ligand or not is currently unresolved. Nonetheless, the collective structural and cellular data establish definitively that the binding capacity and ligand requirement vary drastically among the LBDs within all seven receptor subfamilies. More importantly, newly available structures of orphan nuclear receptor LBDs beg the question as to how tractable all nuclear receptors will be as drug targets? Below we will illustrate both the successes and challenges of ligand discovery for different nuclear receptor proteins.

ER: The Drug Target Darling

Given the significant role of steroid receptors in human biology and disease, especially breast and prostate cancer, it is not surprising that some of the first LBD crystal structures were those of the steroid receptors [19]. The steroid receptors were also the first to be targeted by pharmaceutical compounds, even prior to the availability of the high-resolution LBD structures that paved the way for structure-based drug design. The estrogen receptor (ER, NR3A) is the best example of successful manipulation of a nuclear receptor with synthetic ligands. Crystal structures of the ER LBD bound by

several distinct ligands reveal the exquisite specificity with which these ligands manipulate ER into active and inactive conformations. Each ER-ligand complex presents a distinct set of structural changes in the position of the AF2 relative to the core LBD, suggesting that standard approaches can be used in designing specific agonists or antagonists for this receptor. When bound by the natural ligand estradiol (E2), ER possesses a relatively small and well-defined ligand-binding pocket, and multiple contacts between the receptor and the ligand result in high specificity of interaction [37]. These features allow one to design ER modulatory ligands ranging from selective ER modulators (SERMs) such as tamoxifen, that exhibit mixed agonist/antagonist properties depending on the tissue or promoter, to complete antagonists like ICI 164,384 [38, 39]. In the latter case, the ER/ICI structure revealed how addition of bulky constituents to an agonist scaffold results in a protrusion from the ligand-binding pocket and movement of the AF2 helix into non-productive conformation, thus providing a paradigm for designing steroid nuclear receptor antagonists (Figure 2C) [39]. Regrettably, this approach has not worked for other ligand dependent receptors. Indeed, in a search for high affinity thyroid hormone receptor (TR, NR1A, Figure 1A) antagonist for treatment of hyperthyroidism, adding bulky constituents onto the endogenous TR ligand, triiodothyronine (T_3) does not create a true antagonist as would be predicted from studies on synthetic ER ligands [40, 41]. On the other hand, novel synthetic TR agonists have emerged based on the structure of T_3 complexed with the LBD [42].

The existence of SERMs raises some intriguing questions: what does the inactive LBD structure mean at a cellular level and do only active nuclear receptors interact with the

genome? Based on the ER LBD structures with tamoxifen and raloxifene, no productive interactions with coactivator proteins are possible because the AF2 helix adopts an inactive conformation [37, 38]; however, paradoxically, SERM-bound ER receptors retain transcriptional activity in certain tissues and on certain promoters. Thus, the diverse transcriptional programs for different SERMs, when profiled in an osteosarcoma cell line, are consistent with small overlap in tamoxifen and raloxifene regulated genes [43]. Similarly, an extremely small overlap was noted between groups of genes regulated by tamoxifen and E2 in a uterine cell line, despite the fact that tamoxifen is thought to be a partial agonist in this tissue, promoting endometrial cancer [44]. While further studies are needed, these results illustrate how ligands can dramatically alter gene expression. With the onset of new genome-wide technologies one can begin to examine how promoter occupancy is affected by ligands. Recent studies using chromatin immunoprecipitation combined with microarray analyses (ChIP/CHIP) reveal that many ER binding sites are located at a great distance from the proximal promoters and that some sites could be bound by the receptor even in the absence of E2 [45]. For receptors fortunate enough to have high affinity, specific ligands, as found for steroid receptors (NR3A, NR3C), the collective information gathered from these genome-wide approaches is likely to shed new insights into the physiological consequences of drug and provide for further refinement of drug structure.

PPAR and LXR: Orphans Adopted By Pharmaceuticals

Peroxisome Proliferator-Activated Receptor (PPAR, NR1C) and Liver X Receptor (LXR, NR1F) represent two clear examples where the lack of structural information did not

hinder the development of efficacious high affinity pharmaceutical compounds. PPAR and LXR are responsive to glucose and lipid levels, and play important roles in inflammation, cholesterol and lipid metabolism, and energy balance [46-48]. Despite the fact that natural ligands for PPAR remain controversial, with fatty acids and ecosinoids as the proposed low affinity endogenous ligands for PPARα [49], highly specific synthetic agonists and antagonists have been developed (Figure 2A) [24, 50]. Indeed, thiazolidinediones and the structurally related fibrates are used widely in treatment of diabetes and cardiovascular disease [51, 52].

Oxysterols are the proposed endogenous LXR ligands and are able to bind the ligand-binding pocket of LXR and activate its transcription in cellular assays [53, 54]. Additionally, genetic disruption of oxysterol biosynthesis in mice greatly attenuates LXR function [55]. Existing synthetic LXR agonists show potential in treating cardiovascular disease, although their collective role in controlling liver and gut metabolism may impose unwanted off-target effects [54, 56]. Oxysterols may not be the only endogenous ligands for LXR. Remarkably, a recent study reports that LXR also acts as a glucose sensor, where high concentrations of glucose (2 mM) displace oxysterols from the ligand-binding pocket, bind directly to the LBD, and also appear to act synergistically with the synthetic LXR ligand to affect endogenous target gene expression in the liver [57]. If true, LXR would be the first intracellular glucose sensor to be discovered and could provide a molecular explanation for the prominent linkage of diabetes with cardiovascular disease. Mechanistically, the authors suggest that glucose binds directly to the LXR LBD, perhaps in combination with oxysterol or alternatively binds elsewhere in the pocket or on the

solvent exposed surface of the LBD to allosterically modulate LXR activity [57, 58]. If structural analysis upholds the latter, it would raise an interesting dilemma – how would a hydrophilic molecule, such as glucose or the cellular glucose-6-phosphate, bind tightly into the hydrophobic pocket of LXR.   Nonetheless, this finding is provocative and potentially provides a new paradigm for targeting nuclear receptors.

NR5A Receptors: Large Pockets In Search of Large Ligands

The NR5A subfamily of nuclear hormone receptors includes LRH-1 and SF-1, as well as the *Drosophila* nuclear receptor Ftz-F1.   SF-1 is required for endocrine tissue development and sexual differentiation, and is a major regulator of steroid biosynthesis [59].   LRH-1 is essential in embryonic development, and, in adult, regulates bile acid production, cholesterol transport, and ovarian function [60].   All LBD crystal structures of murine and human members of this subfamily revealed large ligand-binding pockets and structural inflexibility as evidenced by the minimal changes observed with or without ligand or coactivator peptide [33-36, 61].   The overall stability of the NR5A subfamily can be partially explained by the presence of an additional stabilizing layer due to a well-formed and elongated helix H2.

Phospholipids were found in the ligand-binding pockets of mouse and human SF-1 and human LRH-1 and are relatively large (~750Da) compared to other ligands such as steroid derivatives.   The lipid tails fit exceptionally well into the ligand-binding pocket and make a number of specific contacts with helix H12 and the hydrophobic cavity (Figure 1B) [33, 34].   In addition to being integral membrane components, phospholipids

20

also bind in the ligand pocket of START domain and in phosphatidyl inositol transport proteins [62]. All NR5A receptors exhibit constitutive activity in cells, thus it is unclear whether these ligands simply serve to stabilize the LBD helical bundle or whether they act as regulatory ligands. Notably, filling the pocket with bulky residues diminishes ligand uptake in biochemical assays [33] and also attenuates transcriptional activity in cells [34-36]. The challenge in designing synthetic ligands for NR5A receptors is two fold. First, finding a ligand that recapitulates the positioning of the acyl chains and the phosphate head group might be problematic and second, while the SF-1 LBD protein readily exchanges the bacterial phosphatidyl glycerol with PIP3 or PIP2 (HAI, unpublished data), and might naturally be bound by phosphatidic acid [63], displacing the endogenous phospholipid with a small molecule in a cellular environment might prove difficult. However, a recent report describes a small molecule that at nanomolar concentrations promotes coactivator peptide recruitment to SF-1 and LRH-1, displaces the phospholipid ligand, and evokes a modest increase in endogenous target gene activation in human hepatocytes [64]. These studies suggest that, perhaps, this family of nuclear receptors is still tractable for drug discovery.

PXR and CAR: Too Much Receptor For A Single Ligand

Pregnane X receptor (PXR, NR1I2) and the constitutive androstane receptor (CAR, NR1I3) are highly promiscuous nuclear receptors that bind a variety of structurally diverse compounds. Thus there is no difficulty in identifying ligands for these receptors – finding highly specific ligands appears to be the challenge. This is especially true for PXR and most likely reflects its role in the xenobiotic response. A number of different

compounds are accommodated in its ligand-binding pocket ranging from small hydrophobic drugs to the large antibiotic rifampicin [65-67]. Five expandable β-sheets, unique to PXR, allow for this dramatic increase in the size of the ligand-binding pocket [65]. Interestingly, a similar structural feature is also found in START domain proteins and may represent a critical structural arrangement for binding a wide variety of lipophilic molecules [62]. CAR exhibits a large, but empty ligand-binding pocket and high constitutive activity that results from two structural features: αX helix preceding helix H12 that stabilizes AF2 in an active conformation and an extended helix H2, similar to the NR5A receptors [68-72]. For an organism, the promiscuity of PXR and CAR activation is an indispensable feature because it assures protection from a variety of harmful xenobiotics and metabolites. However, this characteristic also presents a formidable challenge to rational drug design. Once again, and as found with TR, bulky constituents added onto existing PXR agonist scaffolds fail to yield suitable antagonists [73]. For CAR, it appears that a significant mode of regulation occurs by shuttling between the nucleus and cytoplasm rather than by ligand activation [74]. Interestingly, inverse agonists or ligands that reduce the constitutive activity of CAR have been reported [68, 75]. Whether natural ligands exist for PXR and CAR remains unclear, and it may be more likely that these receptors are designed to constantly sample their chemical environment in order to protect the organism from harmful cellular metabolites or environmental toxins.

True Orphans Without Pockets

Finally, structural information on other receptor subfamilies reveals some receptors either to be complexed with "structural non-exchangeable ligands" or to simply have inadequate capacity in their pockets to accommodate the smallest of ligands. To date, hepatocyte nuclear factor 4 (HNF4, NR2A) is an example of a receptor with a structural ligand. Structures of the rat HNF4α (NR2A1) and human HNF4γ (NR2A2) LBDs showed a mixture of bacterial fatty acids occupying the ligand-binding pocket [76, 77]. Although HNF4 is found complexed with only a small selection of fatty acids among an assortment of many, this fatty acid ligand is completely entrenched in the ligand-binding pocket and is dislodged only after complete denaturation of the protein, suggesting that in vitro approaches to ligand identification may not be feasible. Similar to this finding, the phospholipid ligand in human LRH-1 is also resistant to in vitro exchange with other phospholipids, perhaps, presenting another case of a structural ligand for a nuclear receptor [35].

As mentioned above, the lack of a conventional hydrophobic cavity makes the ligand hunt extremely difficult. Two receptor subfamilies appear to be "pocketless," including members of NR4A ([NGFI-B/Nurr77, NR4A1], [Nurr1, NR4A2], [NOR1, NR4A3]) and their fly ortholog DHR38, and members of the NR0 subfamily including Dax-1 (NR0B1). All three structures of NR4 LBDs adopt a canonical protein fold but lack any ligand-binding pocket due to obstruction by bulky side chains (Figure 1C). These LBDs also lack a hydrophobic coactivator cleft that is instead replaced with a charged surface [78-80]. Another case of an empty pocket is the new structure of the atypical orphan nuclear receptor Dax-1 complexed with LRH-1. Both Dax-1 and SHP (NR0B2) lack a DBD

altogether, and thus rely on interactions with other NRs and transcription factors to be recruited to the DNA, but both are potent repressors in cellular reporter assays [81, 82]. From the crystal structure it is evident that the ligand-binding pocket of Dax-1 (80 Å^3) is unable to accommodate even the smallest ligand (E. Sablin and R. J. Fletterick, personal communication). Based on this structure and given the high identity with Dax-1, SHP is also predicted to be refractory to ligand regulation.

Finding pharmaceutical ligands for receptors with very small pockets still remains a feasible option as illustrated by recent discovery of a synthetic agonist for the estrogen-related receptor γ (ERRγ, NR3B3). ERRγ is a constitutively active nuclear receptor with no known natural ligand, and the crystal structure of the ERRγ LBD revealed an extremely small ligand-binding pocket (220 Å^3) [83]. Remarkably, in a new crystal structure of ERRγ LBD with a synthetic agonist, GSK4716, the ligand-binding pocket expanded to a notable 610 Å^3 [84]. This result underscores the ability of the LBD to accommodate ligands of varying size, and suggests that continuing the hunt for ligands might yield some future surprises.

Alternative Surfaces for Regulation

Despite the fact that ligand discovery has historically focused on the ligand-binding domain, emerging data suggest that alternative surfaces might be targeted to regulate receptor activity. Alternative binding surfaces have been suggested by structural studies on the NR4A subfamily member, Nurr1. Nuclear magnetic resonance footprinting

studies of Nurr1 LBD with peptides derived from the nuclear receptor corepressor (NCoR) and a related corepressor SMRT identified a hydrophobic binding site on the surface of the LBD between helices H11 and H12 [85]. Mutational disruption of this interaction surface abolished transcriptional activity of Nurr1 underscoring its importance in Nurr1 function. Since the canonical coactivator groove is absent in the NR4A subfamily, this additional LBD surface is possibly the major site for interaction with the coregulators. On that note, it is of interest that crystal structures of the rat farnesoid X receptor LBD (FXR, NR1H4) and the human LRH-1 LBD revealed two coactivator LXXLL peptides bound to the receptor [33, 36, 86]; in these cases, the relevance of this additional bound peptide remains to be determined.

New pharmaceuticals might act by covalent modification of a key protein-protein interaction surface, by blocking an interaction surface, or by allosterically affecting the ligand-binding pocket, as suggested for glucose binding to LXR. Presumably, for the majority of these interactions, one would disrupt the assembly of receptor-coregulator complexes and in essence mimic conventional antagonists [87, 88]. Rodriguez and colleagues synthesized a small molecule inhibitor of coactivator binding that structurally mimics key contacts of a coactivator LXXLL motif with the hydrophobic binding groove of the nuclear receptor [89]. The authors used a crystal structure of agonist-bound ERα in a complex with a coactivator peptide to guide small molecule design followed by a screen to identify molecules that abolish peptide recruitment but do not directly compete with ligand binding. A similar high-throughput approach was used to identify novel covalent inhibitors of TRβ (NR1A2), β-aminoketones. These inhibitors irreversibly react

with a cysteine residue located in the coactivator groove of TRβ LBD thus disrupting the interaction between TRβ and an LXXLL-containing coactivator peptide [90]. TRb has multiple solvent-exposed cysteines on the LBD, yet these compounds show high selectivity towards a single residue, unique to the TR family of receptors. Additionally, some of the tested compounds appear to be isoform-specific, demonstrating vastly different affinities for TRα and TRβ. Similar to these findings, 4-hydroxytamoxifen (OHT) was found to inhibit coactivator recruitment to ERβ and surprisingly, the crystal structure of the ERβ LBD revealed two bound OHT molecules [91]. One molecule was bound in the ligand-binding pocket, and another molecule was revealed in the coactivator groove, displacing the AF2 away from the LBD, into inactive conformation [38]. While the exact contribution of this external OHT binding site to the antagonistic effects of OHT on ERβ function is unclear, this binding event could be uncovering a subtle structural difference between the two ER isoforms. Finally, another allosteric inhibitor compound has been identified for the androgen receptor (AR, NR3C4). It shows reversible binding at a novel hydrophobic LBD surface, conserved in other steroid receptors, and this binding allosterically moves the AF2 helix into an inactive conformation (E. Estébanez-Perpiñá and R. J. Fletterick, personal communication). Collectively, these studies raise the possibility that new drugs may emerge that target additional surfaces other than the hydrophobic ligand-binding pocket.

Summary and Future Directions

The ability of nuclear hormone receptors to bind small molecules with high affinity and high specificity places them squarely at the interface between biology and chemistry. As

such, the nuclear receptor field has historically been focused primarily on the identification of regulatory ligands. Now, an alternative approach is needed for those receptors that fail to exhibit classic ligand dependency, but instead appear to be ligand-independent. Domains outside the DBD and the LBD, especially the Activation Function 1 (AF1) offer a regulatory platform for multiple posttranslational events and coregulator interactions [92, 93]. Positioning of the AF1 varies among nuclear receptors, suggesting that it has hopped around throughout evolution, and can be found in the variable N-terminal extension preceding the DBD [94-97] for steroid receptors or in the hinge region close to the LBD for the NR4 and NR5 subfamilies [98-101]. Both the N-terminal extension and the hinge region are highly variable in length and sequence, and most likely disordered and flexible, thus making their structural determination elusive. Multiple sites of posttranslational modifications are found in these variable regions, such as phosphorylation, sumoylation, acetylation, and ubiquitination, and these sites of modification often cluster closely together. Additionally, the AF1 appears to be a major surface for interaction with numerous coregulator proteins [102-106]. Considering the importance of AF1 in nuclear receptor activity, its structure, function, and interaction with the DBD and the LBD are still poorly understood, and there are no pharmaceuticals available for direct manipulation of AF1 function. For the ligand-independent receptors, such as the NR4 subfamily, posttranslational modifications might be crucial in regulating their activity [107, 108].

The interplay between posttranslational modifications and the ligand potentially leads to a myriad of functional outcomes for the nuclear receptors. We are only beginning to map

out the relationships between individual posttranslational events and to understand the specific effects of their combinations on receptor activity. Numerous studies highlight the importance of "the histone code" or how posttranslational modifications of histone proteins affect transcriptional state of the chromatin and dictate transcriptional competency of genes. The abundance of posttranslational modifications on nuclear receptors suggests a similar idea of regulation.

From the extensive cellular, biochemical and structural studies carried out on nuclear hormone receptors it is now appreciated that their ability to be "classically" regulated by ligands is no longer taken for granted. Indeed, we now know that over half of these receptors are not regulated by ligands as discovered for the steroid receptors many decades ago. In the last ten years, intensive research has focused on the so called "orphan receptors" with the goal of finding their high affinity ligands. Now, it is realized that many receptors cannot be bound by ligand or have a non-exchangeable "structural" ligand embedded in their pockets. For ligand-dependent receptors, the challenge for the next decade will be to refine the specificity of the existing known ligands or identify allosteric modulatory ligands. For ligand-independent receptors, research will have to take a new direction to identify other regulatory sites that can then be targeted by small molecules. Given the importance of nuclear receptors in human biology and disease, they are likely to remain a primary focus for both academia and industry for years to come.

References

1.   Germain, P., et al., *Overview of nomenclature of nuclear receptors.* Pharmacol Rev, 2006. **58**(4): p. 685-704.
2.   Bain, D.L., et al., *Nuclear Receptor Structure: Implications for Function.* Annu Rev Physiol, 2006.
3.   Jordan, V.C., *Antiestrogenic action of raloxifene and tamoxifen: today and tomorrow.* J Natl Cancer Inst, 1998. **90**(13): p. 967-71.
4.   Day, C., *Thiazolidinediones: a new class of antidiabetic drugs.* Diabet Med, 1999. **16**(3): p. 179-92.
5.   Baxter, J.D., et al., *Towards selectively modulating mineralocorticoid receptor function: lessons from other systems.* Mol Cell Endocrinol, 2004. **217**(1-2): p. 151-65.
6.   Phelps, C., et al., *Fungi and animals may share a common ancestor to nuclear receptors.* Proc Natl Acad Sci U S A, 2006. **103**(18): p. 7077-81.
7.   Belkhadir, Y., X. Wang, and J. Chory, *Brassinosteroid signaling pathway.* Sci STKE, 2006. **2006**(364): p. cm4.
8.   Vert, G., et al., *Molecular mechanisms of steroid hormone signaling in plants.* Annu Rev Cell Dev Biol, 2005. **21**: p. 177-201.
9.   Ponting, C.P. and L. Aravind, *START: a lipid-binding domain in StAR, HD-ZIP and signalling proteins.* Trends Biochem Sci, 1999. **24**(4): p. 130-2.
10.   Schrick, K., et al., *START lipid/sterol-binding domains are amplified in plants and are predominantly associated with homeodomain transcription factors.* Genome Biol, 2004. **5**(6): p. R41.
11.   Freedman, L.P., *Increasing the complexity of coactivation in nuclear receptor signaling.* Cell, 1999. **97**(1): p. 5-8.
12.   Nettles, K.W. and G.L. Greene, *Ligand control of coregulator recruitment to nuclear receptors.* Annu Rev Physiol, 2005. **67**: p. 309-33.
13.   Ingraham, H.A. and M.R. Redinbo, *Orphan nuclear receptors adopted by crystallography.* Curr Opin Struct Biol, 2005. **15**(6): p. 708-15.
14.   Bourguet, W., et al., *Crystal structure of the ligand-binding domain of the human nuclear receptor RXR-alpha.* Nature, 1995. **375**(6530): p. 377-82.
15.   DeLano, W.L., *MacPyMOL: A PyMOL-based Molecular Graphics Application for MacOS X.* DeLano Scientific LLC, Palo Alto, CA, USA. , 2007.
16.   Glass, C.K., *Differential recognition of target genes by nuclear receptor monomers, dimers, and heterodimers.* Endocr Rev, 1994. **15**(3): p. 391-407.
17.   Pratt, W.B. and D.O. Toft, *Steroid receptor interactions with heat shock protein and immunophilin chaperones.* Endocr Rev, 1997. **18**(3): p. 306-60.
18.   Renaud, J.P., et al., *Crystal structure of the RAR-gamma ligand-binding domain bound to all-trans retinoic acid.* Nature, 1995. **378**(6558): p. 681-9.
19.   Weatherman, R.V., R.J. Fletterick, and T.S. Scanlan, *Nuclear-receptor ligands and ligand-binding domains.* Annu Rev Biochem, 1999. **68**: p. 559-81.
20.   Darimont, B.D., et al., *Structure and specificity of nuclear receptor-coactivator interactions.* Genes Dev, 1998. **12**(21): p. 3343-56.

21. Heery, D.M., et al., *A signature motif in transcriptional co-activators mediates binding to nuclear receptors.* Nature, 1997. **387**(6634): p. 733-6.

22. Glass, C.K., D.W. Rose, and M.G. Rosenfeld, *Nuclear receptor coactivators.* Curr Opin Cell Biol, 1997. **9**(2): p. 222-32.

23. Baniahmad, A., *Nuclear hormone receptor co-repressors.* J Steroid Biochem Mol Biol, 2005. **93**(2-5): p. 89-97.

24. Xu, H.E., et al., *Structural basis for antagonist-mediated recruitment of nuclear co-repressors by PPARalpha.* Nature, 2002. **415**(6873): p. 813-7.

25. Glass, C.K. and M.G. Rosenfeld, *The coregulator exchange in transcriptional functions of nuclear receptors.* Genes Dev, 2000. **14**(2): p. 121-41.

26. King-Jones, K. and C.S. Thummel, *Nuclear receptors--a perspective from Drosophila.* Nat Rev Genet, 2005. **6**(4): p. 311-23.

27. Billas, I.M. and D. Moras, *Ligand-binding pocket of the ecdysone receptor.* Vitam Horm, 2005. **73**: p. 101-29.

28. Motola, D.L., et al., *Identification of ligands for DAF-12 that govern dauer formation and reproduction in C. elegans.* Cell, 2006. **124**(6): p. 1209-23.

29. Beckstead, R.B. and C.S. Thummel, *Indicted: worms caught using steroids.* Cell, 2006. **124**(6): p. 1137-40.

30. Escriva, H., F. Delaunay, and V. Laudet, *Ligand binding and nuclear receptor evolution.* Bioessays, 2000. **22**(8): p. 717-27.

31. Privalsky, M.L., *Activation incarnate.* Dev Cell, 2003. **5**(1): p. 1-2.

32. Thornton, J.W., E. Need, and D. Crews, *Resurrecting the ancestral steroid receptor: ancient origin of estrogen signaling.* Science, 2003. **301**(5640): p. 1714-7.

33. Krylova, I.N., et al., *Structural analyses reveal phosphatidyl inositols as ligands for the NR5 orphan receptors SF-1 and LRH-1.* Cell, 2005. **120**(3): p. 343-55.

34. Li, Y., et al., *Crystallographic identification and functional characterization of phospholipids as ligands for the orphan nuclear receptor steroidogenic factor-1.* Mol Cell, 2005. **17**(4): p. 491-502.

35. Ortlund, E.A., et al., *Modulation of human nuclear receptor LRH-1 activity by phospholipids and SHP.* Nat Struct Mol Biol, 2005. **12**(4): p. 357-63.

36. Wang, W., et al., *The crystal structures of human steroidogenic factor-1 and liver receptor homologue-1.* Proc Natl Acad Sci U S A, 2005. **102**(21): p. 7505-10.

37. Brzozowski, A.M., et al., *Molecular basis of agonism and antagonism in the oestrogen receptor.* Nature, 1997. **389**(6652): p. 753-8.

38. Shiau, A.K., et al., *The structural basis of estrogen receptor/coactivator recognition and the antagonism of this interaction by tamoxifen.* Cell, 1998. **95**(7): p. 927-37.

39. Pike, A.C., et al., *Structural insights into the mode of action of a pure antiestrogen.* Structure, 2001. **9**(2): p. 145-53.

40. Sandler, B., et al., *Thyroxine-thyroid hormone receptor interactions.* J Biol Chem, 2004. **279**(53): p. 55801-8.

41. Borngraeber, S., et al., *Ligand selectivity by seeking hydrophobicity in thyroid hormone receptor.* Proc Natl Acad Sci U S A, 2003. **100**(26): p. 15358-63.

42. Chiellini, G., et al., *A high-affinity subtype-selective agonist ligand for the thyroid hormone receptor.* Chem Biol, 1998. **5**(6): p. 299-306.

43. Kian Tee, M., et al., *Estradiol and selective estrogen receptor modulators differentially regulate target genes with estrogen receptors alpha and beta.* Mol Biol Cell, 2004. **15**(3): p. 1262-72.

44. Wu, H., et al., *Hypomethylation-linked activation of PAX2 mediates tamoxifen-stimulated endometrial carcinogenesis.* Nature, 2005. **438**(7070): p. 981-7.

45. Carroll, J.S., et al., *Chromosome-wide mapping of estrogen receptor binding reveals long-range regulation requiring the forkhead protein FoxA1.* Cell, 2005. **122**(1): p. 33-43.

46. Willson, T.M., et al., *The PPARs: from orphan receptors to drug discovery.* J Med Chem, 2000. **43**(4): p. 527-50.

47. Michalik, L., et al., *International Union of Pharmacology. LXI. Peroxisome proliferator-activated receptors.* Pharmacol Rev, 2006. **58**(4): p. 726-41.

48. Li, A.C. and C.K. Glass, *PPAR- and LXR-dependent pathways controlling lipid metabolism and the development of atherosclerosis.* J Lipid Res, 2004. **45**(12): p. 2161-73.

49. Kliewer, S.A., et al., *Fatty acids and eicosanoids regulate gene expression through direct interactions with peroxisome proliferator-activated receptors alpha and gamma.* Proc Natl Acad Sci U S A, 1997. **94**(9): p. 4318-23.

50. Xu, H.E., et al., *Structural determinants of ligand binding selectivity between the peroxisome proliferator-activated receptors.* Proc Natl Acad Sci U S A, 2001. **98**(24): p. 13919-24.

51. Nolte, R.T., et al., *Ligand binding and co-activator assembly of the peroxisome proliferator-activated receptor-gamma.* Nature, 1998. **395**(6698): p. 137-43.

52. Xu, H.E., et al., *Molecular recognition of fatty acids by peroxisome proliferator-activated receptors.* Mol Cell, 1999. **3**(3): p. 397-403.

53. Janowski, B.A., et al., *An oxysterol signalling pathway mediated by the nuclear receptor LXR alpha.* Nature, 1996. **383**(6602): p. 728-31.

54. Williams, S., et al., *X-ray crystal structure of the liver X receptor beta ligand binding domain: regulation by a histidine-tryptophan switch.* J Biol Chem, 2003. **278**(29): p. 27138-43.

55. Chen, W., et al., *Enzymatic reduction of oxysterols impairs LXR signaling in cultured cells and the livers of mice.* Cell Metab, 2007. **5**(1): p. 73-9.

56. Farnegardh, M., et al., *The three-dimensional structure of the liver X receptor beta reveals a flexible ligand-binding pocket that can accommodate fundamentally different ligands.* J Biol Chem, 2003. **278**(40): p. 38821-8.

57. Mitro, N., et al., *The nuclear receptor LXR is a glucose sensor.* Nature, 2007. **445**(7124): p. 219-23.

58. Lazar, M.A. and T.M. Willson, *Sweet Dreams for LXR.* Cell Metab, 2007. **5**(3): p. 159-61.

59. Val, P., et al., *SF-1 a key player in the development and differentiation of steroidogenic tissues.* Nucl Recept, 2003. **1**(1): p. 8.

60. Fayard, E., J. Auwerx, and K. Schoonjans, *LRH-1: an orphan nuclear receptor involved in development, metabolism and steroidogenesis.* Trends Cell Biol, 2004. **14**(5): p. 250-60.

61. Sablin, E.P., et al., *Structural basis for ligand-independent activation of the orphan nuclear receptor LRH-1.* Mol Cell, 2003. **11**(6): p. 1575-85.

62. Tsujishita, Y. and J.H. Hurley, *Structure and lipid transport mechanism of a StAR-related domain.* Nat Struct Biol, 2000. **7**(5): p. 408-14.

63. Li, D., et al., *cAMP-Stimulated Interaction Between Steroidogenic Factor-1 and Diacylglycerol Kinase-{theta} Facilitates Induction of CYP17.* Mol Cell Biol, 2007.

64. Whitby, R.J., et al., *Identification of small molecule agonists of the orphan nuclear receptors liver receptor homolog-1 and steroidogenic factor-1.* J Med Chem, 2006. **49**(23): p. 6652-5.

65. Watkins, R.E., et al., *The human nuclear xenobiotic receptor PXR: structural determinants of directed promiscuity.* Science, 2001. **292**(5525): p. 2329-33.

66. Chrencik, J.E., et al., *Structural disorder in the complex of human pregnane X receptor and the macrolide antibiotic rifampicin.* Mol Endocrinol, 2005. **19**(5): p. 1125-34.

67. Watkins, R.E., et al., *2.1 A crystal structure of human PXR in complex with the St. John's wort compound hyperforin.* Biochemistry, 2003. **42**(6): p. 1430-8.

68. Shan, L., et al., *Structure of the murine constitutive androstane receptor complexed to androstenol: a molecular basis for inverse agonism.* Mol Cell, 2004. **16**(6): p. 907-17.

69. Suino, K., et al., *The nuclear xenobiotic receptor CAR: structural determinants of constitutive activation and heterodimerization.* Mol Cell, 2004. **16**(6): p. 893-905.

70. Xu, R.X., et al., *A structural basis for constitutive activity in the human CAR/RXRalpha heterodimer.* Mol Cell, 2004. **16**(6): p. 919-28.

71. Maglich, J.M., et al., *Identification of a novel human constitutive androstane receptor (CAR) agonist and its use in the identification of CAR target genes.* J Biol Chem, 2003. **278**(19): p. 17277-83.

72. Haffner, C.D., et al., *Structure-based design of potent retinoid X receptor alpha agonists.* J Med Chem, 2004. **47**(8): p. 2010-29.

73. Xue, Y., et al., *Crystal structure of the PXR-T1317 complex provides a scaffold to examine the potential for receptor antagonism.* Bioorg Med Chem, 2007. **15**(5): p. 2156-66.

74. Swales, K. and M. Negishi, *CAR, driving into the future.* Mol Endocrinol, 2004. **18**(7): p. 1589-98.

75. Honkakoski, P., et al., *The nuclear orphan receptor CAR-retinoid X receptor heterodimer activates the phenobarbital-responsive enhancer module of the CYP2B gene.* Mol Cell Biol, 1998. **18**(10): p. 5652-8.

76. Dhe-Paganon, S., et al., *Crystal structure of the HNF4 alpha ligand binding domain in complex with endogenous fatty acid ligand.* J Biol Chem, 2002. **277**(41): p. 37973-6.

77. Wisely, G.B., et al., *Hepatocyte nuclear factor 4 is a transcription factor that constitutively binds fatty acids.* Structure, 2002. **10**(9): p. 1225-34.

78. Wang, Z., et al., *Structure and function of Nurr1 identifies a class of ligand-independent nuclear receptors.* Nature, 2003. **423**(6939): p. 555-60.

79. Baker, K.D., et al., *The Drosophila orphan nuclear receptor DHR38 mediates an atypical ecdysteroid signaling pathway.* Cell, 2003. **113**(6): p. 731-42.

80. Flaig, R., et al., *Structural basis for the cell-specific activities of the NGFI-B and the Nurr1 ligand-binding domain.* J Biol Chem, 2005. **280**(19): p. 19250-8.

81. Niakan, K.K. and E.R. McCabe, *DAX1 origin, function, and novel role.* Mol Genet Metab, 2005. **86**(1-2): p. 70-83.

82. Bavner, A., et al., *Transcriptional corepression by SHP: molecular mechanisms and physiological consequences.* Trends Endocrinol Metab, 2005. **16**(10): p. 478-88.

83. Greschik, H., et al., *Structural and functional evidence for ligand-independent transcriptional activation by the estrogen-related receptor 3.* Mol Cell, 2002. **9**(2): p. 303-13.

84. Wang, L., et al., *X-ray crystal structures of the estrogen-related receptor-gamma ligand binding domain in three functional states reveal the molecular basis of small molecule regulation.* J Biol Chem, 2006. **281**(49): p. 37773-81.

85. Codina, A., et al., *Identification of a novel co-regulator interaction surface on the ligand binding domain of Nurr1 using NMR footprinting.* J Biol Chem, 2004. **279**(51): p. 53338-45.

86. Mi, L.Z., et al., *Structural basis for bile acid binding and activation of the nuclear receptor FXR.* Mol Cell, 2003. **11**(4): p. 1093-100.

87. Leduc, A.M., et al., *Helix-stabilized cyclic peptides as selective inhibitors of steroid receptor-coactivator interactions.* Proc Natl Acad Sci U S A, 2003. **100**(20): p. 11273-8.

88. Geistlinger, T.R. and R.K. Guy, *Novel selective inhibitors of the interaction of individual nuclear hormone receptors with a mutually shared steroid receptor coactivator 2.* J Am Chem Soc, 2003. **125**(23): p. 6852-3.

89. Rodriguez, A.L., et al., *Design, synthesis, and in vitro biological evaluation of small molecule inhibitors of estrogen receptor alpha coactivator binding.* J Med Chem, 2004. **47**(3): p. 600-11.

90. Arnold, L.A., et al., *Discovery of small molecule inhibitors of the interaction of the thyroid hormone receptor with transcriptional coregulators.* J Biol Chem, 2005. **280**(52): p. 43048-55.

91. Wang, Y., et al., *A second binding site for hydroxytamoxifen within the coactivator-binding groove of estrogen receptor beta.* Proc Natl Acad Sci U S A, 2006. **103**(26): p. 9908-11.

92. Warnmark, A., et al., *Activation functions 1 and 2 of nuclear receptors: molecular strategies for transcriptional activation.* Mol Endocrinol, 2003. **17**(10): p. 1901-9.

93. McEwan, I.J., *Molecular mechanisms of androgen receptor-mediated gene regulation: structure-function analysis of the AF-1 domain.* Endocr Relat Cancer, 2004. **11**(2): p. 281-93.

94. Miesfeld, R., et al., *Glucocorticoid receptor mutants that define a small region sufficient for enhancer activation.* Science, 1987. **236**(4800): p. 423-7.

95.   Dahlman-Wright, K., et al., *Delineation of a small region within the major transactivation domain of the human glucocorticoid receptor that mediates transactivation of gene expression.* Proc Natl Acad Sci U S A, 1994. **91**(5): p. 1619-23.

96.   Hadzopoulou-Cladaras, M., et al., *Functional domains of the nuclear receptor hepatocyte nuclear factor 4.* J Biol Chem, 1997. **272**(1): p. 539-50.

97.   He, B. and E.M. Wilson, *The NH(2)-terminal and carboxyl-terminal interaction in the human androgen receptor.* Mol Genet Metab, 2002. **75**(4): p. 293-8.

98.   Desclozeaux, M., et al., *Phosphorylation and intramolecular stabilization of the ligand binding domain in the nuclear receptor steroidogenic factor 1.* Mol Cell Biol, 2002. **22**(20): p. 7193-203.

99.   Wansa, K.D., J.M. Harris, and G.E. Muscat, *The activation function-1 domain of Nur77/NR4A1 mediates trans-activation, cell specificity, and coactivator recruitment.* J Biol Chem, 2002. **277**(36): p. 33001-11.

100.   Wansa, K.D., et al., *The AF-1 domain of the orphan nuclear receptor NOR-1 mediates trans-activation, coactivator recruitment, and activation by the purine anti-metabolite 6-mercaptopurine.* J Biol Chem, 2003. **278**(27): p. 24776-90.

101.   Maira, M., et al., *Dimer-specific potentiation of NGFI-B (Nur77) transcriptional activity by the protein kinase A pathway and AF-1-dependent coactivator recruitment.* Mol Cell Biol, 2003. **23**(3): p. 763-76.

102.   Faus, H. and B. Haendler, *Post-translational modifications of steroid receptors.* Biomed Pharmacother, 2006. **60**(9): p. 520-8.

103.   Jackson, T.A., et al., *The partial agonist activity of antagonist-occupied steroid receptors is controlled by a novel hinge domain-binding coactivator L7/SPA and the corepressors N-CoR or SMRT.* Mol Endocrinol, 1997. **11**(6): p. 693-705.

104.   Chen, M., et al., *Phosphorylation of the liver X receptors.* FEBS Lett, 2006. **580**(20): p. 4835-41.

105.   Garcia-Pedrero, J.M., et al., *The SWI/SNF chromatin remodeling subunit BAF57 is a critical regulator of estrogen receptor function in breast cancer cells.* J Biol Chem, 2006. **281**(32): p. 22656-64.

106.   Ou, Q., et al., *The DEAD box protein DP103 is a regulator of steroidogenic factor-1.* Mol Endocrinol, 2001. **15**(1): p. 69-79.

107.   Galleguillos, D., et al., *PIASgamma represses the transcriptional activation induced by the nuclear receptor Nurr1.* J Biol Chem, 2004. **279**(3): p. 2005-11.

108.   Wingate, A.D., et al., *Nur77 is phosphorylated in cells by RSK in response to mitogenic stimulation.* Biochem J, 2006. **393**(Pt 3): p. 715-24.

Cross-references

Liver X Receptors (LXR), chemistry of
Peroxisome Proliferator-Activated Receptor (PPAR), chemistry of
Steroid hormones, chemistry of
Orphan nuclear hormone receptors
Crystallization of proteins: overview of applications in chemical biology

Table1

Nuclear Hormone Receptors and Their Ligands

| Subfamily	Examples of Members	Endogenous Ligands	Examples of Synthetic Ligands	Reference
NR1	TR	Thyroid hormone	GC-1	[40-42]
	PPAR	Fatty acids	GW6471(PPARα) Rosiglitazone (PPARγ)	[24, 50-52]
	LXR	Oxysterols	GW3965 T0901317	[54, 56]
	PXR	Not known	Rifampicin SR12813 Hyperforin	[65-67]
	CAR	Not known	Androstanol Phenobarbital CITCO	[68-71, 75]
NR2	RXR	Retinoic Acid	GW0791 (RXRα)	[14, 70, 72]
	HNF4	Fatty acids (?)	None to date	[76, 77]
NR3	ER	Estradiol	Tamoxifen ICI164,384	[37-39]
	ERR	Not known	GSK4716 (ERRγ)	[83, 84]
NR4	NGFI-B	Not known	None to date	[80]
	Nurr1	Not known	None to date	[78, 80]
NR5	SF-1	Phospholipids (?)	GSK8470	[33, 34, 36, 64]
	LRH-1	Phospholipids (?)	GSK8470	[33, 35, 36, 64]
NR6	GCNF	Not known	None to date	[1]
NR0	DAX-1	Not known	None to date	[81]
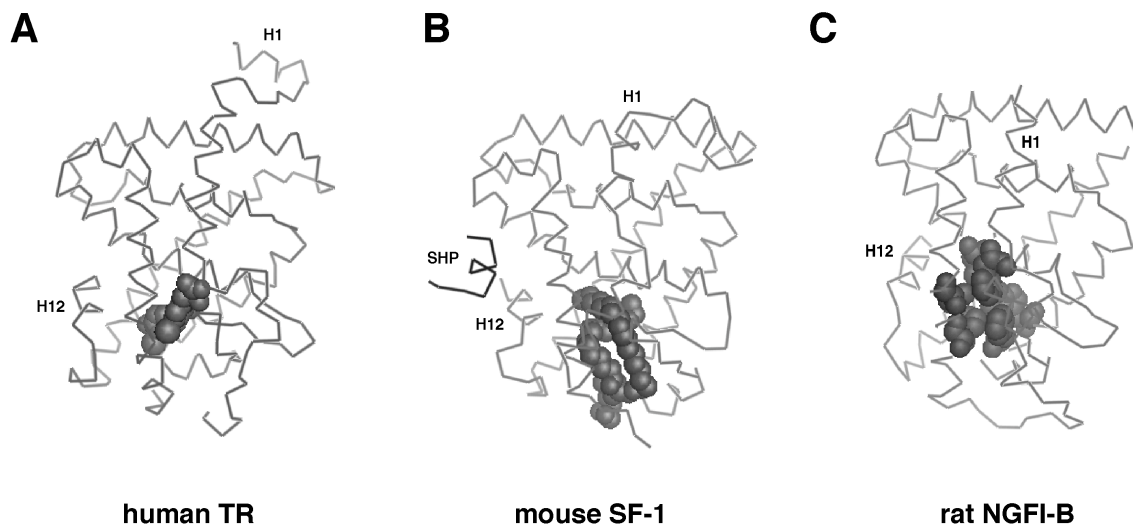
Figure 1. Nuclear receptors can accommodate ligands of various sizes
(A) LBD structure of human TR bound to its ligand, triiodothyronine (shown as spheres) (PDB 1XZX) [40]. **(B)** LBD structure of mouse SF-1 bound to a bacterial phospholipid (shown as spheres) and the mouse SHP peptide (PDB 1YMT) [33]. **(C)** LBD structure of rat NGFI-B with an empty ligand-binding pocket; hydrophobic amino acids occluding the ligand-binding pocket are highlighted as spheres (PDB 1YJE) [80]. Helix 1 (H1) and helix 12 (H12) for each structure are indicated.
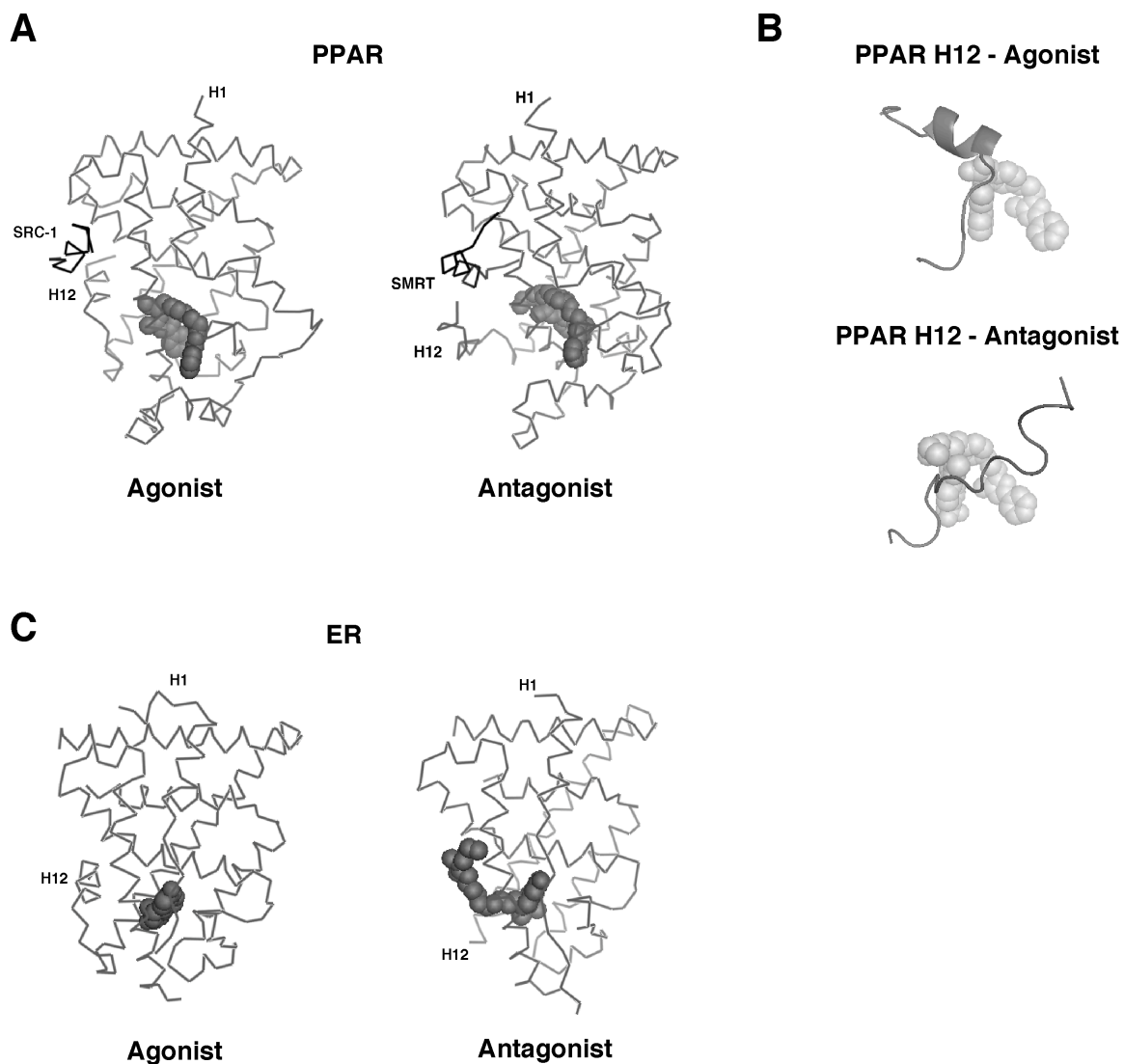
Figure 2. Helix 12 position differs in agonist- and antagonist-bound receptors
(A) LBD structures of human PPAR bound by an agonist ligand and a coactivator peptide (SRC-1) (PDB 1K7L) or an antagonist ligand and a corepressor peptide (SMRT) (PDB 1KKQ) [24, 50]. **(B)** Helix 12 structure from agonist- and antagonist-bound receptor shown in **A**. **(C)** LBD structures of human ER bound by an agonist, 17β-estradiol (PDB 1ERE), and rat ER bound by an antagonist, ICI 164,384 (PDB 1HJ1) [37, 39]. Helix 1 (H1) and helix 12 (H12) for each structure are indicated; ligands are shown as spheres.

CHAPTER III

The DEAD-Box protein DP103 (Ddx20, Gemin-3) represses orphan nuclear receptor activity via SUMO-modification

Martin B Lee*, Lioudmila A Lebedeva*, Miyuki Suzawa, Subhagya A. Wadekar, Marion Desclozeaux, and Holly A Ingraham

Department of Physiology[‡],
Biomedical Sciences Graduate Program
Graduate Program in Biological Sciences
1550 4th Street, GDBS Building
Mission Bay Campus
University of California, San Francisco
Box 0444, San Francisco, CA 94143-2611

Corresponding author: Holly A. Ingraham
E-mail   hollyi@itsa.ucsf.edu
Phone: (415) 476-2731
Fax: (415) 514-3792

Short title: SUMO Repression of NR5A Receptors by a DEAD-Box Protein

* These authors contributed equally to this work

Abstract

Structural analysis of NR5A orphan nuclear receptors suggests that ligand-independent mechanisms must regulate this subclass of receptors. Here, we report that steroidogenic factor 1 (SF-1) and liver receptor homolog (LRH-1) are repressed via posttranslational SUMO modification at conserved lysines within the hinge domain. Indeed, mutating these lysines or adding the SUMO isopeptidase SENP1 dramatically increased both native and Gal4-chimera receptors activities. The mechanism by which SUMO conjugation attenuates SF-1 activity was found to be largely HDAC-independent and was unaffected by the AF2-corepressor, Dax-1. Instead, our data suggest that SUMO-mediated repression involves direct interaction of the DEAD-box protein DP103 with sumoylated SF-1. Of potential E3-SUMO ligase candidates, PIASy and PIASxa strongly promoted SF-1 sumoylation, and addition of DP103 enhanced both PIAS-dependent receptor sumoylation and SF-1 relocalization to discrete nuclear bodies. Taken together, we propose that ATPases/RNA helicases are directly coupled to transcriptional repression by protein sumoylation.

Introduction

Steroidogenic factor 1 (SF-1) and liver receptor homolog 1 (LRH-1) are two closely related transcription factors belonging to the nuclear receptor subfamily V (NR5A) that contain a highly conserved DNA binding domain (DBD), a large hinge domain and a ligand binding domain (LBD, Fig 1A). *Drosophila* Ftz-F1 is the founding member of this subfamily and interacts directly with the pair-rule gene product of *Ftz* to control parasegmention at early embryonic stages (25). The mammalian orthologs SF-1 and

LRH-1 are also critical in tissue development and organogenesis (19, 27, 33). During development SF-1 is essential for male differentiation, adrenogonadal morphogenesis, and terminal differentiation of the ventromedial hypothalamus, and in the adult, this receptor regulates genes involved in steroid biosynthesis and endocrine signaling (34, 44). Although SF-1 null mice die at birth from adrenal failure, SF-1 heterozygous mice live. However, further analyses of these heterozygous mice show that despite seemingly adequate levels of SF-1, the amount of active SF-1 protein is insufficient to overcome defects in adrenal morphogenesis (2, 3). In humans, SF-1 haploinsufficiency is associated with severe adrenal disease, and gonadal dysgenesis (1, 28). LRH-1 acts far earlier in development than SF-1, as evidenced by the embryonic lethality observed in LRH-1 null embryos (33). In vitro and in vivo analyses have implicated LRH-1 in bile acid homeostasis (13, 26), where a heterozygous phenotype has also emerged in the intestine (4). In addition, LRH-1 controls tissue conversion of androgens to estrogen by regulating aromatase gene expression (7, 17)

Despite the fact that the high-resolution crystal structure of LRH-1 revealed a large hydrophobic pocket within the LBD (38), natural ligands have yet to emerge for this subclass of receptors. As such, the question of how subfamily V receptors are regulated is unclear. In many cellular contexts, this subclass of receptors is active and presumably recruits coactivators in a ligand-independent manner. NR5A receptor activity depends on two distinct regions in the LBD, an activation function in helix 1 (AFH1), and the C-terminal AF2 domain (8, 20). In both SF-1 and LRH-1 a "repression domain" has been identified in the hinge region (32, 47). For SF-1, this domain is reported to interact with

the DEAD-box RNA helicase DP103, (Ddx20, Gemin-3) (49), although the precise mechanism of SF-1 repression by DP103 is unknown.

Phosphorylation and sumoylation are posttranslational modifications known to modulate nuclear receptors. Phosphorylation of SF-1 is proposed to increase receptor activity by stabilization of the LBD and enhanced cofactor recruitment (8, 11, 15). On the other hand, sumoylation of transcription factors, such as Elk-1, Lef1, and nearly all steroid nuclear receptors, results in their transcriptional repression (5, 18, 35, 39, 42, 50). Sumoylation occurs at canonical motifs of ψKXE, where ψ is a hydrophobic amino acid and K is the acceptor lysine for covalent attachment of the Small Ubiquitin-like Modifier (SUMO). SF-1, LRH-1, and other invertebrate NR5 receptors are predicted to be sumoylated given the presence of a conserved I**K**SE or I/V**K**QE site in the hinge region (Fig 1A). SUMO modification of proteins is analogous to ubiquitination, involving a three-step ATP-dependent reaction. Processed SUMO protein is loaded onto the heterodimeric E1 enzyme (SAE1/SAE2) and transferred from E1 to the sole E2 enzyme Ubc9, which then mediates SUMO conjugation to the protein substrate with aid from E3-SUMO ligases. Protein Inhibitor of Activated Stats (PIAS) proteins comprise the largest of three identified E3-SUMO ligase classes (29). This protein conjugation is dynamic and easily reversed by Sentrin/SUMO-specific proteases (SENP/SUSP), which cleave SUMO from its substrate. However unlike ubiquitin conjugation, which primarily facilitates protein degradation, SUMO modification of transcription factors often results in transcriptional repression. Others have proposed that this repression involves direct

41

recruitment of histone deacetylases (HDACs) (40, 51), or a relocalization of the SUMO-marked protein to PML nuclear bodies (9, 39).

Here we identify sumoylation as an important posttranslational regulatory mechanism for dampening the activity of subfamily V nuclear receptors. Potential mechanisms for sumoylation-mediated repression were investigated and found to involve a functional interaction between the receptor and the DEAD-box RNA helicase DP103.

Materials & Methods

Plasmids

Full-length mSUMO1 (101aa) was PCR-amplified from embryonic mouse hypothalamic-enriched cDNA using primers: 5'-CTCGAGATGTCTGACCAGGAGGCAAAA-3', 5'-TCTAGACTAAACCGTCGAGTGACCCCC-3', TA-cloned into pCRII (Invitrogen), and subcloned into Xho1-Xba1 pCI-neo. Processed His$_6$-hSUMO1 (97 aa) was subcloned from His$_6$-hSUMO1-pcDNA3 (F. Poulat) into pGEX4T1 at BamH1. HA-tagged mSENP1 was PCR-cloned from mouse hypothalamic-enriched cDNA using primers: 5'-CCGGAATTCATGTACCCATACGACGTACCAGATTACGCTAGCTTGGATGACACAGCTGATGGGGTG-3', 5'-ACCTCTAGAGTCGACTCACAAGAGCTTCCGGTGGAG-3' using EcoR1-Sal1 of pCI-neo. HA-tagged mSF1 in pCI-neo, HA SF-1 S203A and GFP-HA-SF-1 pCMV were described previously (8). K119R, K194R, and 2KR mutants of HA-SF-1-pCI-neo and GFP-HA-SF-1 were created by PCR mutagenesis (Stratagene). All Gal4 constructs contained a HA-epitope tag N-terminal to the Gal4-DBD. A C-terminal fragment containing the hinge-LBD (aa105-462) of SF-1 and mutants was

generated by PCR from HA-SF-1-pCI-neo with primers 5'–ACGCGTCCTTGAAGCAGCAGAAGAAAGCA-3' and 5'–AAGCTTTCAAGTCT GCTTGGCCTG–3', and subcloned 3' to the HA-Gal4-DBD.   A similar strategy was used to create all pGAL-LRH-1 constructs with the LRH-1 (aa198-562) fused to Gal4. FLAG-mPIASxa was cloned from RIKEN clone 4921511I02 with primers 5'-CCGGAATTCATGGACTACAAAGACGACGACGACAAAGCGGATTTCGAGGAG TTG-3', 5'-CCGCTCGAGTCACTGTTGCACAGTATCAGA-3' and FLAG-mPIAS1 was cloned from mouse hypothalamic cDNA using primers: 5'-CTCGAGATGGACTACAAAGACGACGACGACAAAGCGGACAGTGCGGAACTA AAG-3', 5'-CCGCTCGAGTCAG-TCCAATGAGATAATGTC-3'. PCR products were subcloned into pCI-neo, pBH4 and pGADT7. pVP16-PIAS1 and pVP16-PIASxα were generated by inserting FLAG-mPIAS1 and FLAG-mPIASxα PCR fragments downstream of the VP16 activation domain in a pVP16 vector (Clontech). The following constructs were generous gifts: T7 tagged-mPIASy pCMV (R. Grosschedl); FLAG-mPIAS3 pCMV (K. Shuai); full length mDP103 pcDNA3 (Y. Sadovsky); C-terminal hDP103 pGEX (aa414-824) and full length 2FLAG-hDP103 pcDNA3 (from C. Glass).

Cell transfections, luciferase assays, and metabolic labeling

COS-7 cells were plated at a density of 50,000 cells/mL/12-well plates or 1.5×10^6 cells/10cm plate in media (DME H21 4.5g/L glucose with 10% calf serum and antibiotics) 18 hrs prior to transfection. Transfections were carried out using FuGene 6 (Roche). For luciferase assays, cells were transfected with no more than 500ng total DNA per well, and harvested 48 hrs after transfection (BD Pharmingen). All

transfections were performed in triplicate and repeated at least twice. Results were normalized to β-galactosidase activity and are expressed as relative luciferase units (RLU) or fold activation as indicated. For metabolic labeling, COS-7 cells were plated in full media and transfected 18 hrs after plating. Cys/Met deficient media (DME H21 4.5 g/L glucose, 10% dialyzed fetal bovine serum, 2mM glutamine and antibiotics) was added to washed cells 48 hrs post-transfection, followed by 1 hr pulse-labeling with 350 μCi of ^{35}S-Cys and ^{35}S-Met (Redivue, AGQ0080, Amersham), washing, and incubation in full media for relevant chase periods. Cells lysates were subjected to immunoprecipitation, SDS-PAGE electrophoresis, and autoradiography, and signal was quantified by phosphor-imaging.

Yeast Interaction System

An expression cassette containing full-length mouse SF-1 (no heterologous activation domain) was integrated in yeast strain YM4271 containing two integrated reporters, *HIS* and LacZ, driven by four tandem copies of the SF-1 response elements, using the manufacturer's protocols (Clontech). Full length FLAG-tagged mPIAS1, mPIASxα, mPIASy were subcloned into pGADT7 for transformation into yeast reporter strains. Transformants were plated on selective media, and analyzed on X-gal medium and by liquid ß-galactosidase assays.

Western analysis, immunoprecipitation and co-immunoprecipitation

Cells were washed twice in cold PBS (calcium and magnesium free), lysed in 50mM Tris-HCl pH 7.6, 150mM NaCl, 1mM EDTA, 0.1% NP40, 0.5mM PMSF, 0.5mM DTT,

protease inhibitors (Roche), and pre-cleared by centrifugation at 14000 rpm for 30 min. When appropriate, all solutions contained 20 mM N-ethylmaleimide (NEM, Sigma) to inhibit SUMO isopeptidases. Protein concentrations were determined by the Bradford method (Pierce). Equal amounts of total protein were loaded for Western blot analysis. Wild type and mutant receptors were affinity purified using anti-HA affinity matrix (Covance/Babco) in lysis buffer (as described above), washed in a modified lysis buffer containing 300mM KCl and 0.05% NP40, subjected to 8.5% SDS-PAGE and Western blotting following incubation with primary antibodies (anti-HA, 1:2000, Covance/Babco; anti-FLAGM2 1:2000 Sigma; anti-SUMO1, 1:500 Zymed) and a HRP goat anti-mouse 1:10,000 secondary antibody (BioRad). Signal was developed by chemiluminescence (ECL, Amersham). For coimmunoprecipitation of FLAG-hDP103 and sumoylated SF1, cells were transfected and lysed as for in vivo sumoylation in 10 mM NEM. Lysates were incubated with anti-FLAG M2 agarose beads (Sigma) in pull down buffer (50 mM Tris HCl pH 7.6, 150 mM NaCl, 0.5 mM EDTA, 0.01%NP40, 2 mM NEM, protease inhibitors) and precipitates analyzed by Western blotting (anti-HA 1:2000, Covance/Babco; anti-hDP103 1:2000, BD Biosciences).

Recombinant protein expression, in-vitro sumoylation assay and GST pulldowns
Recombinant His_6-hSUMO1 (aa1-97) was expressed and purified by TALON chromatography (Clontech). Recombinant His_6-hE1 (SAE1/SAE2) and His_6-hUbc9 were obtained commercially (LAE Biotech). In vitro transcribed-translated ^{35}S-SF-1 and variants thereof were produced (Promega) and incubated with 150 ng of E1, 750 ng His_6-Ubc9, 900 ng His_6-SUMO1 in 50 mM Tris pH 7.6, 5 mM $MgCl_2$, 1 mM DTT, 2.5 mM

ATP at 37°C for 1.5 hrs and the reaction stopped by boiling in protein loading buffer. Samples were subjected to 8% SDS-PAGE followed by autoradiography. GST pulldown assays were carried out with ^{35}S-SF-1 or variants thereof, and purified GST-C-terminal hDP103 as described (15, 21).

Chromatin Immunoprecipitation Assay

HeLa Luciferase Reporter HLR (Stratagene) cells containing an integrated promoter-reporter of five Gal4 binding sites fused to the luciferase gene were electroporated with pCI-Neo and HA-tagged pGal-SF-1 constructs (4μg). The method used follows that described in (46) with PCR conditions of 25 cycles at 95˚C for 30 s, 53˚C for 1 min, 72˚C for 1 min, using primers as described (40) to amplify a 5' 330 bp region of luciferase cDNA.

NUCLEAR LOCALIZATION AND IMMUNOHISTOCHEMISTRY

COS-7 cells were plated at 6000 cells/well in 4-well chamber slides (Lab-Tek) and transfected in duplicate 24 hrs later (total DNA: 0.5 mg/well). 48 hrs post-transfection, cells were fixed in 4% paraformaldehyde, permeabilized in PBS containing 0.3% Triton-X100 followed by incubation with primary antibodies (rabbit anti-T7 1:300, ICL; mouse anti-FLAGM2 1:5000, Sigma; mouse anti-SF2/ASF 1:1000, Zymed; goat anti-Sp100 1:50 Zymed; and mouse anti-PML (PG-M3), 1:75, Santa Cruz), followed by secondary antibodies (Cy-3 goat anti-rabbit 1:1000; Cy3-donkey anti-mouse 1:1000, Molecular Probes; Texas Red rabbit anti-goat 1:500, Vector) and imaged on a Zeiss LSM510 confocal microscope.

Results

Subfamily V receptors are sumoylated in the hinge region

Although sumoylation is known to repress steroid receptor activity, this modification has not been investigated for so-called orphan nuclear receptors, which can function in a ligand-independent manner. In a modified one-hybrid yeast screen for SF-1 protein partners, we identified Ubc9 or the E2 SUMO conjugating enzyme, as a strong interacting protein (data not shown, see supplementary data, P1). We next asked if SF-1 and LRH-1 could be sumoylated. Indeed, sequence analysis of all vertebrate species of SF-1 and LRH-1 revealed two highly conserved canonical sumoylation motifs at the N- and C-terminal hinge region, while insect Ftz-F1 variants contained one site in the N-terminal hinge region (Fig 1A).

Sumoylation of both SF-1 and LRH-1 was demonstrated in a cellular system as evidenced by slower migrating bands after coexpression of receptor with either SUMO1 or GFP-SUMO1 (Fig 1B). In addition, a similar slower-migrating SF-1 species was detected in NEM-treated lysates made from both Y1 and aT3 cells (Fig 1C, and data not shown), suggesting that endogenous SF-1 is sumoylated. Further analysis revealed that Lys194 served as the major acceptor lysine for SF-1 sumoylation as evidenced by the loss of the slower migrating band with the single mutation K194R and double mutation (K119R and K194R, referred to as 2KR), but not with K119R (Fig 1D). Our results for SF-1 are similar to other recent reports (6, 22). The identity of these slower migrating SF-1 species as sumoylated receptor was confirmed by immunoprecipitation of HA-epitope tagged SF-1, followed by Western blotting with an anti- SUMO1 antibody (Fig

1E), and as predicted no sumoylated species were observed with K194R or 2KR mutant proteins. Their results were confirmed in an in vitro sumoylation assay with Lys194 identified as a major site and Lys119 presumed to be a minor sumoylation site (Fig 1F). Amounts of sumoylated SF-1 diminish in both the K194R and 2KR mutants; the faint residual upshifted band observed in the 2KR variant imply that a minor third site can be sumoylated in vitro. Taken together, we conclude that subfamily V receptors are sumoylated in vivo and in vitro.

Sumoylation of SF-1 attenuates transcriptional activity

Previous studies identified a regulatory domain which when mutated led to increased receptor activity; this domain contained the major sumoylation site for SF-1 and LRH-1 (Fig 1A and (32, 47). Consistent with these reports, we found increased activity of NR5A promoter reporters with either SF-1 or LRH-1 sumoylation mutants (Fig 2A). Increased receptor activity observed with both the K194R and 2KR receptor mutants was not due to increased protein stability, as judged by results from pulse-chase metabolic labeling experiments (Fig 2B). Gal4-SF-1/LRH-1 fusion receptors containing the full hinge and LBD also showed a dramatic increase in activity following mutation of the sumoylation acceptor sites. Strikingly, the single mutant K194R was at least 70-fold more active than wild type and mutation of both sumoylation sites (2KR) resulted in greater than 300-fold activation (Fig 2C, left panel). While K119R exhibited comparable activation to that of wild type, the double mutant at both Lys119 and Lys194 showed remarkable synergism; this is consistent with Lys119 as a minor site. Similar to native receptors, Gal4-SF-1 and Gal4-K119R are efficiently sumoylated, whereas Gal4-K194R

and Gal4-2KR exhibit no detectable sumoylation (Fig 2C, left lower panel). Nearly identical results were observed for Gal4-LRH-1 constructs, where double mutation of K213R/K289R in the hinge region led to strong receptor activation (Fig 2C, right panel).

To confirm that receptor sumoylation served to repress SF-1 activity, we asked if removing the SUMO conjugate from SF-1 with the SUMO isopeptidase, SENP1 would yield similar results as observed with the SF-1 lysine mutants. Indeed, coexpression of SENP1 with SF-1 and SUMO1 resulted in a marked attenuation of sumoylated SF-1 (Fig 3A). Furthermore, activities of both wild type and the K119R mutant were enhanced after addition of small amounts of SENP1 expression vector (25ng); reaching levels observed with the K194R mutant (Fig 3B, left panel). Addition of SENP1 failed to activate the 2KR variant providing further evidence that Lys119 and Lys194 are the major sites of sumoylation (Fig 3B, right panel). Collectively our data suggest that Lys194 plays a dominant role in mediating repression of SF-1 via sumoylation and that receptor sumoylation represents a major silencing mechanism.

A DEAD-box protein mediates repression via SF-1 sumoylation.

The mechanisms by which protein sumoylation leads to transcriptional repression are diverse. Recent literature suggests that repression by sumoylation involves 1) nuclear relocalization with a concomitant decrease of promoter occupancy, or 2) direct recruitment of histone deacetylases (HDACs). Therefore, we asked whether sumoylation mutants differ in their subnuclear localization. Both GFP-wild type and GFP-SUMO mutants yielded nearly identical patterns of nuclear localization (Fig 4A). Consistent

49

with these results, no apparent differences were noted in the promoter occupancy of Gal4-wild type compared to the K194R mutant as judged by chromatin immunoprecipitation (ChIP) results using a HeLa cell line containing a stably integrated Gal4 reporter (Fig 4B). We next asked if SF-1 sumoylation promotes recruitment of HDACs by using the Class I and II HDAC inhibitors, trichostatin A (TSA) and sodium butyrate (NaBT). If HDAC recruitment is essential for SUMO-mediated repression, mutating the major sumoylation sites within SF-1 should prevent derepression by TSA or NaBT. Instead, addition of TSA or NaBT led to a dramatic increase in the activity of all receptor variants (Fig 4C, D). Our results differ from those recently shown for Elk-1 where loss of sumoylation eliminates TSA sensitivity (51), and thus, we suggest that repression of SF-1 via sumoylation is largely HDAC-independent.

For subfamily V, two types of repressors have been identified. The first includes the orphan nuclear receptors Dax-1 or SHP, which interfere with the AF2 in the LBD. The second is the RNA helicase DEAD-box protein DP103 (32). Indeed, while Dax-1 was able to repress the Gal4-K194R mutant as effectively as Gal4-WT (Fig 5A, left panel), DP103 was ineffective at repressing the Gal4-K194R and 2KR mutants (Fig 5A, right panel and data not shown). Moreover, addition of SENP1 failed to abolish Dax-1-mediated repression of SF-1 (Fig 5B, left panel). In contrast, addition of SENP1 completely abrogated DP103-mediated repression of Gal4-SF-1 (Fig 5B, right panel). Our work contrasts a recent report showing no difference between DP103-mediated repression in wild-type and K194R (22). This discrepancy may reflect a difference in cell-types or the significantly greater amounts of DP103 used compared to experiments

shown here. Nonetheless, our data agree with those reported by Sadovsky and colleagues showing Lys194 to be essential for DP103 repression of SF-1 (32).

To test the hypothesis that sumoylation at Lys194 allows DP103 to function as a repressor, interaction between DP103 and sumoylated SF-1 was explored by direct binding assays. As shown previously, only the C-terminal half of DP103 interacts with SF-1 (32). Mutation of the Lys194 and/or Lys119 did not result in an appreciable loss of binding, suggesting that Lys194 is not the sole determinant for DP103 interaction with SF-1 (Fig 5C). Furthermore, DP103 is able to interact efficiently with in vitro sumoylated forms of SF-1 (Fig 5D). These results provide evidence that the DEAD-box protein, DP103 interacts with sumoylated SF-1 and directly participates in receptor repression.

DP103 promotes PIAS-dependent sumoylation and subnuclear relocalization of SF1

To further explore how DP103 might affect SF-1 activity, we first defined the optimal E3-SUMO ligase in vivo. One of the defining characteristics of an E3-SUMO ligase is its ability to interact with, and promote sumoylation of a given substrate. In both the yeast and mammalian two hybrid assays, SF-1 interacted strongly with PIAS1, and less well with PIASxa and PIASy (Fig 6A, B). However, despite this strong interaction, PIAS1 does not serve as an efficient E3-SUMO ligase for SF-1, in vivo. In a survey of four PIAS members, only PIASxa and PIASy promoted SF-1 sumoylation in a dose-dependent manner; this effect was not observed for PIAS1 or PIAS3 (Fig 6C, left panel and supplementary data, P2). In contrast to results from the in vitro assay,

51

overexpression of PIAS proteins in vivo does not reveal detectable sumoylation at non-canonical sites, as evidenced by the 2KR mutant (Fig 6C, right panel). Interestingly, mutating the major phosphorylation site of SF-1 adjacent to Lys194 (S203A) had no effect on receptor sumoylation (Fig 6C, right panel). Next, the functional effects of overexpressing PIAS proteins on wild type and 2KR receptors were determined. Consistent with PIAS-dependent activation of other nuclear receptors (24), we observed an initial activation phase, followed by repression when PIASxa is added to wild type receptor (Fig 6D). Addition of SUMO1 further enhanced receptor repression, suggesting that increased sumoylation does silence SF-1 activity. In contrast, increased repression was not observed with the double 2KR mutant (Fig 6E). The global repression observed with increasing amounts of SUMO1 to either wild type or mutant receptors most likely reflects the multiple nuclear substrates affected by the sumoylation machinery, including corepressors and coactivators (23).

To determine how sumoylation affects interaction between DP103 and SF-1, the levels of receptor sumoylation were driven by the optimal E3-SUMO ligase, PIASy. DP103 interacted with SF-1 in the presence of PIASy, but not under basal levels of sumoylation or after addition of SENP1 (Fig 7A). Surprisingly, DP103 enhanced PIAS-mediated sumoylation (2-3 fold) for all PIAS proteins, except PIAS3 (Fig 7B and supplementary data, P3). No significant increase in sumoylation was observed with DP103 alone (Cont.). Whether this effect arises from increased ligase activity of PIAS proteins or by protecting sumoylated SF-1 from desumoylation, remains to be determined. Finally, we asked if DP103 would alter the subnuclear localization of SF-1. Although our previous

52

results suggested that the nuclear pattern of SF-1 does not change under basal levels of sumoylation, a dramatic relocalization of GFP-SF-1 was revealed when DP103 was coexpressed with PIASy and SUMO1; two representative cells with prominent nuclear bodies are shown (Fig 7C). Addition of SUMO1, PIASy, or DP103, alone, or a combination of DP103 plus PIAS1, PIASxa, or PIAS3, showed no SF-1 relocalization (Fig 7C, data not shown). However, we noted the presence of fine GFP-SF-1 foci in some cells with PIASy alone (Fig 7C). The ability of DP103 and PIASy to shuttle SF-1 to discrete nuclear bodies does not apparently require SF-1 sumoylation, as evidenced by a speckled pattern after addition of SENP1 or with the K119R, K194R and 2KR GFP-SF-1 mutants (Fig 7C and data not shown). Further analysis revealed colocalization of GFP-SF-1 with PIASy, but not with DP103, which localizes to Cajal bodies or gems (Fig 7D). These GFP-SF-1 nuclear bodies appear distinct from endogenous splicing speckles as shown by the non-overlapping patterns between GFP-SF-1 and SF2/ASF. Moreover these foci do not resemble PML nuclear bodies (PML-NBs), given that we failed to detect obvious PML-NBs in COS-7 cells under our culture conditions using two markers, Sp100 and PML (Fig 7D and data not shown). Collectively, our data suggest that DP103 promotes PIAS-mediated sumoylation, and together with PIASy, relocalizes SF-1 to discrete nuclear foci. Whether these foci are functionally significant remains to be determined, however, their formation correlates well with optimal receptor sumoylation suggesting a functional complex between SF-1, PIASy, and DP103.

Discussion

In this study we report that subfamily V nuclear receptors are sumoylated at evolutionarily conserved sites. As established for other transcription factors, SUMO-modification of SF-1 and LRH-1 attenuates significantly transcriptional activity. Mutating the acceptor lysines in both SF-1 and LRH-1 resulted in a more active receptor, and at least in the Gal4 context, the fold-increase is reminiscent of ligand-dependent receptor activation. Thus, for subfamily V receptors the extent of sumoylation represents one mechanism to both regulate and restrain receptor activity. Our data also suggest that sumoylation of the so-called "repression domain" in SF-1/LRH-1 marks the receptor for repression by the DEAD-box protein DP103. Moreover, this ATPase/RNA helicase was found to enhance PIAS-dependent receptor sumoylation and to promote PIASy-dependent shuttling of SF-1 to discrete nuclear bodies or foci. Subnuclear relocalization of SF-1 correlated strongly with conditions that promote extensive receptor sumoylation, suggesting that physical interaction between SF-1, DP103 and PIASy are linked to transcriptional repression.

Repression of SF-1 via sumoylation

In contrast to the ubiquitously expressed E1 and E2 sumoylation enzymes, most of the known E3-SUMO ligases exhibit restricted expression patterns, and therefore might direct tissue specific sumoylation of protein substrates (48). In considering SF-1 sumoylation, three E3-SUMO ligases (PIASxa, PIASy and PIAS1) are all highly expressed in the adult testes (14, 48), where SF-1 regulates multiple genes. SF-1 is also needed for male sexual differentiation (37, 45), and it is possible that sumoylation of SF-

54

1 is sexually dimorphic during development. Thus, silencing of male-specific genes in the ovary can be partially explained by lowered levels of SF-1 or by the actions of Dax1 (30, 41), but may also involve sumoylation. Interestingly, other factors that function in sexual differentiation, namely Sox9 and WT-1, contain sumoylation sites and the combinatorial effects of sumoylation may ensure gene silencing in the female. Finally, it is worth considering the in vivo ratio of non-sumoylated to sumoylated receptor. In this regard, SF-1 haploinsufficiency (2, 28) may stem from inadequate SF-1 activity due to a reduction of protein levels coupled with extensive receptor sumoylation.

Currently, our studies are limited to a loss-of-function analysis. Attempts to provide SUMO1 in *cis* to SF-1, as shown for other proteins (18, 50), have failed due to the precise excision of SUMO1 in COS-7 cells (L.A.L. and H.A.I., unpublished data). Whether SF-1 or LRH-1 sumoylation confers any structural changes to the DBD, hinge, or LBD remains unclear, however results from our ChIP analysis suggest that sumoylation does not alter the apparent DNA binding of a heterologous DBD. Moreover, given that Dax-1-mediated repression of K194R SF-1 mutant is intact, we suggest that no gross conformational changes occur in the LBD of a sumoylation defective receptor. Further structural analyses are needed and will require an appropriate SUMO-SF-1 chimera or SUMO stably conjugated to SF-1/LRH-1. Although our findings point to a functional role for Lys194 and Lys289 in SF-1 and LRH-1, respectively, the role of the minor sumoylation sites at Lys119 or Lys213 (Fig 1A) is less apparent. Despite the fact that disumoylated SF-1 is only observed in vivo under conditions that promote efficient sumoylation, our functional analyses show that both the minor and major sumoylation

sites act in concert to dampen receptor activity. In this regard, it remains to be established if an ordered sumoylation of SF-1/LRH-1 occurs.

Recent studies report interdependency between sumoylation and phosphorylation. MAP-kinase-mediated phosphorylation of Elk-1 greatly reduced sumoylation at adjacent lysines and led to increased transcriptional activity (50) and phosphorylation of heat shock factor 1 is a prerequisite for stress-induced sumoylation (16). Currently, we find no apparent relationship between phosphorylation of Ser203 and sumoylation of SF-1. Indeed, the phospho-deficient S203A mutant was efficiently sumoylated, and all SF-1 SUMO mutants showed equivalent levels of phosphorylated Ser203 in SF-1 (unpublished data). However, it remains possible that the rate and extent of either phosphorylation or sumoylation are altered following modification of the Ser203 or Lys194, respectively.

DEAD-Box Proteins and Transcriptional Repression

Historically, DEAD-box (Ddx) RNA helicases are associated with splicing, in part, because they were initially identified as protein components of the spliceosome (43). However, other functions for Ddx family members have been noted and there is mounting evidence that they function to silence transcription factors, including nuclear receptors, Egr1-4, and the Ets-like repressor, METS (12, 21, 36, 49). Additionally, GRTH (Ddx25), which is expressed in testes, is reported to attenuate expression of SF-1 target genes, including steroidogenic enzymes (10). For DP103 and another DEAD-box protein DP97, the repression domain has been mapped to the C-terminal region and does not require the N-terminal ATPase/helicase domain characteristic of this gene family (21,

36). Attenuation and silencing of transcription are multi-layered and multi-dimensional.
So how might Ddx proteins and sumoylation lead to transcriptional repression?
Recruitment of HDACs upon protein sumoylation, or by Ddx proteins, offers the most
plausible explanation and is consistent with prior literature. Indeed, DP103 interacts with
the N-terminal repression domain of METS and promotes HDAC recruitment (21).
However, our data imply that repression through DP103 is TSA- and NaBT-insensitive;
and suggest that repression by Ddx proteins must involve additional mechanisms other
than recruitment of Class I or II HDACs. In considering other mechanisms, it is possible
that DP103 protects SF-1 from desumoylation. This hypothesis is consistent with the
observations that DP103 increased PIAS-dependent SF-1 sumoylation, and that
additional SENP1 eliminates repression by DP103. The interaction between DP103 and
SF-1 remains to be mapped and is likely to involve multiple interfaces based on our
finding that Lys194 and/or sumoylation at Lys119/Lys194 are not sole determinants of
this interaction. Another possible scenario is that DP103 represses SF-1 by facilitating
PIASy-mediated relocalization of SF-1. However, we noted that sumoylation is
dispensable for movement of SF-1 to nuclear bodies; this observation is reminiscent of
PIASy-dependent relocalization of both wild type and sumoylation defective Lef1 into
nuclear bodies that partially overlap with PML-NBs (39). Thus, while sumoylation is not
required for subnuclear relocalization of SF-1 (or Lef1), conditions that promote optimal
sumoylation do correlate with altered nuclear distribution of SF-1.

Given that DEAD-box proteins are present in both splicing and translational complexes
(31), repression might be coupled to transcript processing or translational control.

However, studies to date, including ours, have yet to identify a function for the RNA helicase (unwindase) and RNA binding motifs in repression. Indeed, the N-terminal portion of DP103 is dispensable for interaction and repression of SF-1 and METS (21, 49), and for relocalization of SF-1 to nuclear bodies (our unpublished data). Further in vitro and in vivo experiments aimed at delineating the precise role of sumoylation in DEAD-box-mediated transcriptional repression will be of interest.

References

1. **Achermann, J. C., J. J. Meeks, and J. L. Jameson.** 2001. Phenotypic spectrum of mutations in DAX-1 and SF-1. Mol Cell Endocrinol **185:**17-25.
2. **Bland, M. L., R. C. Fowkes, and H. A. Ingraham.** 2004. Differential requirement for steroidogenic factor-1 gene dosage in adrenal development versus endocrine function. Mol Endocrinol **18:**941-52.
3. **Bland, M. L., C. A. Jamieson, S. F. Akana, S. R. Bornstein, G. Eisenhofer, M. F. Dallman, and H. A. Ingraham.** 2000. Haploinsufficiency of steroidogenic factor-1 in mice disrupts adrenal development leading to an impaired stress response. Proc Natl Acad Sci U S A **97:**14488-93.
4. **Botrugno, O. A., E. Fayard, J. S. Annicotte, C. Haby, T. Brennan, O. Wendling, T. Tanaka, T. Kodama, W. Thomas, J. Auwerx, and K. Schoonjans.** 2004. Synergy between LRH-1 and beta-catenin induces G1 cyclin-mediated cell proliferation. Mol Cell **15:**499-509.
5. **Chauchereau, A., L. Amazit, M. Quesne, A. Guiochon-Mantel, and E. Milgrom.** 2003. Sumoylation of the progesterone receptor and of the steroid receptor coactivator SRC-1. J Biol Chem **278:**12335-43.
6. **Chen, W. Y., W. C. Lee, N. C. Hsu, F. Huang, and B. C. Chung.** 2004. SUMO Modification of Repression Domains Modulates Function of Nuclear Receptor 5A1 (Steroidogenic Factor-1). J Biol Chem **279:**38730-5.
7. **Clyne, C. D., A. Kovacic, C. J. Speed, J. Zhou, V. Pezzi, and E. R. Simpson.** 2004. Regulation of aromatase expression by the nuclear receptor LRH-1 in adipose tissue. Mol Cell Endocrinol **215:**39-44.
8. **Desclozeaux, M., I. N. Krylova, F. Horn, R. J. Fletterick, and H. A. Ingraham.** 2002. Phosphorylation and intramolecular stabilization of the ligand binding domain in the nuclear receptor steroidogenic factor 1. Mol Cell Biol **22:**7193-203.
9. **Dobreva, G., J. Dambacher, and R. Grosschedl.** 2003. SUMO modification of a novel MAR-binding protein, SATB2, modulates immunoglobulin mu gene expression. Genes Dev **17:**3048-61.
10. **Dufau, M. L., C. Tsai-Morris, P. Tang, and A. Khanum.** 2001. Regulation of steroidogenic enzymes and a novel testicular RNA helicase. J Steroid Biochem Mol Biol **76:**187-97.
11. **Fowkes, R. C., M. Desclozeaux, M. V. Patel, S. J. Aylwin, P. King, H. A. Ingraham, and J. M. Burrin.** 2003. Steroidogenic factor-1 and the gonadotrope-specific element enhance basal and pituitary adenylate cyclase-activating polypeptide-stimulated transcription of the human glycoprotein hormone alpha-subunit gene in gonadotropes. Mol Endocrinol **17:**2177-88.
12. **Gillian, A. L., and J. Svaren.** 2003. The Ddx20/DP103 dead Box protein represses transcriptional activation by Egr2/Krox-20. J Biol Chem.
13. **Goodwin, B., S. A. Jones, R. R. Price, M. A. Watson, D. D. McKee, L. B. Moore, C. Galardi, J. G. Wilson, M. C. Lewis, M. E. Roth, P. R. Maloney, T. M. Willson, and S. A. Kliewer.** 2000. A regulatory cascade of the nuclear

receptors FXR, SHP-1, and LRH-1 represses bile acid biosynthesis. Mol Cell **6:**517-26.

14. **Gross, M., B. Liu, J. Tan, F. S. French, M. Carey, and K. Shuai.** 2001. Distinct effects of PIAS proteins on androgen-mediated gene activation in prostate cancer cells. Oncogene **20:**3880-7.

15. **Hammer, G. D., I. Krylova, Y. Zhang, B. D. Darimont, K. Simpson, N. L. Weigel, and H. A. Ingraham.** 1999. Phosphorylation of the nuclear receptor SF-1 modulates cofactor recruitment: integration of hormone signaling in reproduction and stress. Mol Cell **3:**521-6.

16. **Hietakangas, V., J. K. Ahlskog, A. M. Jakobsson, M. Hellesuo, N. M. Sahlberg, C. I. Holmberg, A. Mikhailov, J. J. Palvimo, L. Pirkkala, and L. Sistonen.** 2003. Phosphorylation of serine 303 is a prerequisite for the stress-inducible SUMO modification of heat shock factor 1. Mol Cell Biol **23:**2953-68.

17. **Hinshelwood, M. M., J. J. Repa, J. M. Shelton, J. A. Richardson, D. J. Mangelsdorf, and C. R. Mendelson.** 2003. Expression of LRH-1 and SF-1 in the mouse ovary: localization in different cell types correlates with differing function. Mol Cell Endocrinol **207:**39-45.

18. **Holmstrom, S., M. E. Van Antwerp, and J. A. Iniguez-Lluhi.** 2003. Direct and distinguishable inhibitory roles for SUMO isoforms in the control of transcriptional synergy. Proc Natl Acad Sci U S A **100:**15758-63.

19. **Ingraham, H. A., D. S. Lala, Y. Ikeda, X. Luo, W. H. Shen, M. W. Nachtigal, R. Abbud, J. H. Nilson, and K. L. Parker.** 1994. The nuclear receptor steroidogenic factor 1 acts at multiple levels of the reproductive axis. Genes Dev **8:**2302-12.

20. **Ito, M., Y. Park, J. Weck, K. E. Mayo, and J. L. Jameson.** 2000. Synergistic activation of the inhibin alpha-promoter by steroidogenic factor-1 and cyclic adenosine 3',5'-monophosphate. Mol Endocrinol **14:**66-81.

21. **Klappacher, G. W., V. V. Lunyak, D. B. Sykes, D. Sawka-Verhelle, J. Sage, G. Brard, S. D. Ngo, D. Gangadharan, T. Jacks, M. P. Kamps, D. W. Rose, M. G. Rosenfeld, and C. K. Glass.** 2002. An induced Ets repressor complex regulates growth arrest during terminal macrophage differentiation. Cell **109:**169-80.

22. **Komatsu, T., H. Mizusaki, T. Mukai, H. Ogawa, D. Baba, M. Shirakawa, S. Hatakeyama, K. I. Nakayama, H. Yamamoto, A. Kikuchi, and K. I. Morohashi.** 2004. SUMO-1 Modification of the Synergy Control Motif of Ad4BP/SF-1 Regulates Synergistic Transcription between Ad4BP/SF-1 and Sox9. Mol Endocrinol.

23. **Kotaja, N., U. Karvonen, O. A. Janne, and J. J. Palvimo.** 2002. The nuclear receptor interaction domain of GRIP1 is modulated by covalent attachment of SUMO-1. J Biol Chem **277:**30283-8.

24. **Kotaja, N., M. Vihinen, J. J. Palvimo, and O. A. Janne.** 2002. Androgen receptor-interacting protein 3 and other PIAS proteins cooperate with glucocorticoid receptor-interacting protein 1 in steroid receptor-dependent signaling. J Biol Chem **277:**17781-8.

25. **Lavorgna, G., H. Ueda, J. Clos, and C. Wu.** 1991. FTZ-F1, a steroid hormone receptor-like protein implicated in the activation of fushi tarazu. Science **252:**848-51.

26. **Lu, T. T., M. Makishima, J. J. Repa, K. Schoonjans, T. A. Kerr, J. Auwerx, and D. J. Mangelsdorf.** 2000. Molecular basis for feedback regulation of bile acid synthesis by nuclear receptors. Mol Cell **6:**507-15.

27. **Luo, X., Y. Ikeda, and K. L. Parker.** 1994. A cell-specific nuclear receptor is essential for adrenal and gonadal development and sexual differentiation. Cell **77:**481-90.

28. **Mallet, D., P. Bretones, L. Michel-Calemard, F. Dijoud, M. David, and Y. Morel.** 2004. Gonadal dysgenesis without adrenal insufficiency in a 46, XY patient heterozygous for the nonsense C16X mutation: a case of SF1 haploinsufficiency. J Clin Endocrinol Metab **89:**4829-32.

29. **Melchior, F., M. Schergaut, and A. Pichler.** 2003. SUMO: ligases, isopeptidases and nuclear pores. Trends Biochem Sci **28:**612-8.

30. **Nachtigal, M. W., Y. Hirokawa, D. L. Enyeart-VanHouten, J. N. Flanagan, G. D. Hammer, and H. A. Ingraham.** 1998. Wilms' tumor 1 and Dax-1 modulate the orphan nuclear receptor SF-1 in sex-specific gene expression. Cell **93:**445-54.

31. **Nelson, P. T., A. G. Hatzigeorgiou, and Z. Mourelatos.** 2004. miRNP:mRNA association in polyribosomes in a human neuronal cell line. Rna **10:**387-94.

32. **Ou, Q., J. F. Mouillet, X. Yan, C. Dorn, P. A. Crawford, and Y. Sadovsky.** 2001. The DEAD box protein DP103 is a regulator of steroidogenic factor-1. Mol Endocrinol **15:**69-79.

33. **Pare, J. F., D. Malenfant, C. Courtemanche, M. Jacob-Wagner, S. Roy, D. Allard, and L. Belanger.** 2004. The fetoprotein transcription factor (FTF) gene is essential to embryogenesis and cholesterol homeostasis, and regulated by a DR4 element. J Biol Chem.

34. **Parker, K. L., and B. P. Schimmer.** 1997. Steroidogenic factor 1: a key determinant of endocrine development and function. Endocrine Reviews **18:**361-77.

35. **Poukka, H., U. Karvonen, O. A. Janne, and J. J. Palvimo.** 2000. Covalent modification of the androgen receptor by small ubiquitin-like modifier 1 (SUMO-1). Proc Natl Acad Sci U S A **97:**14145-50.

36. **Rajendran, R. R., A. C. Nye, J. Frasor, R. D. Balsara, P. G. Martini, and B. S. Katzenellenbogen.** 2003. Regulation of nuclear receptor transcriptional activity by a novel DEAD box RNA helicase (DP97). J Biol Chem **278:**4628-38.

37. **Roberts, L. M., J. Shen, and H. A. Ingraham.** 1999. New solutions to an ancient riddle: defining the differences between Adam and Eve. Am J Hum Genet **65:**933-42.

38. **Sablin, E. P., I. N. Krylova, R. J. Fletterick, and H. A. Ingraham.** 2003. Structural basis for ligand-independent activation of the orphan nuclear receptor LRH-1. Mol Cell **11:**1575-85.

39. **Sachdev, S., L. Bruhn, H. Sieber, A. Pichler, F. Melchior, and R. Grosschedl.** 2001. PIASy, a nuclear matrix-associated SUMO E3 ligase, represses LEF1 activity by sequestration into nuclear bodies. Genes Dev **15:**3088-103.

40. **Shiio, Y., and R. N. Eisenman.** 2003. Histone sumoylation is associated with transcriptional repression. Proc Natl Acad Sci U S A **100:**13225-30.

41. **Swain, A., V. Narvaez, P. Burgoyne, G. Camerino, and R. Lovell-Badge.** 1998. Dax1 antagonizes Sry action in mammalian sex determination. Nature **391:**761-7.

42. **Tallec, L. P., O. Kirsh, M. C. Lecomte, S. Viengchareun, M. C. Zennaro, A. Dejean, and M. Lombes.** 2003. Protein inhibitor of activated signal transducer and activator of transcription 1 interacts with the N-terminal domain of mineralocorticoid receptor and represses its transcriptional activity: implication of small ubiquitin-related modifier 1 modification. Mol Endocrinol **17:**2529-42.

43. **Tanner, N. K., and P. Linder.** 2001. DExD/H box RNA helicases: from generic motors to specific dissociation functions. Mol Cell **8:**251-62.

44. **Tran, P. V., M. B. Lee, O. Marin, B. Xu, K. R. Jones, L. F. Reichardt, J. R. Rubenstein, and H. A. Ingraham.** 2003. Requirement of the orphan nuclear receptor SF-1 in terminal differentiation of ventromedial hypothalamic neurons. Mol Cell Neurosci **22:**441-53.

45. **Vilain, E.** 2000. Genetics of sexual development. Annu Rev Sex Res **11:**1-25.

46. **Wu, W. S., S. Vallian, E. Seto, W. M. Yang, D. Edmondson, S. Roth, and K. S. Chang.** 2001. The growth suppressor PML represses transcription by functionally and physically interacting with histone deacetylases. Mol Cell Biol **21:**2259-68.

47. **Xu, P. L., S. F. Shan, Y. Y. Kong, Y. H. Xie, and Y. Wang.** 2003. Characterization of a strong repression domain in the hinge region of orphan nuclear receptor hB1F/hLRH-1. Sheng Wu Hua Xue Yu Sheng Wu Wu Li Xue Bao (Shanghai) **35:**909-16.

48. **Yan, W., H. Santti, O. A. Janne, J. J. Palvimo, and J. Toppari.** 2003. Expression of the E3 SUMO-1 ligases PIASx and PIAS1 during spermatogenesis in the rat. Gene Expr Patterns **3:**301-8.

49. **Yan, X., J. F. Mouillet, Q. Ou, and Y. Sadovsky.** 2003. A novel domain within the DEAD-box protein DP103 is essential for transcriptional repression and helicase activity. Mol Cell Biol **23:**414-23.

50. **Yang, S. H., E. Jaffray, R. T. Hay, and A. D. Sharrocks.** 2003. Dynamic interplay of the SUMO and ERK pathways in regulating Elk-1 transcriptional activity. Mol Cell **12:**63-74.

51. **Yang, S. H., and A. D. Sharrocks.** 2004. SUMO promotes HDAC-mediated transcriptional repression. Mol Cell **13:**611-7.

Figure Legends

Figure 1. Subfamily V receptors are sumoylated in the hinge region. (A) A schematic of the domain structures and percentage protein identity for *Drosophila* Ftz-F1 and mouse SF-1 and LRH-1 are shown with SUMO sites (S), and phosphorylation sites (P) indicated. The "repression domain" is also shown (R, black square). **(B)** Anti-HA Western blot of COS-7 lysates is shown after transfection with HA-epitope tagged SF-1 or LRH-1 and SUMO1 or GFP-SUMO1. The slower-migrating forms of each receptor are indicated (arrowheads) and all lysates were prepared in the presence of N-ethylmaleimide (NEM), an inhibitor of SUMO isopeptidases. **(C)** Western blots are shown for Y1 whole cell lysates treated with (+) or without (-) 20 mM NEM. Protein was detected with an anti-SF-1 antibody. Upshifted SF-1 after NEM treatment is indicated with arrowhead. **(D)** Anti-HA Western blot of COS-7 cells is shown for empty vector control (pCI), HA-SF-1 wild type or lysine mutants with sumoylated SF-1 (*SF-1) and non-sumoylated SF-1 (SF-1) indicated; SUMO1 was coexpressed in all conditions. A control immunoblot for SUMO1 is shown below. **(E)** Anti-SUMO1 Western blot of HA-immunoprecipitated lysates from COS-7 cells transfected with wild type or lysine mutants of SF-1 is shown with sumoylated SF-1 (arrowhead) and non-specific bands (NS) indicated. A control immunoblot for HA-SF-1 expression is shown below. One µg of each plasmid was added for all transfections. **(F)** In vitro sumoylation of in vitro transcribed-translated ^{35}S-labeled wild type and lysine mutants of SF-1 (1µl) was carried out as described in Materials and Methods. Unmodified SF-1 (SF-1) and sumoylated SF-1 (arrowheads) are indicated.

Figure 2. Sumoylation represses SF-1 transcriptional activity. **(A)** Transcriptional activity of wild type and lysine mutants of SF-1 (50ng) on the aromatase-luciferase reporter (Aro-Luc, 500ng) is shown for both COS-7 (no SUMO1 added), and HepG2 cells (50ng SUMO1 added). Other promoter-luciferase reporters used in HepG2 cells were the 3b-hydroxysteroid dehydrogenase promoter (3bHSD Luc, −153/+2 bp); a synthetic promoter containing tandem SF-1 response elements from the mouse Müllerian Inhibiting Substance promoter (2XRE MIS); and the StAR promoter (StAR Luc, -966/+1), 250 ng of each promoter used. **(B)** Stability of wild type (WT) and lysine mutant (K194R or 2KR) SF-1 proteins in COS-7 cells was determined after metabolic labeling, followed by chase for 0, 2, 5 and 12 hours. An autoradiogram of immunoprecipitated HA-proteins from whole cell lysates is shown with phosphor-image data graphed as the percentage of labeled protein remaining after each chase period; levels of protein at time 0 were taken to be 100%. **(C)** Transcriptional activity is shown for Gal4-SF-1 wild type (pGalWT, aa105-462, 25ng) or Gal4-SF-1-Lys mutants (pGalK119R, pGalK194R, pGal2KR, 25ng) on the Gal4-luciferase reporter (pFR-Luc, Stratagene, 200ng) in COS-7 cells (left panel). Anti-HA Western blotting shows expression levels of the Gal4-SF-1 WT or KR mutants with slower-migrating forms of sumoylated Gal4-SF-1 protein indicated (arrowhead). Transcriptional activity of Gal4-LRH-1 wild type (pGalWT, aa198-561, 25ng) and lysine mutants (pGalK213R, pGalK289R, pGal2KR) are shown (right panel). All luciferase activity is expressed as fold-activation over parent vectors: pCI-neo (C) for panels in A and pM (pGal) for panels in C.

Figure 3. Adding SENP1 increases activity of wild type SF-1, but not the 2KR sumoylation mutant. **(A)** A western blot is shown for COS-7 cells transfected with empty vector (pCI) or wild type SF-1 (1μg each) in the presence or absence of SUMO1 and SENP1 (1μg each); with sumoylated SF-1 (*SF-1), and non-sumoylated SF-1 (SF-1) indicated. **(B)** Effects of increasing amounts of SENP1 (0, 25 and 50ng) are shown for transcriptional activity of Gal4-SF-1 wild type (pGalWT) and Gal4-SF-1-lysine mutants (pGalK119R, pGalK194R, pGal2KR) on the pFR-Luc Gal4 reporter in COS-7 cells. Luciferase activity is expressed as relative light units. Amounts of transfected plasmids are identical to those used in Fig 2C.

Figure 4. SF-1 sumoylation mutants exhibit wild-type localization, promoter occupancy and sensitivity to HDAC inhibitors. **(A)** Nuclear localization is shown for transfected GFP-SF-1 wild-type (WT) and lysine mutants (K119R, K194R, 2KR) in COS-7 cells; 100ng of each plasmid was used and resulted in expression of GFP-SF-1 in 15% of all cells. **(B)** ChIP assays are shown for control vector, N-terminal HA tagged Gal4-SF-1 (WT) or Gal4-lysine mutants in HeLa cells containing integrated Gal4 response elements fused to luciferase using anti-HA or control IgG. **(C)** Trichostatin A (TSA) and **(D)** sodium butyrate (NaBT) effect on transcriptional activity of Gal4-SF-1 wild-type (pGalWT) and lysine mutants (pGalK119R, pGalK194R, pGal2KR) in COS-7. TSA (0, 333 nM) or NaBT (0, 0.1, 1 or 10 mM) was added to cells 12 hrs post transfection and incubated for 24 hrs.

Figure 5. The DEAD-box protein DP103 mediates SF-1 repression by sumoylation at Lys194 and binds to sumoylated SF-1. **(A)** Repression of SF-1 by Dax1 and DP103. Increasing amounts of mDax1 or mDP103 (0, 25, 50, 150ng) were cotransfected in COS-7 cells with Gal4-SF-1 (pGalWT) or the pGalK194R (25ng each) on Gal4-luciferase reporter (pFR-Luc, 250ng). **(B)** Effect of SENP1 on repression by Dax1 and mDP103. Increasing amounts of Dax1 or DP103 (as in A) cotransfected with control vector (pGal) or Gal4-SF1 wild type (pGal-SF-1), with or without SENP1 (25ng), on Gal4-luciferase reporter (as in A). **(C)** GST pulldown assays show binding of [^{35}S]-SF-1 wild-type and lysine mutants (WT, K194R, 2KR) to increasing amounts of GST-hDP103 C-terminal aa414-824 (GST-DPC; 1X, 2X indicate relative amounts used). 10% input and GST controls are indicated. A schematic of human DP103 shows the unique C-terminal region and conserved helicase domain motifs (gray rectangles), including the signature DEAD-box motif (black rectangle). **(D)** GST pulldown assays show binding of in vitro sumoylated [^{35}S]-SF-1 (+E1, *SF-1, upper panel) to increasing amounts of GST-hDP103 C-terminal (1X, 3X, 9X) and to a nonsumoylated SF-1 control made in reactions lacking E1 enzyme (-E1, SF-1, lower panel). Amounts of GST proteins used in panels C, D are shown in Supplemental Data P3.

Figure 6. PIASxa and PIASy are E3-SUMO ligases for SF-1. **(A)** Interactions between full-length SF-1 and Gal4AD-PIASxa -PIAS1, -PIASy fusion proteins are shown in yeast expressing SF-1, driving SF-1 response elements fused to LacZ (bar graph) and when grown on X-gal medium. The bottom sector (SF-1) shows b-galactosidase activity resulting from yeast expressing SF-1 (SF-1), and the empty vector

pGADT7. Other sectors show activity in strains with SF-1 and transformed PIAS fusion proteins, as indicated. **(B)** A mammalian two-hybrid system showing transcriptional activity of wild type pGal4-SF-1 with increasing concentrations of VP16-PIAS fusion proteins (25, 50, 150ng). Empty vector control is shown (pGal). **(C)** Western blots for COS-7 cells cotransfected with wild type HA-SF-1, SUMO1, and with individual PIAS proteins or SENP1 are shown (left panel). The right panel shows a Western blot for HA-SF-1 (WT), lysine or S203A SF-1 mutants after addition of PIASxa and SUMO1; 1μg of each plasmid was added. Sumoylated SF-1 (*SF-1) and non-sumoylated SF-1 (SF-1) are indicated. **(D)** The transcriptional activity of pGal-SF-1 is shown after increasing amounts of PIASxa were cotransfected in COS-7 cells (25, 50, 150ng), in the absence or presence of SUMO1 (25ng) with the Gal4-luciferase reporter (pFR-Luc, 250ng). Empty vector control is shown (pGal). **(E)** Activity of the double sumoylation mutant of SF-1 (pGal-2KR) is shown after increasing amounts of PIASxa, as described in C.

Figure 7. DP103 interacts with sumoylated SF-1 in vivo and promotes PIASy-mediated SF-1 relocalization into nuclear bodies. (A) Coimmunoprecipitation of sumoylated SF-1 from COS-7 cells transfected with wild-type HA-SF-1, SUMO1 and combinations of FLAG-DP103, T7-PIASy and SENP1 (1μg each) using anti-FLAG-M2 agarose beads. Western blots for HA-SF-1 (3% input lysate, upper panel) and immunoprecipitated (IPed) DP103 (10% IPed protein, middle panel) are shown with sumoylated SF-1 indicated (*SF-1). SF-1 (black arrowhead) and non-specific bands (NS) are indicated in an anti-HA Western blot of IPed DP103 protein (lower panel). **(B)** Western blots are shown of COS-7 cells cotransfected with wild type HA-SF-1, SUMO1

and individual PIAS proteins or empty vector (Cont.) with (+) or without (-) DP103 (1 mg). Sumoylated SF-1 (*SF-1) and non-sumoylated SF-1 (SF-1) are indicated. **(C)** Nuclear localization of GFP-SF-1 transfected into COS-7 cells is shown with different combinations of FLAG-DP103, T7-PIASy, FLAG-PIAS1 and SENP1, as indicated. All cells were transfected with SUMO1 (100ng). **(D)** Subnuclear signals are shown for wild-type GFP-SF-1 (green) and indirect immunofluorescence is shown for T7-PIASy (red), or FLAG-hDP103 (red). Colocalization of GFP-SF-1 and T7-PIASy signals are shown in the merged figure (upper panels), or the endogenous DP103 signal (lower panels) are indicated (arrowheads). Staining for endogenous SF2/ASF (marker for splicing speckles) or Sp100 (marker for PML-NBs) is shown (red). Note that no positive staining is observed for endogenous Sp100. In all conditions, cells were transfected with 100 ng each of GFP-SF-1, PIASy, hDP103, and SUMO1.

Supplemental Data P1

Interaction of SF-1 and LRH-1 with Ubc9. **(A)** Ubc9 interacts strongly with SF-1 and LRH-1 in vivo. ß-galactosidase activity was measured in yeast expressing stably integrated mouse SF-1 and driven by four tandem copies of the SF-1 response element from the Müllerian Inhibiting Substance promoter. Values are shown for SF-1/LRH-1 binding alone (-), with the control pGADT7 vector (pGADT7) or with Ubc9 fused to the Gal4AD (AD-UBC9) (left and middle panels). Values for yeast with no integrated SF-1 or LRH-1 are also shown (right panel). **(B)** Mouse Ubc9 was cloned into GST expression vector pGEX4T1 and expressed in BL21 after induction with IPTG. Equivalent amounts of GST-Ubc9 and GST proteins were bound to radiolabeled SF-1 in 20mM Tris pH 8.0,

0.1M NaCl, 1mM EDTA, 1mM DTT, 0.01% NP40, 10% glycerol, 0.1mM PMSF at 4°C

for 3 hrs. Beads were washed, and analyzed by SDS-PAGE gel, followed by

autoradiography.

Supplemental Data P2

Supplemental Data for Figure 4. Anti-HA immunoblot of COS-7 cells co-transfected

with wild-type HA-SF1 and SUMO1 (1 μg each) with increasing amounts of PIASxa or

PIASg (0, 0.15, 0.5, 1.5 mg), with sumoylated SF-1 (*SF-1) and non-sumoylated SF-1
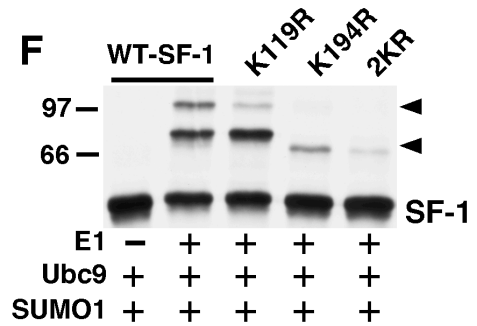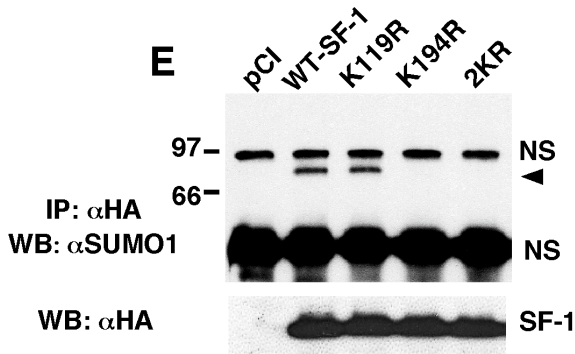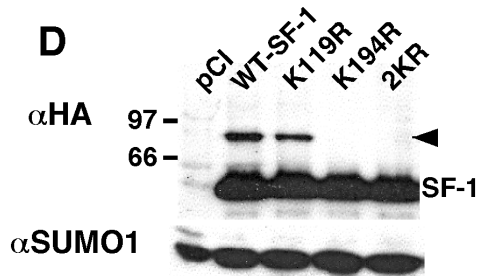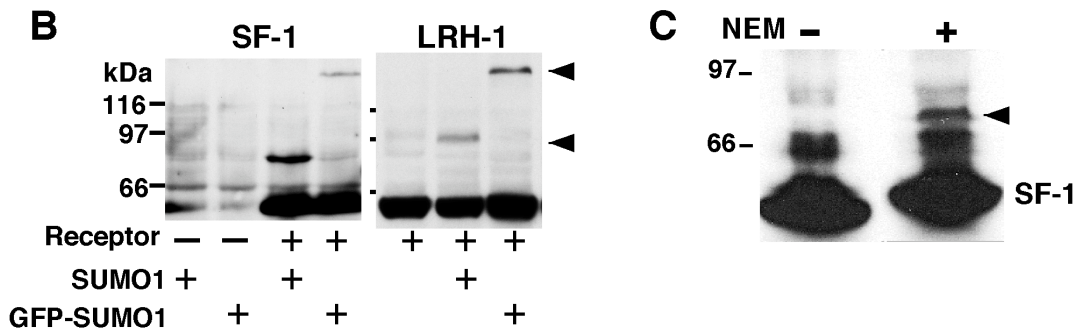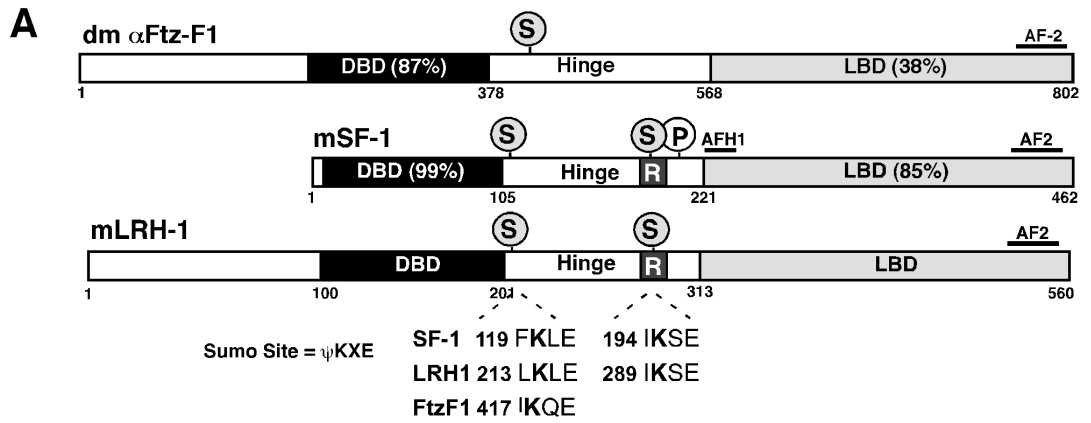
(SF-1) indicated.

Supplemental Data for Figure 5. Effect of trichostatin A (TSA: 0, 10, 100, 333 nM; *left*

panel) and sodium butyrate (NaBT: 0, 0.1, 1, 10 mM; *right* panel) on transcriptional

activity of Gal4SF-1 wild-type (pGalSF-1), lysine mutants (pGalK119R, pGalK194R,

pGal2KR) or empty vector (pGal) in HeLa cells containing integrated Gal4 response

elements driving luciferase, with data expressed as raw light units uncorrected for beta-

galactosidaase activity.
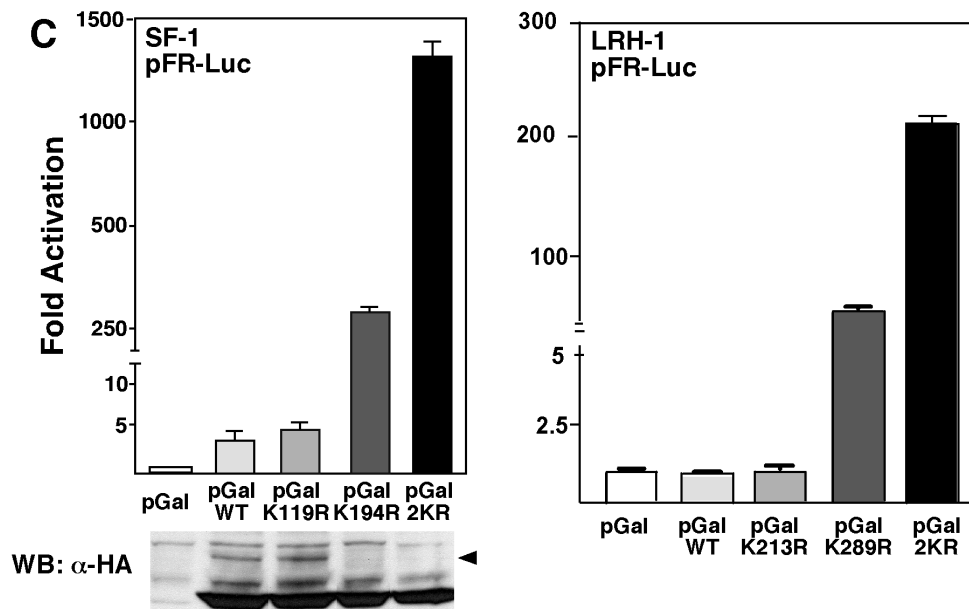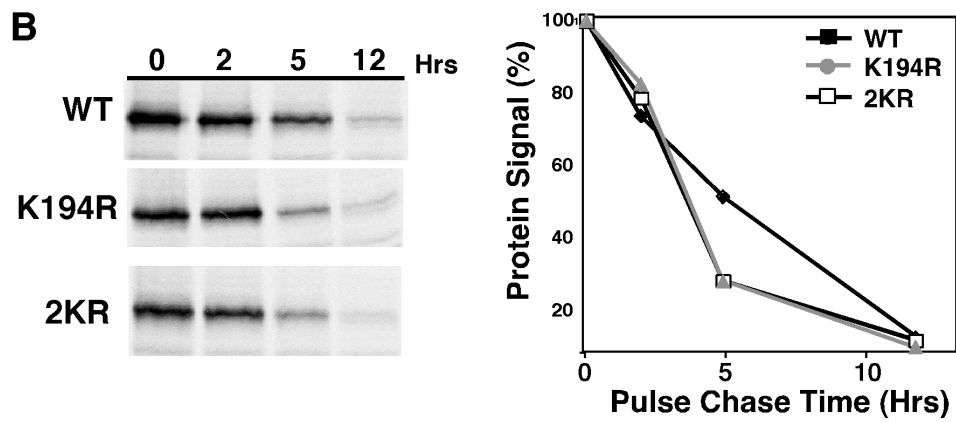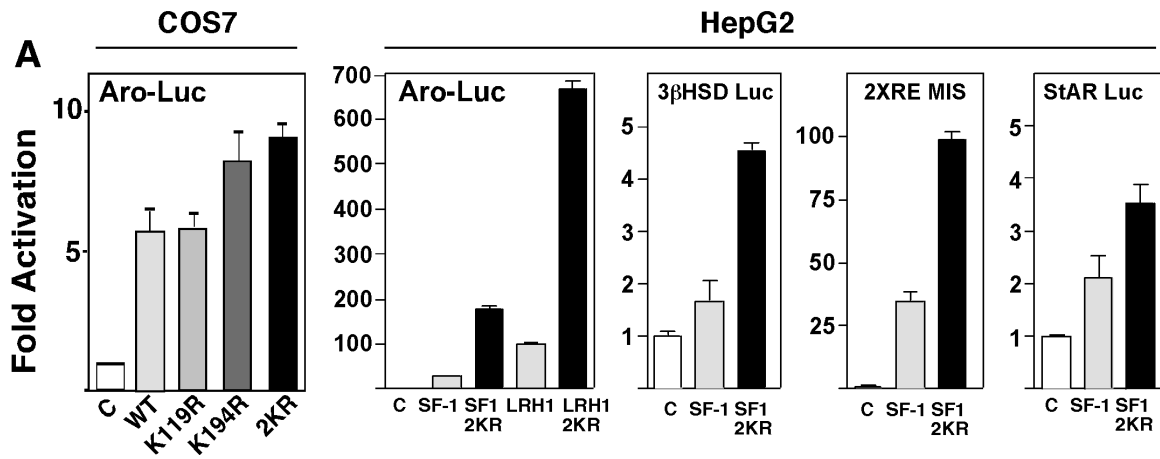

Supplemental Data P3

Supplemental Data for Figure 6. Coomassie-stained 10% SDS-PAGE gel of GST and

increasing amounts of GST-DP103 C terminal (1X, 3X, 9X, GST-DPC) used in GST-

pull-downs.

Supplemental Data for Figure 7. *Left* panels: Control Western blots for PIAS

expression: anti-FLAG immunoblot for FLAG tagged PIAS1, PIASxa and PIAS3 (*top*)

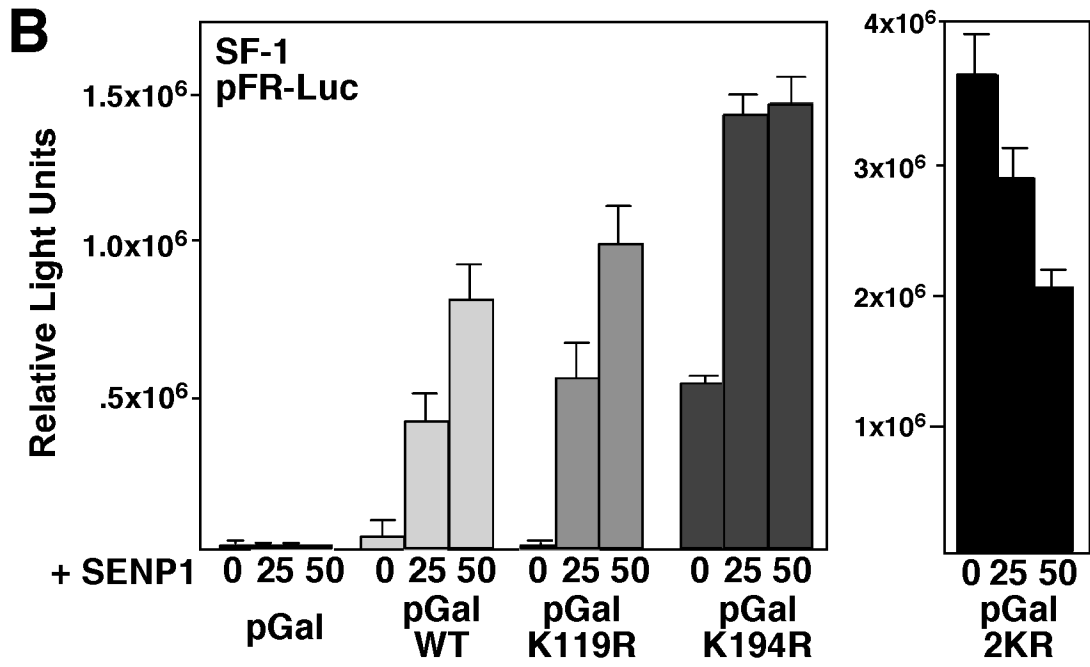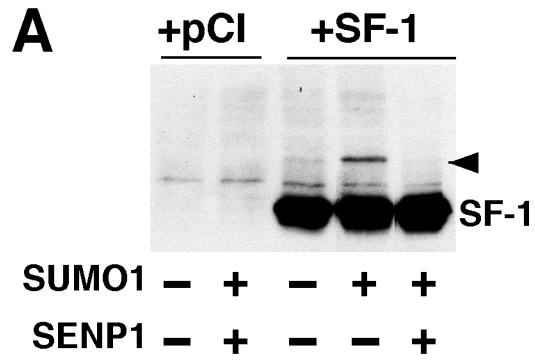and anti-T7 tag immunoblot for T7 tagged PIASg (*bottom*). Equal amounts of total

protein were loaded. *Right* panel: Anti-HA immunoblot of lysates from COS-7 cells co-transfected with wild type HA-SF-1, SUMO1 (1µg each) and PIASxa alone (0.15 µg) with increasing amounts of DP103 (0, 0.15, 0.5, 1.5 mg). Sumoylated SF-1 (*SF-1) and non-sumoylated SF-1 (SF-1) are indicated.
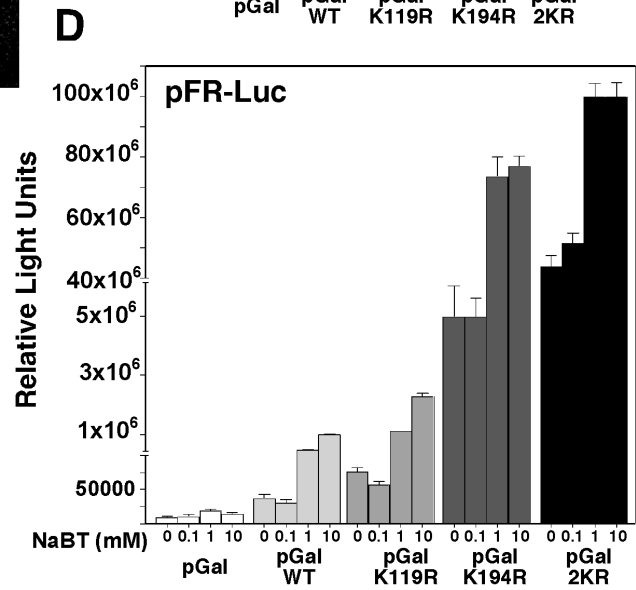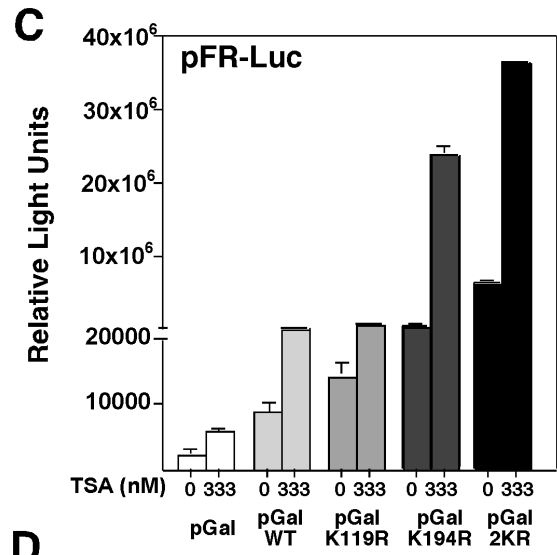
A

dm αFtz-F1

| DBD (87%) | Hinge | LBD (38%) |

S
AF-2
1   378   568   802

mSF-1

| DBD (99%) | Hinge | R | LBD (85%) |

S   S P AFH1   AF2
1   105   221   462

mLRH-1

| DBD | Hinge | R | LBD |

S   S   AF2
1   100   201   313   560

Sumo Site = ψKXE

SF-1   119 FKLE   194 IKSE
LRH1   213 LKLE   289 IKSE
FtzF1 417 IKQE

B

SF-1   LRH-1

kDa
116—
97—
66—

Receptor   —   —   +   +   +   +   +
SUMO1   +   +   +
GFP-SUMO1   +   +   +

C

NEM   —   +

97—
66—

SF-1

D

pCI WT-SF-1 K119R K194R 2KR

αHA

97—
66—

SF-1

αSUMO1

E

pCI WT-SF-1 K119R K194R 2KR

97—   NS
66—

IP: αHA
WB: αSUMO1   NS

WB: αHA   SF-1

F

WT-SF-1   K119R K194R 2KR

97—
66—

SF-1

E1   —   +   +   +   +
Ubc9   +   +   +   +   +
SUMO1   +   +   +   +   +

Lee et al. - Figure 1

71

Lee et al. - Figure 2

A

+pCI   +SF-1

◄ SF-1

| SUMO1 | − | + | − | + | + |
| SENP1 | − | + | − | − | + |

B

SF-1
pFR-Luc

Relative Light Units

1.5x10⁶
1.0x10⁶
.5x10⁶

+ SENP1   0 25 50   0 25 50   0 25 50   0 25 50

pGal   pGal WT   pGal K119R   pGal K194R

4x10⁶
3x10⁶
2x10⁶
1x10⁶

0 25 50
pGal 2KR

Lee et. al. - Figure 3

73

A

WT K119R

K194R 2KR

B

	Cont.	WT	K119R	K194R	2KR
αHA					
IgG					

C

pFR-Luc

Relative Light Units

40x10⁶
30x10⁶
20x10⁶
10x10⁶
20000
10000

TSA (nM) 0 333   0 333   0 333   0 333   0 333

pGal pGal WT pGal K119R pGal K194R pGal 2KR

D

pFR-Luc

Relative Light Units

100x10⁶
80x10⁶
60x10⁶
40x10⁶
5x10⁶
3x10⁶
1x10⁶
50000

NaBT (mM) 0 0.1 1 10   0 0.1 1 10   0 0.1 1 10   0 0.1 1 10   0 0.1 1 10
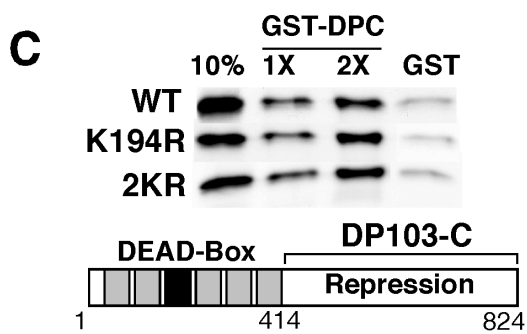
pGal pGal WT pGal K119R pGal K194R pGal 2KR

Lee et al - Figure 4

74

Lee et al. - Figure 5

75

A

β-Gal Activity

+SF-1: Cont. PIASy PIASxα PIAS1

SF-1 PIAS1
SF-1 PIASxα
SF-1 PIASy
SF-1

B

Fold activation

+pGal   + + + +   + + + +   + + + +
+pGalSF-1   + + + +   + + + +
+VP16-PIASxα   −   ◢   −   ◢
+VP16-PIAS1   −   ◢   −   ◢

C

   −   +SENP1 +PIAS1 +PIASxα +PIASy +PIAS3   WT K119R K194R 2KR S203A

◄ SF-1   ◄ SF-1

+SUMO1   **+SUMO1/+PIASxα**

D

Relative Light Units

+PIASxα   0   25   50   150   0   25   50   150   0   25   50   150
+SUMO1   25   25   25   25

pGal   **+pGal SF-1**

E

+PIASxα   0   25   50   150   0   25   50   150
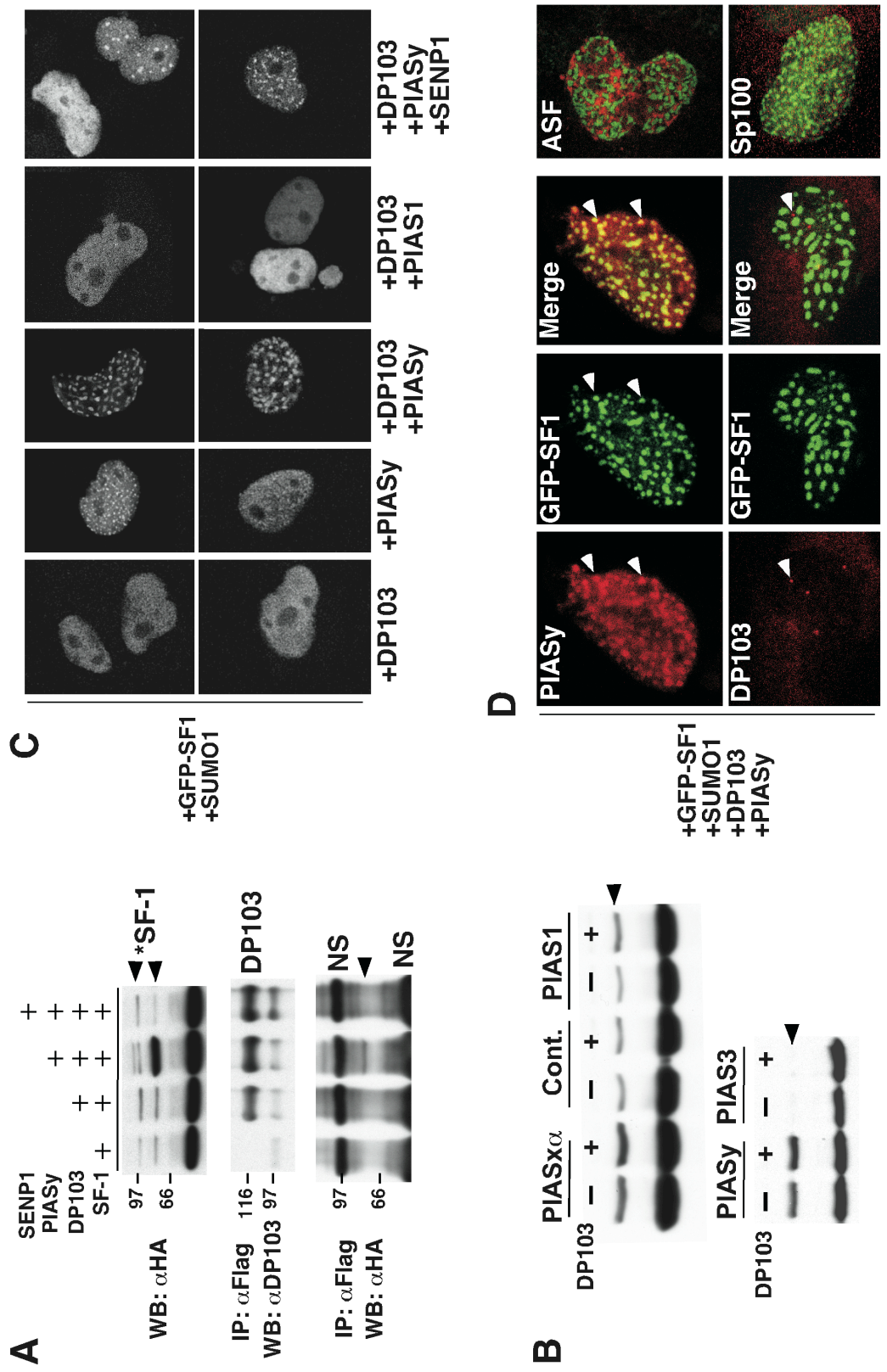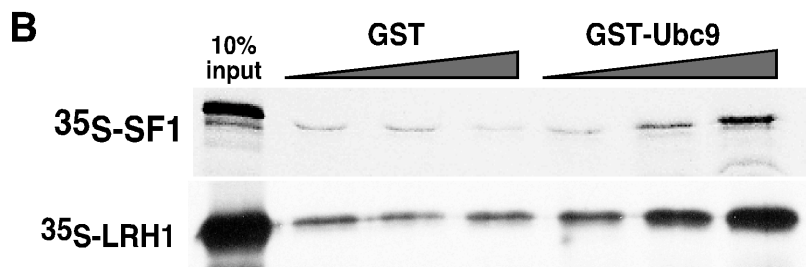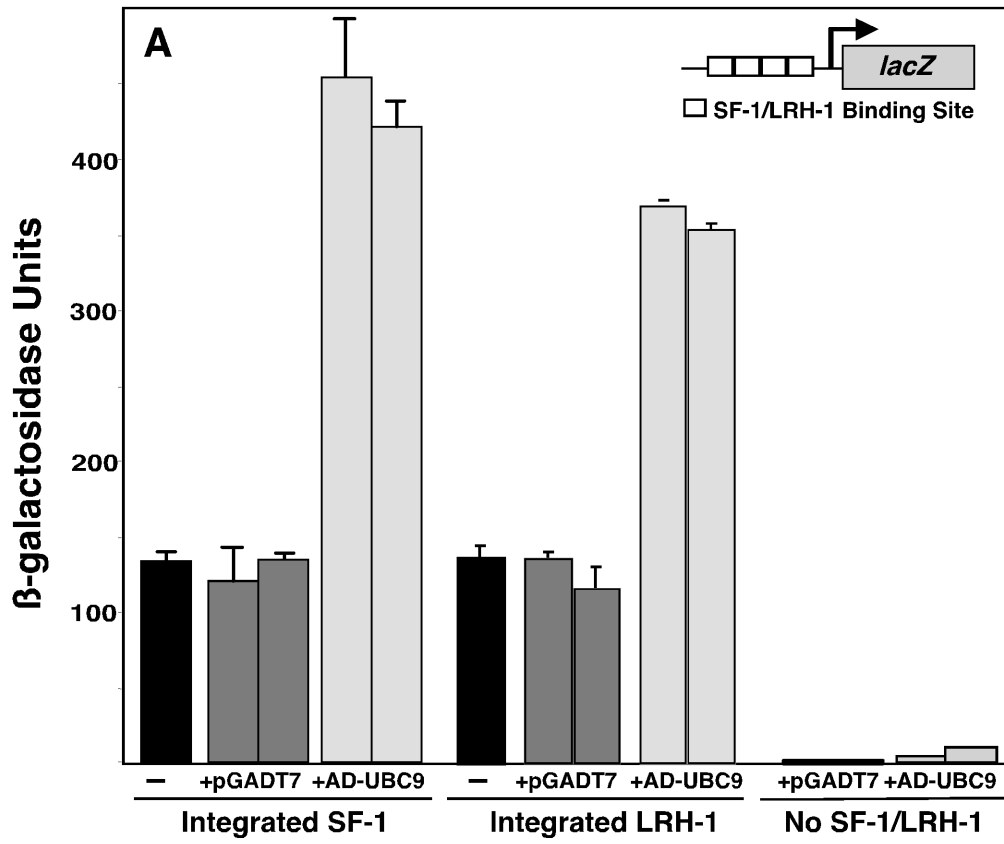+SUMO1   25   25   25   25

+pGal 2KR

Lee et al. - Figure 6
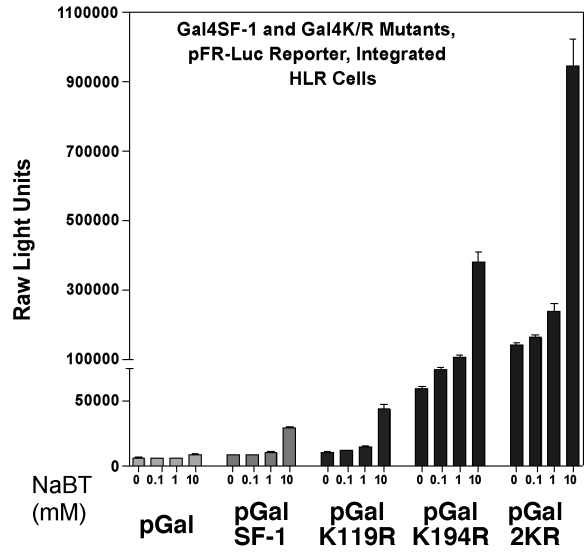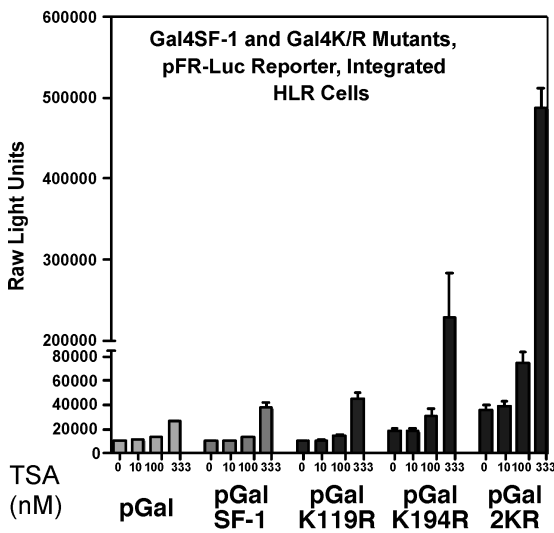
Lee et al. - Figure 7

77

Supplemental Data P2

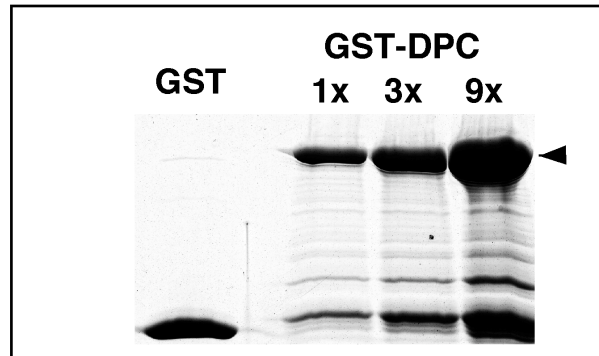## Supplemental Data for Figure 4
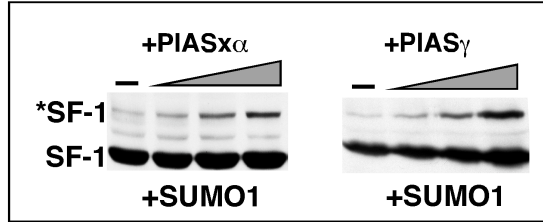
Supplemental Data for Figure 5

Supplemental Data P3

Supplemental Data for Figure 6

Supplemental Data for Figure 7
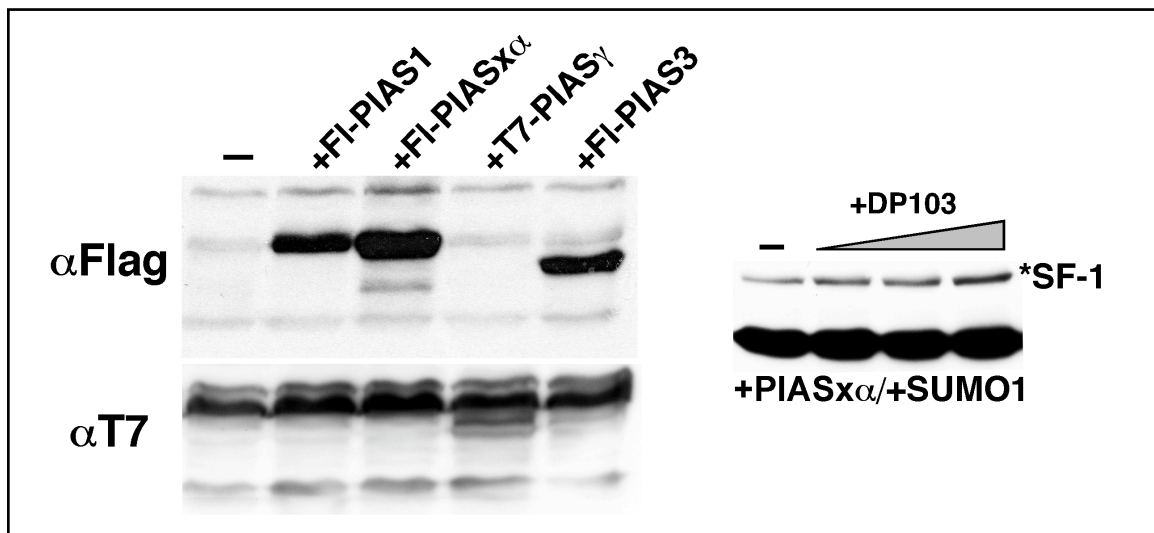
CHAPTER IV

Sumoylation Selectively Regulates Steroidogenic Factor 1 (NR5A1) Target Genes.

Lioudmila A. Campbell[1,3]**, Matthew D. Show**[1]**, Jared G. Ingraham**[1]**, Jeremy Flinders**[2]**, John D. Gross**[2]**, and Holly A. Ingraham**[1]

Department of Chemical and Molecular Pharmacology and Physiology[1],

Department of Pharmaceutical Chemistry[2],

Graduate Program in Biological Sciences[3],

1550 4th Street, Rock Hall

Mission Bay Campus

University of California, San Francisco
Box 0444, San Francisco, CA 94143-2611

Corresponding author: Holly A. Ingraham

E-mail   Holly.Ingraham@ucsf.edu

Phone: (415) 476-2731
Fax: (415) 514-4459

Short title: SUMO1 regulates SF-1 DNA-binding

Abstract

Steroidogenic Factor 1 (SF-1, NR5A1) is a transcription factor that belongs to the nuclear hormone receptor superfamily. SF-1 activity is repressed via posttranslational modification with SUMO1 (Small Ubiquitin-like Modifier) on two conserved lysines. Here, we examine molecular consequences of SF-1 sumoylation. We show that SF-1 sumoylation at Lysine 194 reduces in vitro MAP kinase phosphorylation at Ser203, but does not affect interactions with SF-1 coregulators or the structure of SF-1 LBD, as demonstrated by nuclear magnetic resonance spectroscopy. Additionally, we discover a regulatory loop between SF-1 sumoylation at Lysine 119 and SF-1 DNA-binding. DNA-binding inhibits SF-1 sumoylation at Lysine 119, and SF-1 sumoylation at Lysine 119 abolishes DNA-binding at low affinity SF-1 binding sites. These studies suggest that sumoylation may contribute to differential target gene regulation by SF-1, highlighting importance of posttranslational modifications in regulating nuclear receptor function.

Introduction

Post-translational modification with ubiquitin-like proteins has emerged as an important regulatory mechanism in many aspects of cellular function[1]. In particular, post-translational modification of transcription factors with a Small Ubiquitin-like Modifier (SUMO) results in repression of target genes[2]. Sumoylation occurs on lysines within consensus φKxE sites through an enzymatic mechanism analogous to ubiquitination[3]. Whether specific SUMO E3 ligases exist is an intriguing question as in vitro sumoylation occurs in their absence; however, several proteins exhibit SUMO E3 ligase activity in cells with the largest group belonging to the Protein Inhibitors of Activated STATs family[4]. Sumoylation is easily reversible by the action of SUMO isopeptidases (SUSPs or SENPs) with seven members identified in humans so far[5].

Significant progress has been made in structural studies of SUMO pathway enzymes, providing important insights into SUMO biochemistry and substrate selection[6,7]. And, the discovery of SUMO-interacting motif elucidated how SUMO acts as a molecular tag to sequester transcription factors into PML bodies[8,9]. Yet whether SUMO represses transcription factors through specific molecular changes to the structure of transcription factor protein-protein and protein-DNA interfaces remains less clear. Structural analysis of a sumoylated transcription factor Ets-1 revealed a "beads-on-a-string" conformation where SUMO1 and the substrate behaved as two independent domains[10]. Thus far, there is only one report where sumoylation modulates DNA-binding function of a DNA repair enzyme, thymine-DNA glycosylase, through a conformational change that results in dissociation of the enzyme from DNA[11,12]. Similarly, SUMO modification is known to

reduce DNA-binding ability of several transcription factors; it is yet to be determined whether this is due to specific effects of SUMO on transcription factor structure[13-16].

Nuclear hormone receptors are unique among sequence-specific transcription factors because, in addition to the DNA-binding domain (DBD), they contain a ligand-binding domain (LBD) that responds to lipophilic molecules[17]. Most nuclear hormone receptors are sumoylated, and SUMO inhibits their transcriptional activity or even leads to transrepression[18-21]. Additionally, nuclear hormone receptors such as Peroxisome Proliferator-activated Receptor γ (PPARγ, NR1C) and Estrogen-Related Receptor α (ERRα, NR3B) are also sequentially phosphorylated and sumoylated on the extended sumoylation consensus sites φKxExxSP, demonstrating a complex regulatory relationship between the two post-translational modifications[22-24]. Recent advancements in obtaining nuclear receptor LBD and DBD structures have led to improved mechanistic understanding of how these domains carry out their function[17]. Therefore, it is now feasible to explore how sumoylation might affect ligand-binding and coactivator recruitment to the LBD and the DNA-binding ability of nuclear receptors at the molecular level.

Nuclear hormone receptor Steroidogenic Factor 1 (SF-1, NR5A) coordinates male sexual differentiation and adrenal organogenesis during development and controls the expression of steroidogenic enzymes in the adult[25,26]. Similar to other nuclear hormone receptors, SF-1 transcriptional activity is potently repressed by sumoylation[20,27,28]. The location of SF-1 sumoylation sites next to the DBD and the LBD suggests that

sumoylation may directly influence the function of the nearby domain. Additionally, Mitogen-Activated Protein kinase (MAPK) phosphorylation site in SF-1 is located in close proximity to the LBD sumoylation site resembling a phospho-sumoyl switch site described for other proteins[20,22,29]. Recent SF-1 LBD and DBD structures provide molecular-level understanding of how SF-1 is regulated by ligand, coactivators, and DNA-binding[30-33]. The discovery of a phospholipid in the pocket of SF-1 LBD suggests a possibility of ligand-dependent regulation[31-33]. And, the DBD-DNA complex structure reveals the contribution of specific DBD features, the A-box loop and the Ftz-F1 helix, to recognition of the binding site[30,34]. In light of these findings, SF-1 represents a good model system to understand how modification with SUMO affects structure and function transcription factors. Here, we combined biochemical and structural approaches to examine the molecular effects of sumoylation on SF-1 LBD structure and interactions with coregulators, and SF-1 DNA-binding activity. Our studies suggest that SUMO1 regulates SF-1 activity in a promoter-specific manner by inhibiting SF-1 phosphorylation and DNA-binding activity. These results offer a molecular explanation for the repressive effect of sumoylation on SF-1 transcriptional activity.

Methods

Plasmids

Mouse SF-1 fragment containing SF-1 LBD (aa178-462, cysteine mutant, described previously)[31] was cloned using *BamHI-XhoI* sites into the bacterial expression vector pBH4 (modified pET-based vector, described previously)[35] and the GST-fusion vector pGEX6P2 (Clontech). For MBP-fusion proteins, mouse SF-1 full length (aa1-462),

Hinge-LBD (aa106-462), and DBD (aa1-122) were cloned into the pMALp2X vector

(New England Biolabs) modified to contain N-terminal tobacco etch virus protease

cleavage site using *EcoRI/XbaI* (full length SF-1), *EcoRI-HindIII* (Hinge-LBD), and

EcoRI (DBD) sites. Human SUMO1 (aa1-97) and mouse Ubc9 (aa1-158) were cloned

into the bacterial expression vector pBH4 using *BamHI-XhoI* sites. S203D and S203E

SF-1 phosphomimics were created using pBH4-SF-1 LBD as a template and

K100D/R103D and R92Q SF-1 DBD point mutants were created using pMAL-SF-1

DBD as a template by PCR mutagenesis (QuikChange Site-Directed Mutagenesis kit,

Stratagene). All DNA concentrations were measured using NanoDrop spectrophotometer

(NanoDrop Technologies), and validated by DNA sequencing.

Recombinant protein expression

Proteins were expressed using *E. coli* BL21 (lDE3) cells grown in LB media and lysed

using a microfluidizer (Microfluidics, Inc). His_6-^{15}N-SUMO1 was expressed in M9

medium containing 1g/L $^{15}NH_4Cl$ (Cambridge Isotope Labs). His_6-^{13}C-SF-1 LBD was

expressed in M9 medium containing 1g/L unlabeled NH_4Cl with 100mg/L g-^{13}C-α-

ketoisovalerate (Cambridge Isotope Labs) and 50mg/L g-^{13}C-α-ketobuterate (Cambridge

Isotope Labs) added to the culture 30 minutes prior to induction with IPTG. His_6-

hSUMO1 and wild type and phosphomimic His_6-SF-1 LBD proteins were purified by

TALON chromatography (Clontech) followed by anion exchange chromatography using

HiTrapQ column (GE Healthcare) in a buffer containing 20mM HEPES [pH7.5], 1mM

EDTA, 2mM CHAPS, eluted with an ammonium acetate gradient and concentrated by

ultrafiltration. All MBP-fusion proteins were purified using amylose resin (New England

Biolabs) according to the manufacturer's protocol except the MBP buffer contained
400mM NaCl and 50mM $ZnCl_2$ for SF-1 DBD-containing constructs. Full length MBP-
SF-1 and MBP-Hinge-LBD SF-1 were subsequently purified by size exclusion
chromatography in the MBP buffer using Superdex 200 column (GE Healthcare) and
concentrated by ultrafiltration. Wild type and mutant SF-1 DBD proteins were further
purified by anion exchange chromatography as described above. Recombinant His_6-hE1
(SAE1/SAE2) and recombinant His_6-mUbc9 were expressed and purified as described
previously except TALON chromatography (Clontech) was used, and His_6-tags were not
removed[36]. GST fusion proteins were expressed and purified as described previously[35].
Sumoylated His_6-SF-1 LBD protein was purified by anion exchange chromatography as
described above for His_6-SF-1 LBD, followed by size exclusion chromatography in a
buffer containing 50mM HEPES [pH7.0], 150mM NaCl, 50mM Arginine, 50mM
Glutamine, 2% glycerol (v/v), 2mM CHAPS, 2mM DTT using Superdex 200 column
(GE Healthcare).

In-vitro sumoylation assays

SF-1 sumoylation assays were carried out in 30μL (small-scale) or 10μL (large-scale)
with 0.1mM E1, 10mM Ubc9, 30mM SUMO1 in a sumoylation buffer containing 50mM
Tris-HCl [pH8.0], 100mM NaCl, 10mM $MgCl_2$, 10mM ATP, 2mM DTT at 4°C,
overnight. For DNA inhibition assays, DNA at varying concentrations was incubated
with SF-1 DBD protein in the sumoylation buffer at room temperature for 30 minutes
followed by addition of E1, Ubc9, and SUMO1 at the above concentrations and
incubation overnight at 4°C. Reactions were resolved on 4-12% NuPAGE Bis-Tris gel

(Invitrogen) in NuPAGE MOPS buffer (Invitrogen), stained either with Coomassie or with SYPRO Red dye (Molecular Probes) according to the manufacturer's protocol, and visualized using Typhoon laser scanner (Molecular Dynamics). The signal was quantified using NIH Image and plotted using GraphPad Prism software. GST-SF-1 protein bound to glutathione-agarose 4B beads (Pharmacia) was sumoylated overnight at 4°C in the sumoylation buffer containing 2.5mM E1, 8mM Ubc9, 40mM SUMO1, and the extent of sumoylation was checked by SDS-PAGE and Coomassie staining. Sumoylated GST-SF-1 was extensively washed with a buffer containing 20mM Tris-HCl [pH8.0], 100mM NaCl, 10% glycerol (v/v), 0.01% NP-40 (v/v), 1mM EDTA, 1mM DTT to remove sumoylation enzymes and free SUMO1.

Nuclear Magnetic Resonance Spectroscopy

^{15}N-HSQC and and ^{13}C-HSQC spectra for His$_6$-SUMO1 (aa1-97) and sumoylated His$_6$-SF-1 (aa178-462) were recorded on a Bruker 800MHz ^{1}H frequency spectrometer equipped with a triple-resonance cryogenic probe. The samples were in a buffer containing 50mM HEPES [pH7.0], 150mM NaCl, 50mM Arginine, 50mM Glutamine, 2% glycerol (v/v), 2mM CHAPS, 2mM TCEP, and 10% (v/v) D$_2$O. Final protein concentration was estimated to be 200μM. NMR data were processed with NMRPipe and spectral analysis was carried out using Sparky[37,38]. Assignments for residues in SUMO1 were previously reported[39,40].

GST Pulldown Assay

Sumoylated or unsumoylated GST-SF-1 protein was incubated with full-length [^{35}S]-GRIP1 and [^{35}S]-Dax1 produced using the T7 TNT-coupled transcription/translation system (Promega) in a buffer containing 20mM Tris-HCl [pH8.0], 100mM NaCl, 10% glycerol (v/v), 0.01% NP-40 (v/v), 1mM EDTA, 1mM DTT, 0.1mM phenylmethylsulfonyl fluoride, and Complete protease inhibitor cocktail (Roche) overnight at 4°C. Bound proteins were washed three times with the above buffer, eluted by boiling in SDS loading buffer, separated by SDS-PAGE, dried, and exposed to film (Pierce).

In Vitro Kinase Assay

Purified recombinant in vitro sumoylated or unmodified His$_6$-SF-1 LBD protein was incubated with 20 units of Erk2 kinase (New England Biolabs) in a buffer containing 50mM Tris-HCl [pH7.5], 10mM MgCl$_2$, 1mM EGTA, 2mM DTT, 0.01% Brij35, 200mM [γ^{32}P]-ATP and excess cold ATP to a final specific activity of 500mCi/mmol. Reactions were performed at 30°C for 0, 10, 20, 40, 80, and 160 minutes, stopped by addition of SDS loading buffer, and resolved by SDS-PAGE. The gel was dried, exposed to film (Pierce) or phosphoimager, and quantified using a Storm 860 (Molecular Dynamics). The data was plotted and curve-fitted using GraphPad Prism software.

Electrophoretic Mobility Shift Assays

EMSAs were performed to measure the DNA-binding ability of sumoylated or unmodified SF-1 using in vitro sumoylation reactions containing purified recombinant

SF-1 DBD and double-stranded oligonucleotides taken from the promoters of SF-1 target genes (see Supplementary Table 1 for sequences). In short, 2-5μL of the in vitro sumoylation reaction (300-700nM unmodified SF-1 DBD) was incubated in a total volume of 20μL at room temperature for 30 minutes in a buffer containing 50mM Tris [pH8.0], 150mM NaCl, 10mM MgCl$_2$, 10mM DTT, 10mM ATP, and 1mM of double-stranded oligonucleotide. Ulp1 samples were subsequently treated with 0.5μL of recombinant Ulp1 (LifeSensors) at room temperature for 20 minutes. 10μL of the EMSA reaction was loaded on a native 6% polyacrylamide gel, and electrophoresis was carried out in 1XTBE buffer at room temperature. The gel was analyzed using fluorescence-based EMSA kit (Molecular Probes). To detect total protein, the remaining 10μL of the EMSA reaction was added to SDS-PAGE loading dye, boiled, and resolved on a 4-12% NuPAGE Bis-Tris gel (Invitrogen) in NuPAGE MOPS buffer (Invitrogen), stained with SYPRO Red dye (Molecular Probes) according to the manufacturer's protocol. Gels stained with fluorescent dyes were visualized using Typhoon laser scanner (Molecular Dynamics).

Results

Recombinant SF-1 DBD and LBD are sumoylated in vitro.

Recent studies demonstrate that sumoylation is an important post-translational modification transcriptional regulation comes from cellular studies that relied on loss-of-function and overexpression approaches[20]. Yet, no gain-of-function approaches have emerged to study functional consequences of protein sumoylation in cellular or genetic systems. To understand molecular effects of sumoylation on function of SF-1 DBD and

90

LBD, we set up an in vitro sumoylation assay with recombinant SF-1 protein. In vitro

sumoylation assay components have been previously described, and sumoylation of SF-1

occurs in vitro in the absence of SUMO E3 ligases[36]. SF-1 DBD and LBD were

expressed as separate proteins, including proximal sumoylation sites at Lys119 for DBD

and Lys194 for LBD (Figure 1A and B). Bacterially-expressed full-length SF-1 of

desired purity could not be obtained to effectively investigate its sumoylation. SF-1 DBD

was expressed as an N-terminal maltose binding-protein (MBP) fusion to enhance protein

solubility and stability[34]. SF-1 is efficiently sumoylated in cellular systems and in the in

vitro assay on two sites, Lys119 and Lys194 (Figure 1B)[20,27,28]. Sumoylation of SF-1 was

abolished with addition of yeast SUMO isopeptidase, Ulp1, demonstrating that in vitro

SF-1 sumoylation is reversible (Figure 1B). Additionally, we established that

sumoylation at Lys119 or Lys194 does not depend on the neighboring functional domain.

Both SF-1 sites were sumoylated in constructs lacking SF-1 DBD (Figure 1B) or SF-1

LBD (data not shown) suggesting that sumoylation of Lys119 and Lys194 does not

require the proximal DBD and LBD domains, respectively.

Sumoylation at Lys194 does not change SF-1 LBD structure.

Transcriptional regulation of nuclear hormone receptors involves structural changes in

the LBD elicited by binding of ligand or coregulator proteins[17]. For SF-1, neither the

phospholipid ligand or corepressor peptide binding does not result in large movement of

SF-1 LBD[31]. We asked by using nuclear magnetic resonance spectroscopy (NMR)

whether enzymatic conjugation of SUMO1 at Lys194 could induce a conformational

change in SF-1 LBD. Nearly complete overlap of the HSQC spectrum of free [15]N-

SUMO1 compared to ^{15}N-SUMO1 conjugated to unlabeled SF-1 LBD revealed no non-covalent interactions between SUMO1 and SF-1 LBD (Figure 2A). The only observed changes in SUMO1 were Gly97, the site of isopeptide bond formation with Lys194 in SF-1, and in the neighboring Gly96. Similarly, no interaction between SUMO1 and the N-terminal domain of another transcription factor Ets-1 was detected by NMR in the structure of sumoylated Ets-1[10]. We also saw no changes in free SUMO1 structure upon addition of excess free SF-1 LBD (data not shown). Additionally, in a reciprocal experiment using [13]C- ILV-methyl-labled SF-1 LBD, no changes were detected in the HSQC spectrum of ^{13}C-ILV-methyl-SF-1 LBD conjugated to unlabeled SUMO1 as compared to the spectrum of [13]C- ILV-methyl-SF-1 LBD alone (Supplementary Figure 1). Due to high α-helical content of the SF-1 LBD resulting in poor spectrum dispersion, we did not attempt backbone assignment of SF-1 LBD. At present, we cannot rule out the possibility that sumoylation at Lys194 indirectly changes the structure of SF-1 LBD without detectable interaction with SUMO1. Consistent with our NMR data, sumoylation at Lys194 did not abolish the interaction with full-length coactivator GRIP1 and did not enhance recruitment of full-length corepressor DAX1 (Figure 2B). The slight reduction in coregulator recruitment to the sumoylated SF-1 LBD is, most likely, due to non-specific blocking of the coactivator groove by SUMO1 or unfavorable electrostatic interactions between SUMO1 and full-length coregulators due to highly negatively charged nature of SUMO1. Taken together, these data suggest that molecular mechanism of SF-1 repression by Lys194 sumoylation is unlikely to involve a conformational change in SF-1 LBD or altered coregulator recruitment.

Sumoylation at Lys194 reduces in vitro phosphorylation by MAP kinase.

Close proximity of phosphorylation and sumoylation sites in numerous transcription factors suggests a regulatory relationship between theses post-translational modifications[22,41]. The SF-1 LBD sumoylation site at Lys194 is proximal to SF-1 MAP kinase phosphorylation site at Serine 203 (Ser203) (Figure 3A). We wanted to examine the whether Lys194 and Ser203 are a part of the phospho-sumoyl switch in SF-1. Phosphorylation of SF-1 on Ser203 enhances coactivator recruitment and increases transcription of downstream genes[29,35]. Homologous Ser280 is not phosphorylated by MAP kinase in a close relative of SF-1, Liver Receptor Homolog 1 (LRH-1); yet, sumoylation of LRH-1 was slightly enhanced by acidic residues downstream from the sumoylation site[42,43]. Our previous cellular experiments suggested that phosphorylation of SF-1 at Ser203 was not required for sumoylation at Lys194. Mutation of Ser203 to Alanine had no effect on sumoylation of SF-1 at Lys194, as well as phosphorylated receptor was still efficiently sumoylated in cells (data not shown)[20]. Additionally, no changes in SF-1 sumoylation were observed with Ser203Asp and Ser203Gln phosphomimic mutants or with in vitro phosphorylated wild-type SF-1 (Figure 3B, data not shown). At present, we cannot exclude that, under limiting Ubc9 conditions, phosphorylation at Ser203 might act to enhance sumoylation at Lys194. Instead, the results from the in vitro kinase assay suggest that SF-1 sumoylation at Lys194 may result in inhibition of phosphorylation at Ser203. Recombinant sumoylated SF-1 LBD protein showed reduced level of in vitro phosphorylation by MAP kinase as shown by decreased incorporation of $[^{32}P]$-labeled phosphate with sumoylated receptor compared to the unmodified SF-1 LBD, especially at later time points (Figure 3C).

Sumoylation at Lys119 inhibits DNA-binding at specific sites.

Sumoylation leads to diverse outcomes for the interaction between the transcription factor and the genome. In some cases, sumoylated transcription factors retain their ability to bind DNA, while sumoylation of other transcription factors inhibits their DNA-binding ability[44-46]. Although several nuclear hormone receptors contain sumoylation sites in close proximity to their DBDs, molecular effects of sumoylation on DNA-binding activity of these proteins have not been thoroughly investigated[18]. The mechanisms underlying SUMO-induced inhibition of DNA-binding are likely to be diverse, and may include steric hindrance where SUMO1 attachment directly blocks interaction with DNA[13,15]. Yet, sumoylation could also affect the DBD-DNA interaction by altering the specificity of DBD-DNA contacts. SF-1 is a monomeric nuclear hormone receptor that binds extended nuclear receptor half-site sequences 5' – YCAAGGYCR – 3' (where Y = T/C, R = G/A)[47]. It has been previously shown that SUMO-modified SF-1 retains the ability to bind DNA[27,28]. Using EMSA assay, we show that recombinant SF-1 DBD, in vitro sumoylated at Lys 119, binds to a double-stranded oligonucleotide, containing consensus SF-1 binding site from the mouse Mullerian Inhibiting Substance (mMIS) promoter (Figure 4B and C)[48]. We detected efficient binding of both unmodified and sumoylated SF-1 DBD to DNA, and the binding was completely abolished in the presence of a mutant mMIS binding site (Figure 4C). In the presence of Ulp1, sumoylation of SF-1 DBD was abolished, and the sumoylated protein-DNA complex was no longer detected on the gel (Figure 4B and C). The components of the in vitro sumoylation reaction could not bind DNA (data not shown).

94

Based on SF-1 target gene promoter analyses, a significant number of SF-1 binding sites are known to deviate from the consensus[30]. Similarly, in a recent genome-wide study, 30% of the genes bound by a monomeric nuclear hormone receptor, ERRα, deviated from the TCAAGGTCA consensus ERR site[49]. The structures of SF-1 DBD complexed with an atypical rat inhibinα binding site and LRH-1 bound to a consensus human CYP7A site reveal that the same structural elements in NR5A DBD are important for interaction with DNA[30,34]. To examine whether sumoylated protein could bind to an atypical SF-1 binding site, we used the rat inhibinα site compared to a consensus SF-1 site from the human CYP11A gene (human side-chain cleavage or hSCC) (Figure 4B). In the sequence of the rat inhibinα site, G:C pairing replaces the A:T pairing in the +1 position of the DNA, creating a low affinity site. The CYP11A SF-1 consensus site is identical to the CYP7A site used to determine LRH-1 DBD structure (Figure 4A)[34]. In the EMSA assay, we saw robust binding of the sumoylated SF-1 DBD to the CYP11A site whereas the binding of the sumoylated protein to the atypical rat inhibinα site was abolished (Figure 4C). This effect was still observed even with increasing amounts of sumoylated SF-1 as compared to the amount of protein used to detect binding to a consensus element (Figure 4D). Intriguingly, alignment of the human and rat inhibinα promoter shows that the G in the +1 position of the binding site is changed to a T in human inhibinα SF-1 binding site. This sequence is still atypical with respect to SF-1 binding; however, this single base change now allows sumoylated SF-1 to bind the inhibin site (Figure 4C). These data suggest that sumoylation may be specifically affecting the recognition of the 5' bases in the SF-1 binding site; possibly, through

affecting the interaction of the A-box loop with DNA or packing of the Ftz-F1 helix against the core of the DBD. Alternatively, sumoylation may simply diminish overall affinity of SF-1 for DNA because the signal from the protein-DNA complex with the human inhibinα site appears to be greater than that from the rat inhibinα site at the same sumoylated protein concentration (Figure 4C). Quantitative measurements of sumoylated SF-1 DBD binding to DNA will be necessary to answer this question.

DNA-binding inhibits SF-1 sumoylation at Lys119.

For nuclear hormone receptors, binding site recognition depends on numerous specific contacts between the DBD and the DNA and is accompanied by conformational changes in the DBD[50]. Since the DBD sumoylation site is in close proximity to SF-1 DBD, we wanted to ask whether DNA-bound protein could be sumoylated. We found that in vitro sumoylation of SF-1 DBD at Lys119 in the presence of SF-1 binding site from the mMIS promoter was significantly inhibited (Figure 5A). This effect was specific for SF-1-DNA interaction because a mutant binding site that does not interact with the protein had no effect on Lys119 sumoylation. This affect was also observed with a full-length SF-1 protein (Supplementary Figure 1). Sumoylation of SF-1 DBD in the presence of high concentration of single-stranded DNA was not affected suggesting that DNA had no non-specific inhibitory effect on SF-1 sumoylation (Figure 5A). Sumoylation of SF-1 DBD at Lys119 was inhibited by DNA in a dose-dependence suggesting sumoylation is inhibited as a direct result of SF-1 DNA-binding (Figure 5B). Residual sumoylation observed in the presence of high concentrations of DNA was probably due to a small fraction of misfolded or oxidized SF-1 DBD, unable to bind DNA.

A unique feature of the SF-1 DBD, Ftz-F1 helix, stabilizes SF-1 DNA-binding through non-covalent interactions with the core DBD and contacts with the DNA phosphate backbone[30,34]. Because Lys119 sumoylation site lies in close proximity to the Ftz-F1 helix, we hypothesized that Ftz-F1 helix interaction with DNA would act to hinder sumoylation at Lys119. Therefore, mutations reducing the interaction of the Ftz-F1 helix with DNA would enhance sumoylation at Lys119. From the structure of SF-1 DBD, Lys100 and Arg103 create a positively-charged surface in the Ftz-F1 helix that lies proximal to the phosphate backbone of the DNA, potentially coordinating the backbone through electrostatic interactions (Figure 6A)[30]. We mutated both Lys100 and Arg103 in the Ftz-F1 helix to aspartic acid (K100D, R103D mutant). The DNA-binding ability of this mutant was significantly impaired compared to wild type protein confirming the importance of the Ftz-F1 helix for SF-1 DNA-binding ability (Figure 6B). As predicted, in the presence of SF-1 consensus site, SF-1 DD mutant was sumoylated at Lys119 to a much higher extent than wild-type protein (Figure 6B). We conclude that SF-1 bound to DNA is resistant to sumoylation at Lys119 DBD site, providing a way to regulate SF-1 DBD sumoylation.

Sumoylated SF-1 human mutant does not bind DNA.

A single base-pair difference between rat inhibina and human inhibina modulates sumoylated SF-1 DNA-binding. Based on SF-1 and LRH-1 DBD structures, the G:C (+1) pairing in rat inihbina interacts with conserved RGGR sequence in the A-box loop region of mSF-1 DNA-binding domain. This region is conserved in other monomeric

nuclear hormone receptors and was shown to be important for binding site recognition[51,52]. Intriguingly, a homozygous loss-of-function human mutation in the A-box loop of SF-1 was recently reported where Arginine 92 is substituted by Glutamine (Arg92Gln)[53]. The Arg92Gln mutant SF-1 protein is still able to bind DNA in gel-shift assay and retains appreciable transcriptional activity in cell-based assays[53,54]. However, this mutation results in a complete loss of SF-1 function in vivo. We hypothesized that Arg92Gln mutation affects regulation of SF-1 DNA-binding by sumoylation at Lys119. The structure of LRH-1 DBD suggests structural basis for reduced interaction of Arg92Gln (Arg165 in LRH-1) mutant with DNA[34]. In wild-type protein, the side-chain of Arg165 donates hydrogen bonds to the bases in the minor groove of DNA (Figure 7A)[55]. In the Arg92Gln mutant, these coordinating interactions are reduced, thus impairing DNA-binding ability of the mutant protein. In the presence of DNA, SF-1 Arg92Gln mutant was sumoylated to a slightly greater extent than wild-type protein (Figure 7B, lower panel). This observation was predicted as a result of reduced interaction of Arg92Gln mutant with DNA. However, addition of a wild-type SF-1 binding site still efficiently inhibited sumoylation of Arg92Gln mutant at Lys119, demonstrating that regulation of Lys119 sumoylation by DNA-binding is intact in this mutant. Supporting this observation, SF-1 Arg92Gln protein bound to DNA well as demonstrated by EMSA assay (Figure 7C). Surprisingly, sumoylated SF-1 Arg92Gln mutant could no longer bind DNA at either consensus or atypical SF-1 binding sites (Figure 7C), suggesting that this mutation impairs the ability of the sumoylated protein to bind DNA. This result highlights the contribution of the A-box loop region of SF-1 DBD to the ability of the sumoylated protein to bind DNA.

Discussion

In this study, we report that repression of SF-1 function by sumoylation involves two distinct molecular mechanisms that target both the DBD and the LBD of SF-1. We show that sumoylation close to SF-1 LBD does not change the structure of the LBD, but instead interferes with MAPK phosphorylation of SF-1. Sumoylation of SF-1 at the DBD site modulates DNA-binding activity of the receptor in a sequence-dependent manner. Additionally, the DNA-bound SF-1 is refractory to sumoylation at the DBD site suggesting that DBD sumoylation does not affect the receptor while it is interacting with the genome. As such, we uncover a regulatory relationship between transcription factor sumoylation and DNA-binding ability.

Our experimental approach was focused on evaluating biochemical and structural properties of sumoylated SF-1 LBD and DBD. Current cellular and in vivo studies on the consequences of sumoylation are limited to loss-of-function approaches due to lack of technology for creating permanently-sumoylated substrates. Unlike with phosphorylation, where phospho-specific antibodies are available and phosphomimics may offer plausible gain-of-function substitutes, the mimicking approach is not feasible with a small protein like SUMO, and there are currently no specific antibodies that are able to specifically detect the sumoylated form of the protein. Direct terminal SUMO-fusions are informative when SUMO acts as a separate domain or as a molecular tag to direct proteins PML bodies[56]. Yet, such approach may fail to reveal the importance of specific structural changes in the substrate protein induced by conjugation of SUMO at its native site. Therefore, biochemical and structural analyses of purified sumoylated

99

proteins complement cellular loss-of-function studies and address the question of whether sumoylation directly affects structure and function of its substrates.

Conformational changes in nuclear receptor LBD through ligand-binding, coactivator, or corepressor recruitment is an established mechanism to regulate nuclear receptor activity. Yet, recent structural studies of the so-called orphan nuclear hormone receptors bring into question whether such mechanism applies to all nuclear receptors[17,31]. Although sumoylation of SF-1 potently represses its activity, we propose that the mechanism by which SUMO represses SF-1 does not involve changes in SF-1 LBD. Our nuclear magnetic resonance spectroscopy studies with sumoylated SF-1 LBD revealed that the structure of SUMO1 is not affected upon conjugation to SF-1, suggesting that SF-1 LBD and SUMO1 do not interact (Chaper 4, Figure 2A). The recruitment of full-length coregulators or corepressor Dax1 LxxLL peptide was also not affected by sumoylation of the LBD (Chapter 4, Figure 2 and L.A.C. unpublished results). These results confirm previous observations that Dax1 could potently repress both wild type and SUMO-mutant SF-1[20]. However, from our cellular assays it is evident that LBD sumoylation site functions in concert with the DBD sumoylation site as the double mutant exhibits greater activity than either of single SUMO-mutants. Therefore, the LBD site may employ additional mechanisms to repress SF-1 activity, and the function of the LBD site in the context of full-length receptor would be interesting to examine. At this time, we were not able to obtain large quantities of recombinant full-length SF-1 sufficient for structural studies in the *E.coli* expression system; therefore, the function of the LBD and the DBD sumoylation sites in the context of a full-length SF-1 could not addressed in this study.

Our study demonstrates that SF-1 sumoylation selectively regulates SF-1 DNA-binding ability at low affinity binding sites. Molecular mechanism underlying this selectivity is likely to involve functional interaction of SUMO1 with regions of SF-1 DBD that are crucial in recognition of low affinity SF-1 binding sites. For example, the selective nature of sumoylated SF-1-DNA interaction is abolished by Arg92Gln DBD mutation. This mutation resides in the A-box loop of SF-1 DBD and disrupts DBD-DNA contacts important for recognition of the 5' region of the SF-1 binding site[34]. Similarly, SUMO1 conjugation affects DNA recognition in thymine-DNA glycosylase (TGD) enzyme[11,12]. Sumoylation induces a conformational change in TGD that may enhance its dissociation from DNA, possibly, through interfering with the N-terminal domain of TGD that facilitates suboptimal substrate processing[12]. It would be interesting to determine whether sumoylation of SF-1 DBD also results in a conformational change that could explain the selectivity in sumoylated SF-1 binding between high and low affinity sites.

Regulation of DNA-binding by sumoylation may be a common feature for nuclear hormone receptors. DNA-binding of LRH-1 will be, most likely, regulated by DBD sumoylation, since DBD structure and the location of the DBD sumoylation site are highly conserved between LRH-1 and SF-1 in both mouse and human. Intriguingly, the founding member of the NR5A subfamily, Drosophila Ftz-F1 has a single sumoylation site located next to the DBD, and Ftz-F1 is sumoylated in vitro (M.D.S., unpublished data). Thus, regulation of NR5 subfamily DNA-binding by sumoylation may be an evolutionarily conserved mechanism. Moreover, several other nuclear hormone receptors, like ERR and PPAR, have sumoylation sites in close proximity to their DNA-

binding domains[21,24]. Studies of whether sequence-specific regulation of nuclear receptor

DNA-binding by sumoylation is a common regulatory mechanism for nuclear hormone

receptors would be of interest.

Many nuclear hormone receptors recognize identical nuclear receptor half-sites. Because

the expression profiles of nuclear hormone receptors also overlap, they could be

competing for regulation of the same target gene. For example, SF-1 and NGFI-B

(NR4A1, Nurr77) have overlapping binding site preferences with SF-1 being able to bind

to many of the NGFI-B binding sites[51]. Our EMSA studies with sumoylated SF-1 DBD

suggest that sumoylation prevents interaction of SF-1 with specific binding sites,

potentially allowing these sites to be bound and regulated by another nuclear hormone

receptor. Therefore, sumoylation may account for one mechanism to ensure specificity

of gene regulation by a particular nuclear hormone receptor during a certain

developmental stage or signaling event. Additionally, changes in SF-1 binding site

sequences among different species thus determine whether a particular SF-1 target gene

is regulated by SF-1 sumoylation in this specie. For example, SF-1 binding site in the

inhibinα promoter varies by one base pair in human and rat. While the rat inhibinα site

is not bound by sumoylated SF-1 in our study, at the human site, the interaction of

sumoylated SF-1 with DNA is restored. Thus, taking into account binding site mutations

between species, SF-1 sumoylation is likely to regulate overlapping but distinct sets of

SF-1 target genes in different organisms.

References

1. Kerscher, O., Felberbaum, R. & Hochstrasser, M. Modification of proteins by ubiquitin and ubiquitin-like proteins. *Annu Rev Cell Dev Biol* 22, 159-80 (2006).
2. Johnson, E.S. Protein modification by SUMO. *Annu Rev Biochem* 73, 355-82 (2004).
3. Capili, A.D. & Lima, C.D. Taking it step by step: mechanistic insights from structural studies of ubiquitin/ubiquitin-like protein modification pathways. *Curr Opin Struct Biol* (2007).
4. Sharrocks, A.D. PIAS proteins and transcriptional regulation--more than just SUMO E3 ligases? *Genes Dev* 20, 754-8 (2006).
5. Mukhopadhyay, D. & Dasso, M. Modification in reverse: the SUMO proteases. *Trends Biochem Sci* 32, 286-95 (2007).
6. Reverter, D. & Lima, C.D. Insights into E3 ligase activity revealed by a SUMO-RanGAP1-Ubc9-Nup358 complex. *Nature* 435, 687-92 (2005).
7. Reverter, D. & Lima, C.D. Structural basis for SENP2 protease interactions with SUMO precursors and conjugated substrates. *Nat Struct Mol Biol* 13, 1060-8 (2006).
8. Heun, P. SUMOrganization of the nucleus. *Curr Opin Cell Biol* 19, 350-5 (2007).
9. Kerscher, O. SUMO junction-what's your function? New insights through SUMO-interacting motifs. *EMBO Rep* 8, 550-5 (2007).
10. Macauley, M.S. et al. Beads-on-a-string, characterization of ETS-1 sumoylated within its flexible N-terminal sequence. *J Biol Chem* 281, 4164-72 (2006).
11. Steinacher, R. & Schar, P. Functionality of human thymine DNA glycosylase requires SUMO-regulated changes in protein conformation. *Curr Biol* 15, 616-23 (2005).
12. Baba, D. et al. Crystal structure of thymine DNA glycosylase conjugated to SUMO-1. *Nature* 435, 979-82 (2005).
13. Anckar, J. et al. Inhibition of DNA binding by differential sumoylation of heat shock factors. *Mol Cell Biol* 26, 955-64 (2006).
14. Tsuruzoe, S. et al. Inhibition of DNA binding of Sox2 by the SUMO conjugation. *Biochem Biophys Res Commun* 351, 920-6 (2006).
15. Chou, C.C. et al. Small ubiquitin-like modifier modification regulates the DNA binding activity of glial cell missing Drosophila homolog a. *J Biol Chem* 282, 27239-49 (2007).
16. Boyer-Guittaut, M. et al. SUMO-1 modification of human transcription factor (TF) IID complex subunits: inhibition of TFIID promoter-binding activity through SUMO-1 modification of hsTAF5. *J Biol Chem* 280, 9937-45 (2005).
17. Ingraham, H.A. & Redinbo, M.R. Orphan nuclear receptors adopted by crystallography. *Curr Opin Struct Biol* 15, 708-15 (2005).
18. Faus, H. & Haendler, B. Post-translational modifications of steroid receptors. *Biomed Pharmacother* 60, 520-8 (2006).

19.   Ghisletti, S. et al. Parallel SUMOylation-dependent pathways mediate gene- and signal-specific transrepression by LXRs and PPARgamma. *Mol Cell* 25, 57-70 (2007).

20.   Lee, M.B. et al. The DEAD-box protein DP103 (Ddx20 or Gemin-3) represses orphan nuclear receptor activity via SUMO modification. *Mol Cell Biol* 25, 1879-90 (2005).

21.   Pascual, G. et al. A SUMOylation-dependent pathway mediates transrepression of inflammatory response genes by PPAR-gamma. *Nature* 437, 759-63 (2005).

22.   Yang, X.J. & Gregoire, S. A recurrent phospho-sumoyl switch in transcriptional repression and beyond. *Mol Cell* 23, 779-86 (2006).

23.   Yamashita, D. et al. The transactivating function of peroxisome proliferator-activated receptor gamma is negatively regulated by SUMO conjugation in the amino-terminal domain. *Genes Cells* 9, 1017-29 (2004).

24.   Vu, E.H., Kraus, R.J. & Mertz, J.E. Phosphorylation-dependent sumoylation of estrogen-related receptor alpha1. *Biochemistry* 46, 9795-804 (2007).

25.   Luo, X., Ikeda, Y. & Parker, K.L. A cell-specific nuclear receptor is essential for adrenal and gonadal development and sexual differentiation. *Cell* 77, 481-90 (1994).

26.   Parker, K.L. & Schimmer, B.P. Steroidogenic factor 1: a key determinant of endocrine development and function. *Endocr Rev* 18, 361-77 (1997).

27.   Komatsu, T. et al. Small ubiquitin-like modifier 1 (SUMO-1) modification of the synergy control motif of Ad4 binding protein/steroidogenic factor 1 (Ad4BP/SF-1) regulates synergistic transcription between Ad4BP/SF-1 and Sox9. *Mol Endocrinol* 18, 2451-62 (2004).

28.   Chen, W.Y., Lee, W.C., Hsu, N.C., Huang, F. & Chung, B.C. SUMO modification of repression domains modulates function of nuclear receptor 5A1 (steroidogenic factor-1). *J Biol Chem* 279, 38730-5 (2004).

29.   Hammer, G.D. et al. Phosphorylation of the nuclear receptor SF-1 modulates cofactor recruitment: integration of hormone signaling in reproduction and stress. *Mol Cell* 3, 521-6 (1999).

30.   Little, T.H. et al. Sequence-specific deoxyribonucleic acid (DNA) recognition by steroidogenic factor 1: a helix at the carboxy terminus of the DNA binding domain is necessary for complex stability. *Mol Endocrinol* 20, 831-43 (2006).

31.   Krylova, I.N. et al. Structural analyses reveal phosphatidyl inositols as ligands for the NR5 orphan receptors SF-1 and LRH-1. *Cell* 120, 343-55 (2005).

32.   Wang, W. et al. The crystal structures of human steroidogenic factor-1 and liver receptor homologue-1. *Proc Natl Acad Sci U S A* 102, 7505-10 (2005).

33.   Li, Y. et al. Crystallographic identification and functional characterization of phospholipids as ligands for the orphan nuclear receptor steroidogenic factor-1. *Mol Cell* 17, 491-502 (2005).

34.   Solomon, I.H. et al. Crystal structure of the human LRH-1 DBD-DNA complex reveals Ftz-F1 domain positioning is required for receptor activity. *J Mol Biol* 354, 1091-102 (2005).

35. Desclozeaux, M., Krylova, I.N., Horn, F., Fletterick, R.J. & Ingraham, H.A. Phosphorylation and intramolecular stabilization of the ligand binding domain in the nuclear receptor steroidogenic factor 1. *Mol Cell Biol* 22, 7193-203 (2002).

36. Yunus, A.A. & Lima, C.D. Purification and activity assays for Ubc9, the ubiquitin-conjugating enzyme for the small ubiquitin-like modifier SUMO. *Methods Enzymol* 398, 74-87 (2005).

37. Delaglio, F et al. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR.* 6, 277-293 (1995).

38. Goddard, T.D., and Kneeler, D. G. Sparky 3. *University of California, San Francisco* (1999).

39. Bayer, P. et al. Structure determination of the small ubiquitin-related modifier SUMO-1. *J Mol Biol* 280, 275-86 (1998).

40. Jin, C., Shiyanova, T., Shen, Z. & Liao, X. Heteronuclear nuclear magnetic resonance assignments, structure and dynamics of SUMO-1, a human ubiquitin-like protein. *Int J Biol Macromol* 28, 227-34 (2001).

41. Hietakangas, V. et al. PDSM, a motif for phosphorylation-dependent SUMO modification. *Proc Natl Acad Sci U S A* 103, 45-50 (2006).

42. Lee, Y.K., Choi, Y.H., Chua, S., Park, Y.J. & Moore, D.D. Phosphorylation of the hinge domain of the nuclear hormone receptor LRH-1 stimulates transactivation. *J Biol Chem* 281, 7850-5 (2006).

43. Yang, S.H., Galanis, A., Witty, J. & Sharrocks, A.D. An extended consensus motif enhances the specificity of substrate modification by SUMO. *Embo J* 25, 5083-93 (2006).

44. Dobreva, G., Dambacher, J. & Grosschedl, R. SUMO modification of a novel MAR-binding protein, SATB2, modulates immunoglobulin mu gene expression. *Genes Dev* 17, 3048-61 (2003).

45. Perdomo, J., Verger, A., Turner, J. & Crossley, M. Role for SUMO modification in facilitating transcriptional repression by BKLF. *Mol Cell Biol* 25, 1549-59 (2005).

46. Callewaert, L., Verrijdt, G., Haelens, A. & Claessens, F. Differential effect of small ubiquitin-like modifier (SUMO)-ylation of the androgen receptor in the control of cooperativity on selective versus canonical response elements. *Mol Endocrinol* 18, 1438-49 (2004).

47. Ueda, H. & Hirose, S. Defining the sequence recognized with BmFTZ-F1, a sequence specific DNA binding factor in the silkworm, Bombyx mori, as revealed by direct sequencing of bound oligonucleotides and gel mobility shift competition analysis. *Nucleic Acids Res* 19, 3689-93 (1991).

48. Shen, W.H., Moore, C.C., Ikeda, Y., Parker, K.L. & Ingraham, H.A. Nuclear receptor steroidogenic factor 1 regulates the mullerian inhibiting substance gene: a link to the sex determination cascade. *Cell* 77, 651-61 (1994).

49. Dufour, C.R. et al. Genome-wide orchestration of cardiac functions by the orphan nuclear receptors ERRalpha and gamma. *Cell Metab* 5, 345-56 (2007).

50. Holmbeck, S.M., Dyson, H.J. & Wright, P.E. DNA-induced conformational changes are the basis for cooperative dimerization by the DNA binding domain of the retinoid X receptor. *J Mol Biol* 284, 533-9 (1998).

51.   Wilson, T.E., Fahrner, T.J. & Milbrandt, J. The orphan receptors NGFI-B and steroidogenic factor 1 establish monomer binding as a third paradigm of nuclear receptor-DNA interaction. *Mol Cell Biol* 13, 5794-804 (1993).
52.   Gearhart, M.D., Holmbeck, S.M., Evans, R.M., Dyson, H.J. & Wright, P.E. Monomeric complex of human orphan estrogen related receptor-2 with DNA: a pseudo-dimer interface mediates extended half-site recognition. *J Mol Biol* 327, 819-32 (2003).
53.   Achermann, J.C. et al. Gonadal determination and adrenal development are regulated by the orphan nuclear receptor steroidogenic factor-1, in a dose-dependent manner. *J Clin Endocrinol Metab* 87, 1829-33 (2002).
54.   Ito, M., Achermann, J.C. & Jameson, J.L. A naturally occurring steroidogenic factor-1 mutation exhibits differential binding and activation of target genes. *J Biol Chem* 275, 31708-14 (2000).
55.   DeLano, W.L. MacPyMOL: A PyMOL-based Molecular Graphics Application for MacOS X. *DeLano Scientific LLC, Palo Alto, CA, USA*. (2007).
56.   Shiio, Y. & Eisenman, R.N. Histone sumoylation is associated with transcriptional repression. *Proc Natl Acad Sci U S A* 100, 13225-30 (2003).
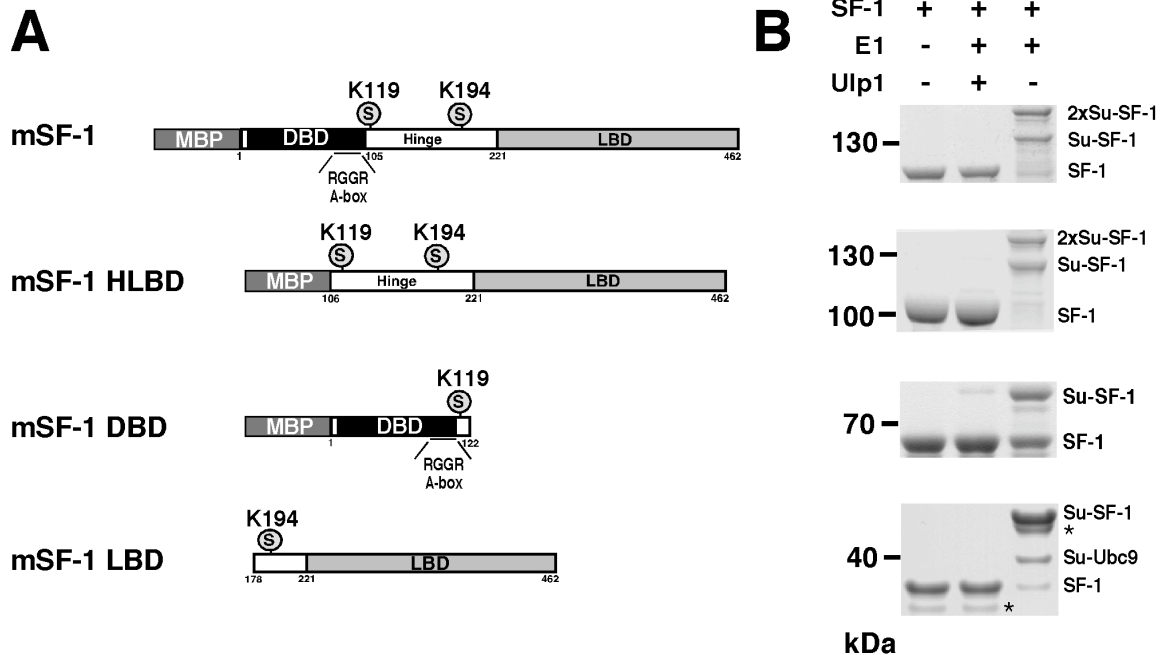
Acknowledgments

Figure 1. *In vitro* **sumoylation of recombinant SF-1 protein.** **(A)** A schematic of mouse SF-1 protein constructs is shown with SUMO sites at Lys119 (K119, S), Lys194 (K194, S) and the A-box sequence indicated. Amino acid numbers are given for the beginning and the end of each construct. DBD, DNA-binding domain; LBD, ligand-binding domain; HLBD, Hinge and ligand-binding domain; MBP, maltose-binding protein. **(B)** Protein gels of unmodified SF-1 (-E1), sumoylated SF-1 treated with Ulp1 enzyme (+E1, +Ulp1), or sumoylated SF-1 (+E1) proteins (10µL in vitro reaction) are shown for the SF-1 constructs depicted in **A**. Protein was detected by Coomassie staining. The * denotes wild type SF-1 degradation product. **SF-1**, unmodified SF-1; **Su-SF-1**, sumoylated SF-1; **2XSu-SF-1**, SF-1 with two SUMO1 attached; **Su-Ubc9**, sumoylated Ubc9.
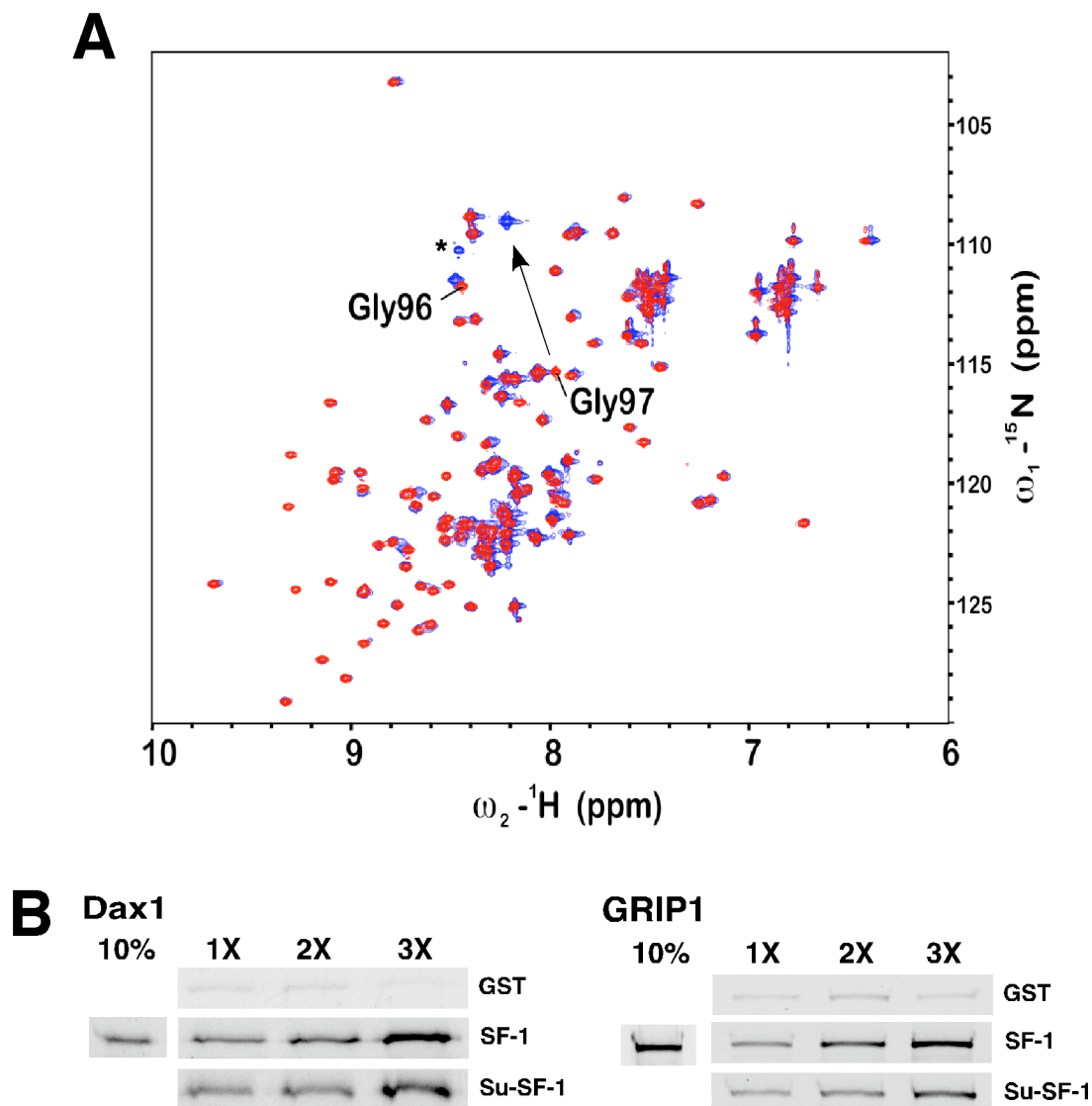
Figure 2. Lys194 sumoylation does not change SF-1 LBD structure. **(A)** NMR analysis revealed no changes in the structure of SUMO1 after conjugation to SF-1, except in the residue involved in the isopeptide bond (Gly97) and the adjacent residue (Gly96). An overlay of ^{1}H-^{15}N HSQC spectra from ^{15}N-SUMO1 (red) and ^{15}N-SUMO1-SF-1 LBD (blue) are shown (800 MHz ^{1}H frequency). Gly97 and Gly96 are labeled. The arrow indicates the chemical shift of Gly97. The * marks a resonance that was visible in ^{15}N-SUMO1 spectrum at a lower contour and, most likely, represents a residue in the His$_6$ tag on SUMO1. **(B)** GST pulldown assays of full-length [^{35}S]-Dax1 and [^{35}S]-GRIP1 bound to increasing amounts of unmodified GST-SF-1 LBD (SF-1) or sumoylated GST-SF-1 LBD (Su-SF-1). 1X, 2X, 3X indicate relative amounts of GST proteins used. 10% input and GST control are indicated.
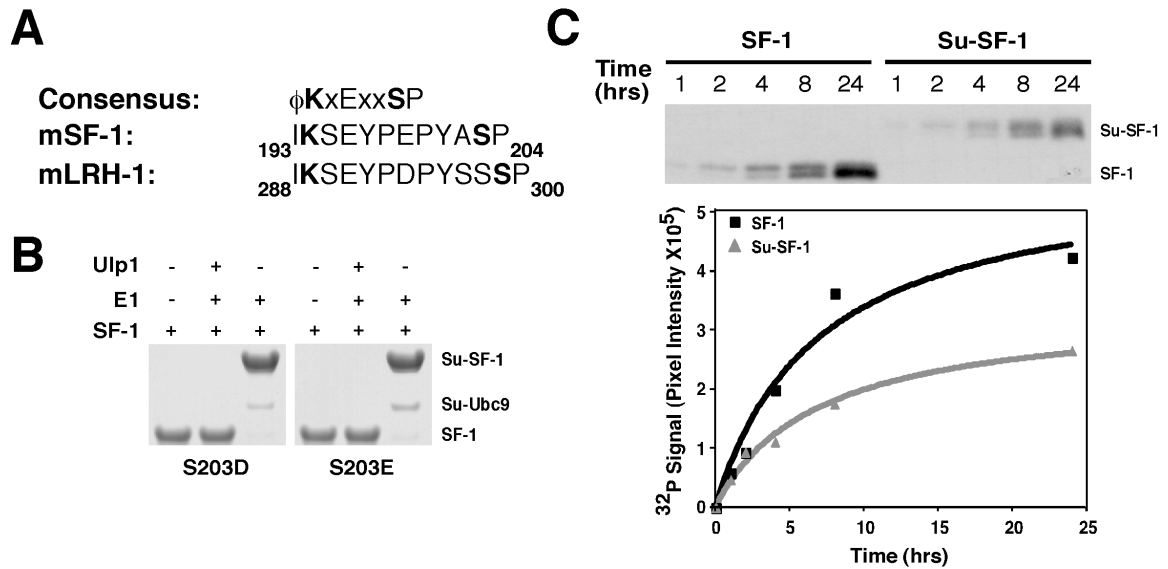
Figure 3. SF-1 sumoylation at Lys194 inhibits *in vitro* SF-1 phorphorylation. (A) Sequences from putative phospho-sumoyl switch sites in mouse SF-1 and mouse LRH-1 are shown as compared to a reported consensus sequence[22]. Amino acid numbers are given for the beginning and the end of each sequence. Sumoylated lysine (K) and phosphorylated serine (S) are shown in bold. **(B)** In vitro sumoylation assays were performed with equimolar quantities (9μM) of purified aspartic acid phosphomimic (S203D) and glutamic acid phosphomimic (S203E) SF-1 LBD protein. Protein gels of reactions containing unmodified SF-1 (-E1, -Ulp1), sumoylated SF-1 treated with Ulp1 enzyme (+E1, +Ulp1), or sumoylated SF-1 (+E1, -Ulp1) proteins are shown. Protein signal was detected by Coomassie staining. **(C)** Equimolar quantities (22μM) of unmodified or sumoylated SF-1 LBD protein was phosphorylated by Erk2 in the presence of radiolabeled [^{32}P]-γ-ATP for the time indicated. Quantified data are represented as a graph below. **SF-1**, unmodified SF-1; **Su-SF-1**, sumoylated SF-1; **Su-Ubc9**, sumoylated Ubc9.
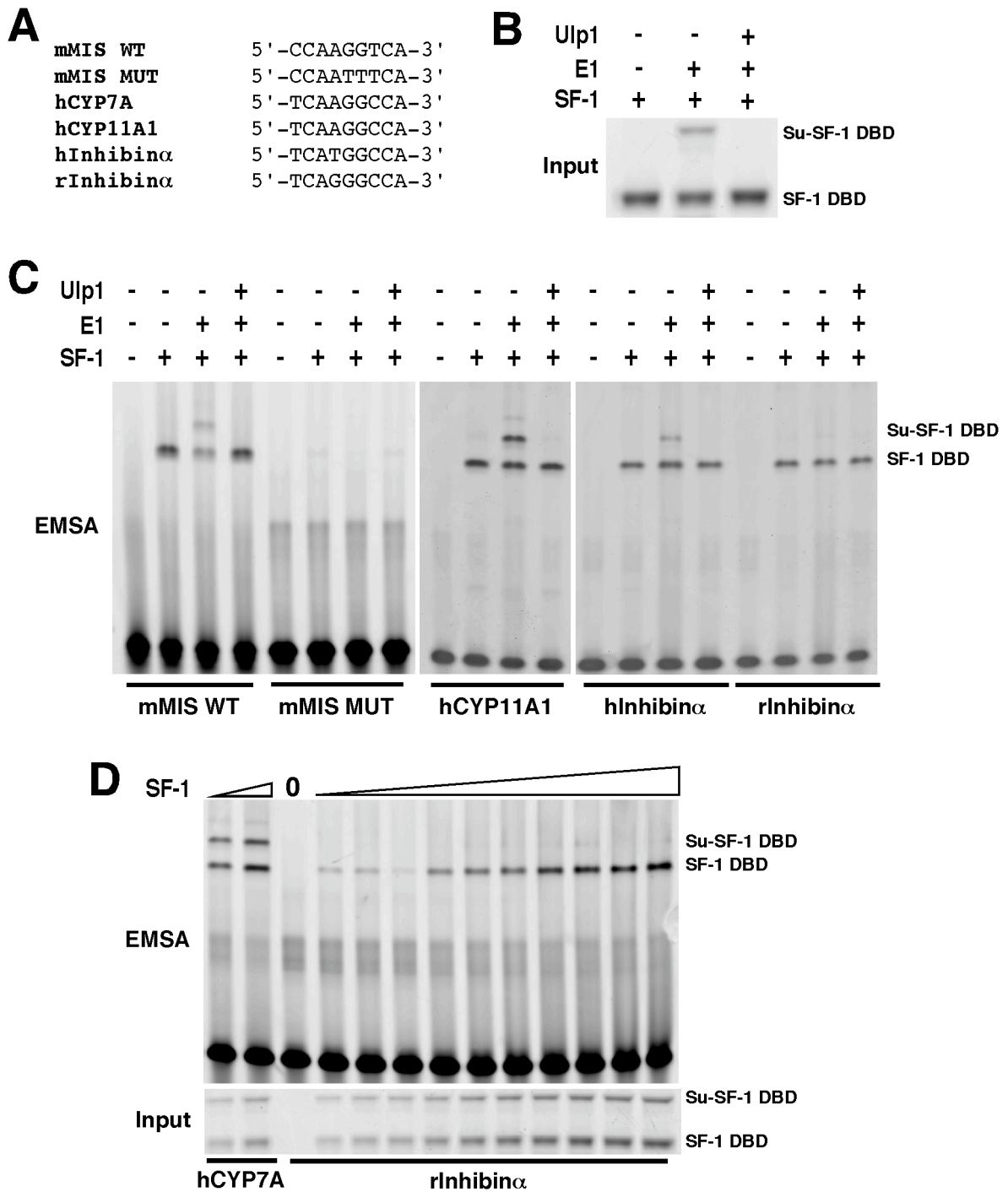
Figure 4. Sumoylated SF-1 DBD selectively binds DNA. **(A)** SF-1 binding sites from SF-1-responsive promoters used in EMSA are shown. Complete oligonucleotide sequences are shown in Table 1. m, mouse; h, human; r, rat; MIS, Mullerian Inhibiting Substance; CYP11A1, cytochrome P450 11A1; CYP7A, cytochrome P450 7A. **(B)** Protein gel of EMSA input (10μL) containing unmodified SF-1 DBD (-E1, -Ulp1),

sumoylated SF-1 DBD (+E1, -Ulp1), or sumoylated SF-1 DBD treated with Ulp1 (+E1, +Ulp1) proteins is shown. Protein signal was detected by SYPRO Red staining. **(C)** EMSA assays using unmodified SF-1 DBD (-E1), sumoylated SF-1 DBD (+E1), or sumoylated SF-1 DBD treated with Ulp1 (+E1, +Ulp1) proteins (2μL in vitro reaction) and double-stranded oligonucleotides (1μM) from SF-1 target gene promoters are shown. Oligonucleotides were derived from SF-1 target genes indicated below the gel. DNA signal was detected by Sybr Green staining. **(D)** EMSA illustrating binding of varying concentrations (small triangle 1.5μL, 2.5μL in vitro reaction; large triangle: 1μL-5.5μL in vitro reaction, in 0.5μL increments) sumoylated SF-1 DBD protein to double-stranded oligonucleotides (1μM) from human CYP7A and rat inhibinα promoters is shown. Protein gel of EMSA input (Input, 10μL) containing unmodified and sumoylated SF-1 DBD protein is shown. For **C** and **D**, protein-DNA complexes (10μL) were resolved on a native polyacrylamide gel (EMSA). 0 denotes no protein control. **SF-1 DBD**, unmodified SF-1 DBD; **Su-SF-1 DBD**, sumoylated SF-1 DBD.
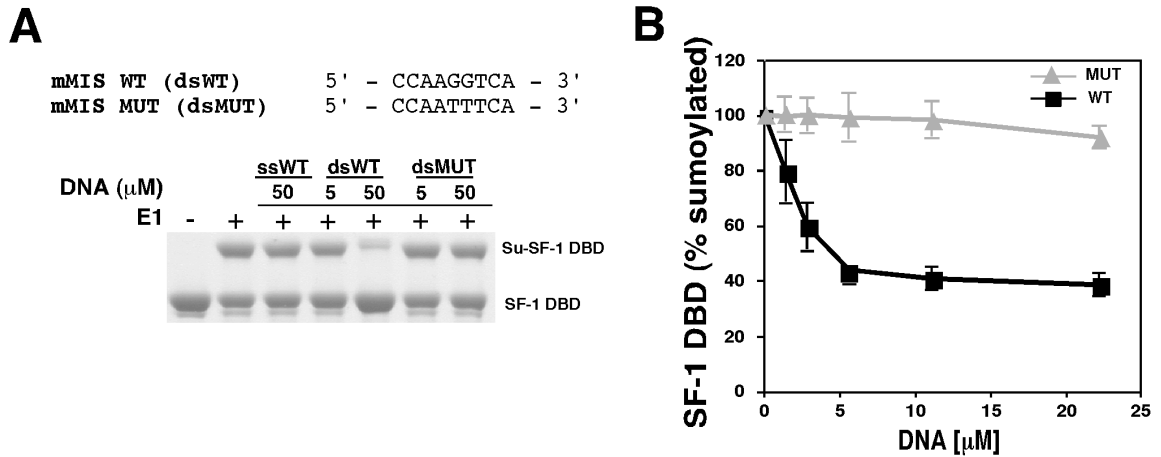
A

			ssWT	dsWT		dsMUT	

mMIS WT (dsWT)   5' – CCAAGGTCA – 3'
mMIS MUT (dsMUT)   5' – CCAATTTCA – 3'

DNA (μM)   50   5   50   5   50
E1   -   +   +   +   +   +   +

Su-SF-1 DBD
SF-1 DBD

B

Figure 5. DNA-binding inhibits SF-1 DBD sumoylation. **(A)** (Top panel) sequences of wild type (dsWT) and mutant (dsMUT) SF-1 binding sites from the mouse Mullerian Inhibiting Substance (mMIS) promoter are shown. Complete oligonucleotide sequences are listed in Supplemental Table 1. (Bottom panel) Sumoylation of SF-1 DBD in the presence of DNA. A Coomassie-stained protein gel of in vitro sumoylation reactions containing SF-1 DBD (5μM) and indicated concentrations of wild type single-stranded (ssWT), wild type double-stranded (dsWT), or mutant double-stranded (dsMUT) SF-1 binding sites is shown. **(B)** The graph shows quantified SF-1 DBD (2.8μM) sumoylation in the presence of increasing concentration (0, 1.4μM, 2.8μM, 5.6μM, 11.1μM, 22.2μM) of wild type (mMIS WT) or mutant (mMIS MUT) SF-1 binding sites from three independent experiments; error bars show s.e.m. Signal was detected by SYPRO Red staining. **SF-1 DBD**, unmodified SF-1 DBD; **Su-SF-1 DBD**, sumoylated SF-1 DBD.
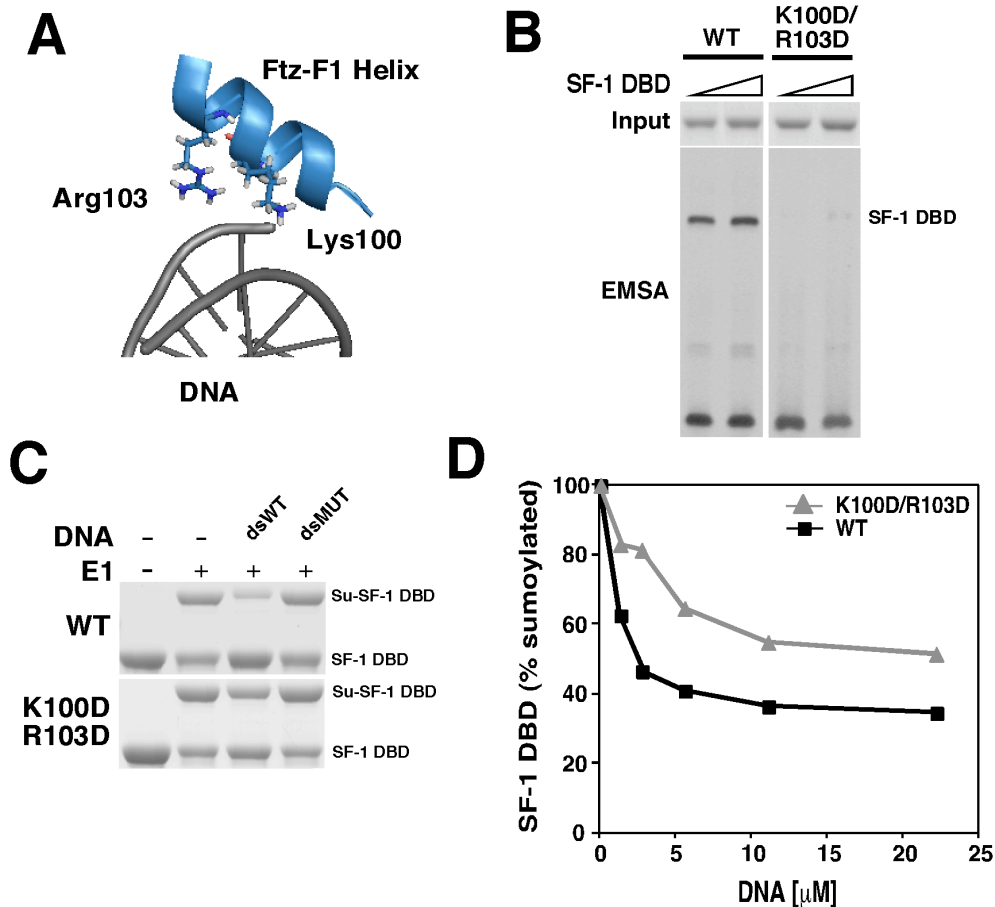
Figure 6. DNA-binding is necessary to inhibit SF-1 Lys119 sumoylation. (A) A ribbon diagram of Lys100 and Arg103 (sticks) in the Ftz-F1 helix of SF-1 in relation to DNA (grey sticks). The diagram was made using mouse SF-1 DBD structure (PDB 2FF0)[30,55]. **(B)** EMSA showing binding of wild type SF-1 DBD (WT, 4μL, 5μL of in vitro reaction) or mutant SF-1 DBD (K100D/R103D, 4μL, 5μL of in vitro reaction) protein to double-stranded oligonucleotide (1μM) from mouse MIS promoter is shown. Protein gel of EMSA input (Input, 10μL) containing wild type or K100D/R103D mutant SF-1 DBD is shown. Protein signal was detected by SYPRO Red staining. Protein-DNA complexes (10μL) were resolved on a native polyacrylamide gel (EMSA). DNA signal was detected by Sybr Green staining. **(C)** Coomassie-stained protein gels of sumoylation reaction containing wild type or K100D/R103D mutant SF-1 DBD (2.8μM, +E1) and wild type double stranded (dsWT, 28μM) or mutant double stranded (dsMUT, 28μM) mouse MIS SF-1 binding sites are shown. No sumoylation control (-DNA, -E1) and a control reaction without DNA are included (-DNA, +E1). **(D)** The graph shows quantified sumoylation of wild type or K100D/R103D mutant SF-1 DBD (2.8μM) in the presence of increasing concentration (0, 1.4μM, 2.8μM, 5.6μM, 11.1μM, 22.2μM) of wild type mMIS SF-1 binding site. Protein signal was detected by SYPRO Red staining. **SF-1 DBD**, unmodified SF-1 DBD; **Su-SF-1 DBD**, sumoylated SF-1 DBD.
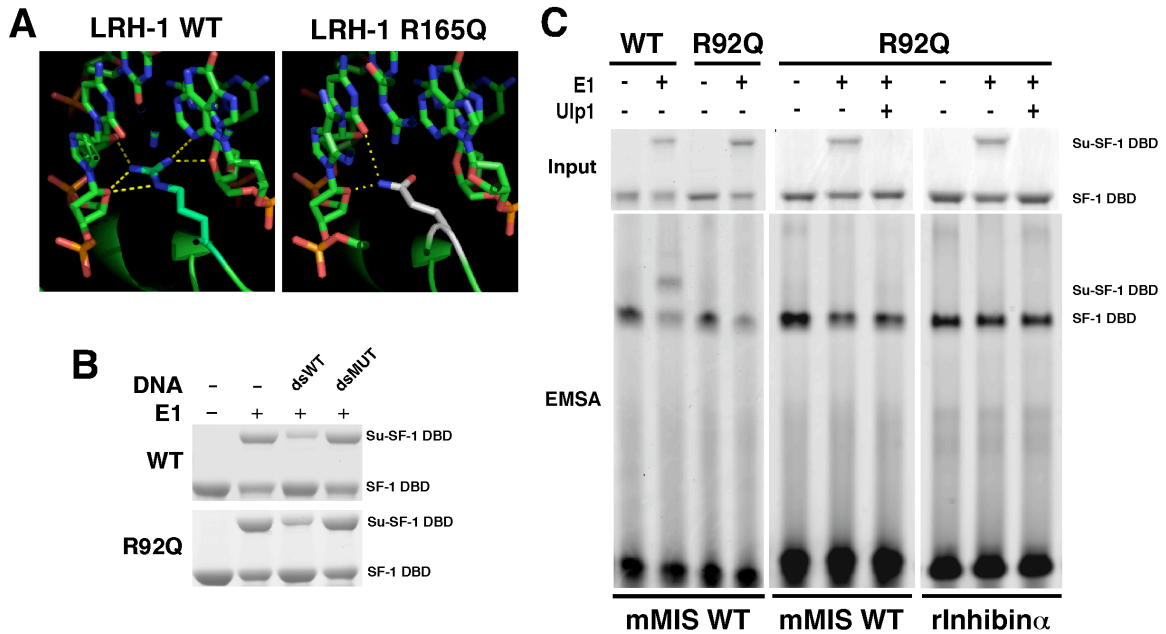
113

Figure 7. Sumoylated R92Q SF-1 human mutant does not bind DNA. (A) A model depicting possible amino acid-DNA contacts for wild type and R165Q mutant LRH-1 proteins. The model was created using human LRH-1 DBD structure and the mutagenesis function in PyMol (PDB 2A66)[34,55]. **(B)** Coomassie-stained protein gels of sumoylation reaction containing wild type or R92Q mutant SF-1 DBD (2.8μM, +E1) and wild type double stranded (dsWT, 28μM) or mutant double stranded (dsMUT, 28μM) mMIS SF-1 binding sites are shown. No sumoylation control (-DNA, -E1) and a control reaction without DNA are included (-DNA, +E1). **(C)** EMSA using unmodified wild type or R92Q mutant SF-1 DBD (-E1), sumoylated wild type or R92Q mutant SF-1 DBD (+E1), or sumoylated R92Q mutant SF-1 DBD treated with Ulp1 (+E1, +Ulp1) proteins (2μL in vitro reaction) and double-stranded oligonucleotides (1μM) from SF-1 target gene promoters are shown. Protein-DNA complexes (10μL) were resolved on a native polyacrylamide gel (EMSA). Oligonucleotides were derived from SF-1 target genes indicated below the gel. DNA was detected by Sybr Green staining. Protein gels of EMSA assay input (Input, 10μL) containing wild type or R92Q mutant SF-1 DBD protein, either unmodified or as a mixture of unmodified and sumoylated proteins, are shown. Protein signal was detected by SYPRO Red staining. **SF-1 DBD**, unmodified SF-1 DBD; **Su-SF-1 DBD**, sumoylated SF-1 DBD.
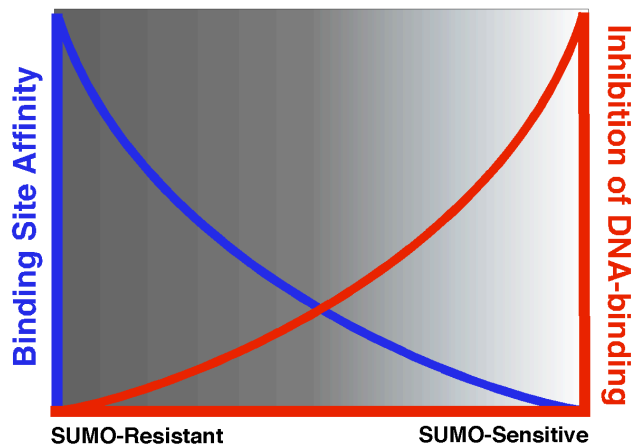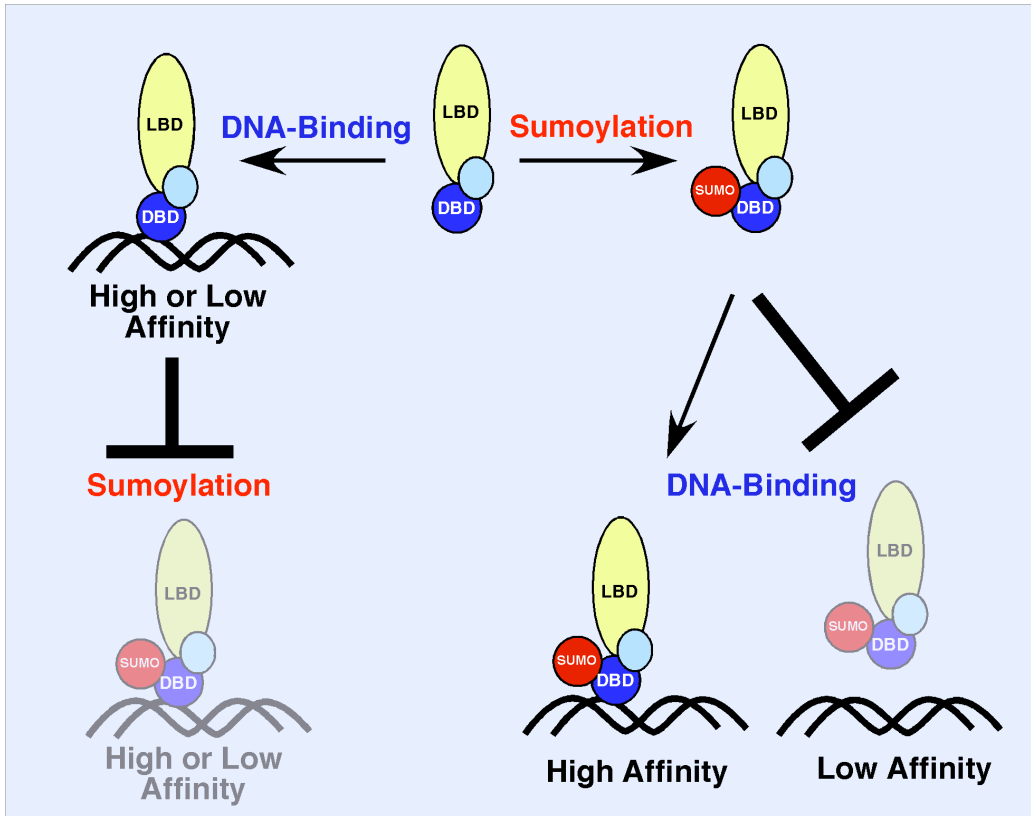
Figure 8. Model: the interplay between SF-1 DBD sumoylation and DNA-binding.
SF-1 exists in two states: DNA-bound at high and low affinity sites and unbound. DNA-bound SF-1 is resistant to sumoylation at DBD site. Only unbound SF-1 can be modified by SUMO1 at the DBD site. Once sumoylated, SF-1 is unable to bind low affinity sites. There is an inverse relationship between SF-1 binding site affinity and the inhibition of DNA-binding by sumoylation. Low affinity sites will be specifically regulated by SF-1 sumoylation, while the interaction of SF-1 with high affinity sites will not be significantly affected.

115
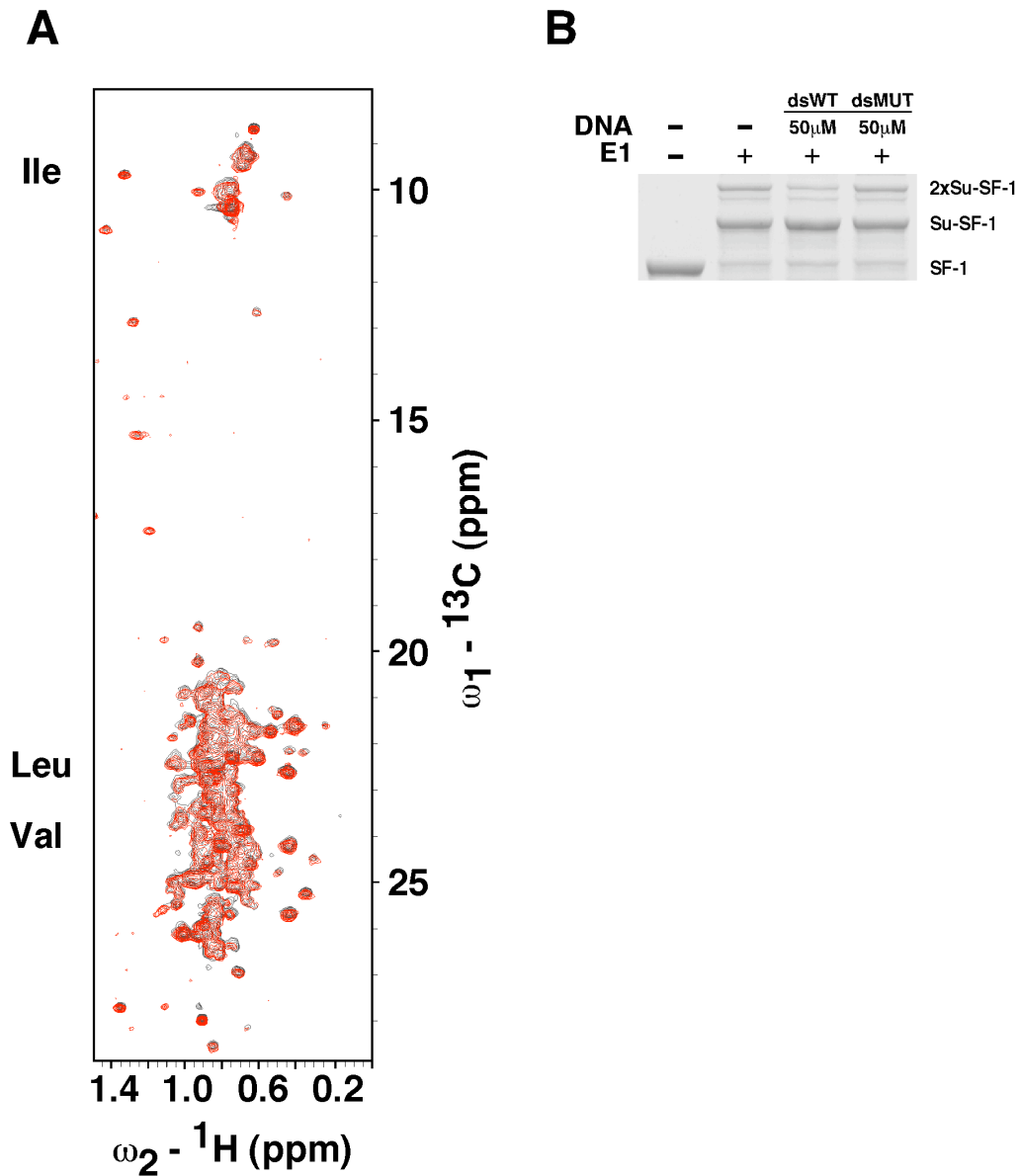
Supplementary Figure 1. **(A)** NMR analysis revealed no changes in the structure of SF-1 LBD after conjugation to SUMO1. An overlay of ^{1}H-^{13}C HSQC spectra from ^{13}C-ILV-methyl-SF-1 LBD (grey) and SUMO1-^{13}C- ILV-methyl-SF-1 LBD (red) are shown (800 MHz ^{1}H frequency). Regions of the spectrum for leucine, valine (Leu, Val) and isoleucine (Ile) resonances are indicated. **(B)** Sumoylation of full length SF-1 in the presence of SF-1 DNA binding site from mMIS promoter. A Coomassie-stained protein gel of in vitro sumoylation reactions containing full length SF-1 (2.5μM) and indicated concentrations of wild type double-stranded (dsWT) or mutant double-stranded (dsMUT) SF-1 binding sites is shown. No sumoylation control (-DNA, -E1) and a control reaction without DNA are included (-DNA, +E1). **SF-1 DBD**, unmodified SF-1 DBD; **Su-SF-1 DBD**, sumoylated SF-1 DBD; **2XSu-SF-1**, SF-1 with two SUMO1 attached.

Assay	Gene Name	Sequence
Sumoylation		
	mMIS wt sense	5'- TCCCC**AAGGTCACC**-3'
	mMIS wt antisense	5'- **GGTGACCTT**GGGGA-3'
	mMIS mut sense	5'- TCCCC**AATTTCACC**-3'
	mMIS mut antisense	5'- **GGTGAAATT**GGGGA-3'
EMSA		
	hCYP7A sense	5'- AACCACTTCTGATACCTGTGGACTTAGTT**CAAGGCC**AGTTACTACCACTTT-3'
	hCYP7A antisense	5'- AAAGTGGTAGTAACT**GGCCTT**GAACTAAGTCCACAGGTATCAGAAGTGGTT-3'
	mMIS wt sense	5'- *gatcc*GCCAGGCACTGTCCCCC**AAGGTC**ACCTTTGGTGTTGATA*agatccc*-3'
	mMIS wt antisense	5'- *tcgagggatcf*TATCAACACCAAAGGT**GACCTT**GGGGACAGTGCCTGGC*g*-3'
	mMIS0 mut sense	5'- *gatcc*GCCAGGCACTGTCCCC**AATTTC**ACCTTTGGTGTTGATA*agatccc*-3'
	mMIS0 mut antisense	5'- *tcgagggatcf*TATCAACACCAAAGGT**GAAATT**GGGGACAGTGCCTGGC*g*-3'
	hInhibinα sense	5'- GGCGGGAGTGGGAGATAAGGC**TCATGGCC**ACAGACATCTGCGTCAGAGATA-3'
	hInhibinα antisense	5'- TATCTCTGACGCAGATGTCTGT**GGCCATGA**GCCTTATCTCCCACTCCCGCC-3'
	rInhibinα sense	5'- CATTCTTGGCGGGAGTGGGAGATAAGGC**TCAGGGCC**ACAGACATCTGCGTC-3'
	rInhibinα antisense	5'- GACGCAGATGTCTGT**GGCCCTGA**TATCTCCCACTCCCGCCAAGAATG-3'
	hCYP11A1 sense	5'- ACATTTTATCAGCTTCTGGTAT**GGCCCTTGA**GCTGGTAGTTATAATCTTGGC-3'
	hCYP11A1 antisense	5'- GCCAAGATTATAAACTACCAGCT**CAAGGCC**ATACCAGAAGCTGATAAAATGT-3'

Supplementary Table 1. SF-1 Binding Sites Sequences of oligonucleotides from promoters of SF-1 target genes used in sumoylation and EMSA assays are shown. SF-1 binding site sequence is highlighted in bold. Partial restriction sites incorporated for cloning purposes are shown in italics. m, mouse; h, human; r, rat. MIS, Mullerian Inhibiting Substance; CYP7A1, cytochrome P450 7A1; CYP11A1, cytochrome P450 11A1.

117

CHAPTER V

Discussion and Future Directions

Discussion

Recent studies demonstrate that the majority of nuclear hormone receptors are modified by SUMO1 with profound effects on their transcriptional activity [1]. To date, sumoylation is the most potent mechanism to control the activity of Subfamily V nuclear receptors, as it appears that SF-1 and LRH-1 are not subject to switch-like regulation by ligand as seen with the steroid receptors.

Haploinsufficiency and modified protein populations

Studies in this thesis suggest that post-translational modifications such as sumoylation can produce protein pools that are functionally distinct. For example, SF-1 sumoylated at Lys194 is resistant to phosphorylation and activation by MAPK, while SF-1 sumoylated at Lys119 is resistant to DNA-binding at low affinity sites. This implies that, in the cell, transcriptionally-active SF-1 protein represents only a fraction of total SF-1. Haploinsufficiency is a phenomenon observed with several transcription factors where loss of one wild type allele manifests in complete loss-of-function phenotype [2]. This observation implies that gene regulation is sensitive to transcription factor concentration and that cells regulate the amount of active transcription factor. Similarly, haploinsufficiency of human SF-1 mutations suggests that the dosage of active SF-1 protein is critical in human physiology [3, 4]. According to my data, sumoylation renders active SF-1 protein unable to bind DNA at low affinity sites, thus altering active SF-1 concentration that these sites encounter. Human SF-1 mutant, Arg92Gln, provides additional support for this hypothesis. Sumoylated Arg92Gln mutant no longer binds

DNA, most likely, failing to activate a critical SF-1 target gene and resulting in complete loss-of-function phenotype. Thus, sumoylation establishes pools of SF-1 with different DNA-binding ability and provides a mechanism to regulate transcriptionally-active SF-1 protein concentration.

A large number of posttranslational protein modifications have been identified to date, including multiple ubiquitin-like proteins, methylation, acetylation, and arginylation. Therefore, differentially modified pools of protein must exist, and posttranslational modifications will have either competing or synergistic effects on specific protein function such as DNA-binding. For example, the relationship between acetylation and sumoylation has not been addressed in my studies. SF-1 is acetylated by the acetyltransferase p300 on multiple lysines in the DBD, and acetylation enhances SF-1 DNA-binding ability [5]. It would be interesting to determine whether acetylation could reverse the loss of DNA-binding due to SF-1 sumoylation at low affinity sites.

Mechanism of SF-1 Repression by SUMO1

Two prominent mechanisms of transcription factor repression by sumoylation involve direct recruitment of a repressor such as HDACs or relocalization of the sumoylated substrates into nuclear subdomains such as PML bodies. Both of these repression mechanisms apply to sumoylated SF-1, but they are mediated by a set of novel effectors. Sumoylation at Lys194 recruits a novel SF-1 repressor DEAD-box helicase DP103. I showed that SUMO-mediated repression of SF-1 does not involve direct interaction with Class I/II HDAC, although DP103 has the ability to directly recruit repressors mSin3A,

NCoR, and HDAC2 and 5 [6]. Instead, DP103 greatly enhanced PIASy-dependent

sumoylation of SF-1 and facilitates relocalization of SF-1 and PIASy into novel nuclear

speckles. There was no co-localization between these nuclear speckles and PML

markers, demonstrating that SF-1/PIASy nuclear speckles are distinct from PML bodies.

Further studies on the sumoylated SF-1/PIASy/DP103 protein complex are necessary to

dissect the molecular mechanism of SUMO-mediated repression of SF-1 by DP103 and

the biological significance of the SF-1/PIASy nuclear bodies.

Regulation of SF-1 DNA-binding activity by sumoylation is an additional molecular

mechanism, as revealed through in vitro studies of sumoylated SF-1 protein. The

discovery that Lys119 SF-1 sumoylation regulates DNA-binding at low affinity sites

supports the idea that sumoylation controls expression of a subset of the genes regulated

by SF-1. One remaining question is how to identify these genes in cells. Bioinformatic

predictions of nuclear receptor binding sites are based on the consensus sequence, thus

they cannot identify all sites that are bound by the receptor in vivo [7]. Therefore, an

unbiased approach such as chromatin immunoprecipitation combined with a microarray

containing sequences from promoter regions of the entire genome should help to identify

binding sites occupied by SF-1 on chromatin, especially the sites that deviate from the

CAAGGTCA consensus [8].

Future experiments need to address the possibility of cross-talk between the two SF-1

sumoylation sites. Cooperativity between these two sites could tie together SUMO-

mediated repression of SF-1by DP103 at Lys194 and regulation of SF-1 DNA-binding by

SUMO1 at Lys119. Cellular loss-of-function experiments demonstrate that Lys194 in the SF-1 LBD is the major site of sumoylation. Mutation of Lys194 to arginine leads to a great increase in SF-1 transcriptional activity, especially in the context of the Gal4 system. Analyses of SUMO-modified SF-1 by western blotting support a possibility of ordered sumoylation where sumoylation of Lys194 is required for sumoylation of Lys119 at the DBD [9]. My data show that DNA-bound SF-1 cannot be sumoylated at Lys119. Thus, the order of sumoylation events, especially at low affinity SF-1 binding sites, may start with sumoylation of DNA-bound SF-1 at Lys194 in the presence of a PIAS protein and DP103. The helicase activity of DP103 may be important in dissociating SF-1 from chromatin, as DEAD-box helicases are known to regulate protein-nucleic acid interactions [10]. Once SF-1 is released from DNA, Lys119 can be sumoylated. This will inhibit further DNA-binding at low affinity sites, and sumoylated SF-1 becomes sequestered in nuclear speckles through interaction with the PIAS protein. Of course, many points in this model still have to be explored experimentally. Cellular and biochemical studies using full-length receptor are needed to address the questions of the relative pools of mono-sumoylated SF-1 versus SF-1 sumoylated at both Lys119 and Lys194, the precise relationship between the two SF-1 sumoylation sites, and the regulation of ordered SF-1 sumoylation.

Biological significance of SF-1 sumoylation

The ultimate goal of these studies is to understand the biological significance of sumoylation in SF-1 in vivo function. While much progress has been made in understanding in vivo importance of sumoylation in yeast and *Drosophila*, whole

organism studies in mammals are just emerging [11-13]. Sumoylation is an essential posttranslational modification in mice, as loss of either Ubc9 or SENP1 leads to embryonic lethality [14, 15]. Specifically, consequences of SF-1 loss-of-sumoylation phenotype would be best explored in a knock-in mouse model where two sumoylation sites in SF-1, Lys119 and Lys194, are replaced with arginines (2KR). These studies are currently underway in our laboratory. One possibility is that creating highly active SF-1 protein due to loss of sumoylation would result in embryonic lethality. Transgenic SF-1 overexpression using endogenous SF-1 promoter leads to adrenocortical cell proliferation and tumors, demonstrating that, normally, SF-1 activity is tightly regulated [16]. Attempts to overexpress SF-1 from stronger promoters have failed to give transgene transmission suggesting that having high levels of active SF-1 protein is harmful during embryonic development (H.A.I, unpublished results). SF-1 expression is detectable in placenta indicating a role for SF-1 in embryo implantation and nurturing [17]. If the 2KR embryos survive until birth, based on SF-1 knockout phenotype of male-to-female sex reversal, lack of adrenals and adrenal steroidogenesis, overactive SF-1 protein would result in massive production of adrenal steroids, increase in adrenal size, and may even lead to the masculanization of females [18]. To identify candidate genes regulated by SF-1 sumoylation, expression profiles of SF-1 target tissues such as the gonads, adrenals, and the hypothalamus between wild type and 2KR animals need to be compared using microarray technology. The results of these gene-profiling studies will help to identify specific tissues where SF-1 sumoylation is functionally important and to delineate specific signaling pathways involved in endocrine organ development and function that are regulated by SF-1 sumoylation.

Regulation of SF-1 sumoylation

Studies in this thesis demonstrate that sumoylation is critically important in controlling

the activity of Subfamily V nuclear receptors. Yet, at present, little is known about

whether sumoylation of SF-1 is a regulated or a constitutive process. SUMO E3s and

SUMO isopeptidases present likely candidates for regulation by signaling through cell

surface peptide receptors, MAPK, or cAMP pathways [19]. The fraction of sumoylated

SF-1 in the cell is yet to be determined. It is possible that most SF-1 exists in

unsumoylated state, and only a small population becomes transiently sumoylated by

activation of Ubc9 and specific SUMO E3 ligase. In collaboration with Dr. Martin Lee, I

examined the role of PIAS proteins as SUMO E3 ligases for SF-1. PIASy and PIASxα

were able to increase levels of sumoylated SF-1 in overexpression cell culture studies.

PIASxα tissue distribution overlaps well with SF-1 expression pattern (M.B.L,

unpublished results). However, mouse knockouts of PIASy, PIAS1, and PIASxα

revealed modest defects demonstrating that they are largely dispensable for development

and in adult function [20-22]. These observations suggest a redundant role for these

proteins in the sumoylation pathway.

Current data also allow for an alternative possibility where most SF-1 in cells is

constitutively sumoylated, perhaps during a critical developmental or differentiation

stage. Thus, regulation of SF-1 sumoylation would happen at the level of desumoylation

where transcriptionally-active SF-1 would be liberated by the action of an isopeptidase.

SUMO isopeptidases have defined subcellular localization patterns implying that their

actions may be more substrate-specific [23]. Future research is needed to understand

how and when SUMO isopeptidases are activated and regulated, as well as to define their

substrate specificity. Moreover, the existence of additional accessory proteins that confer

specificity and regulation to SF-1 sumoylation cannot be discounted at this time.

Perhaps, the in vitro sumoylation system for SF-1, established in the course of this thesis,

could be adapted for an unbiased biochemical screening of SF-1-expressing cell lines to

identify additional players in the pathway of SF-1 sumoylation.

References:

1. Faus, H. and B. Haendler, *Post-translational modifications of steroid receptors.* Biomed Pharmacother, 2006. **60**(9): p. 520-8.

2. Seidman, J.G. and C. Seidman, *Transcription factor haploinsufficiency: when half a loaf is not enough.* J Clin Invest, 2002. **109**(4): p. 451-5.

3. Bland, M.L., et al., *Haploinsufficiency of steroidogenic factor-1 in mice disrupts adrenal development leading to an impaired stress response.* Proc Natl Acad Sci U S A, 2000. **97**(26): p. 14488-93.

4. Kohler, B., et al., *Five novel mutations in steroidogenic factor 1 (SF1, NR5A1) in 46,XY patients with severe underandrogenization but without adrenal insufficiency.* Hum Mutat, 2007.

5. Chen, W.Y., L.J. Juan, and B.C. Chung, *SF-1 (nuclear receptor 5A1) activity is activated by cyclic AMP via p300-mediated recruitment to active foci, acetylation, and increased DNA binding.* Mol Cell Biol, 2005. **25**(23): p. 10442-53.

6. Klappacher, G.W., et al., *An induced Ets repressor complex regulates growth arrest during terminal macrophage differentiation.* Cell, 2002. **109**(2): p. 169-80.

7. Klimova, N.V., et al., *[Recognition of the potential SF-1 binding sites by SiteGA method, their experimental verification and search for new SF-1 target genes].* Mol Biol (Mosk), 2006. **40**(3): p. 512-23.

8. Dufour, C.R., et al., *Genome-wide orchestration of cardiac functions by the orphan nuclear receptors ERRalpha and gamma.* Cell Metab, 2007. **5**(5): p. 345-56.

9. Komatsu, T., et al., *Small ubiquitin-like modifier 1 (SUMO-1) modification of the synergy control motif of Ad4 binding protein/steroidogenic factor 1 (Ad4BP/SF-1) regulates synergistic transcription between Ad4BP/SF-1 and Sox9.* Mol Endocrinol, 2004. **18**(10): p. 2451-62.

10. Fuller-Pace, F.V., *DExD/H box RNA helicases: multifunctional proteins with important roles in transcriptional regulation.* Nucleic Acids Res, 2006. **34**(15): p. 4206-15.

11. Chan, H.Y., et al., *Genetic modulation of polyglutamine toxicity by protein conjugation pathways in Drosophila.* Hum Mol Genet, 2002. **11**(23): p. 2895-904.

12. Shih, H.P., et al., *Identification of septin-interacting proteins and characterization of the Smt3/SUMO-conjugation system in Drosophila.* J Cell Sci, 2002. **115**(Pt 6): p. 1259-71.

13. Steffan, J.S., et al., *SUMO modification of Huntingtin and Huntington's disease pathology.* Science, 2004. **304**(5667): p. 100-4.

14. Nacerddine, K., et al., *The SUMO pathway is essential for nuclear integrity and chromosome segregation in mice.* Dev Cell, 2005. **9**(6): p. 769-79.

15. Cheng, J., et al., *SUMO-specific protease 1 is essential for stabilization of HIF1alpha during hypoxia.* Cell, 2007. **131**(3): p. 584-95.

16. Doghman, M., et al., *Increased steroidogenic factor-1 dosage triggers adrenocortical cell proliferation and cancer.* Mol Endocrinol, 2007. **21**(12): p. 2968-87.
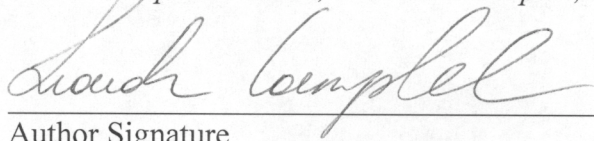
17.	Bamberger, A.M., et al., *Expression of steroidogenic factor-1 (SF-1) mRNA and protein in the human placenta*. Mol Hum Reprod, 1996. **2**(6): p. 457-61.

18.	Sadovsky, Y., et al., *Mice deficient in the orphan receptor steroidogenic factor 1 lack adrenal glands and gonads but express P450 side-chain-cleavage enzyme in the placenta and have normal embryonic serum levels of corticosteroids*. Proc Natl Acad Sci U S A, 1995. **92**(24): p. 10939-43.

19.	Guo, B., et al., *Signalling pathways and the regulation of SUMO modification*. Biochem Soc Trans, 2007. **35**(Pt 6): p. 1414-8.

20.	Roth, W., et al., *PIASy-deficient mice display modest defects in IFN and Wnt signaling*. J Immunol, 2004. **173**(10): p. 6189-99.

21.	Liu, B., et al., *PIAS1 selectively inhibits interferon-inducible genes and is important in innate immunity*. Nat Immunol, 2004. **5**(9): p. 891-8.

22.	Santti, H., et al., *Disruption of the murine PIASx gene results in reduced testis weight*. J Mol Endocrinol, 2005. **34**(3): p. 645-54.

23.	Mukhopadhyay, D. and M. Dasso, *Modification in reverse: the SUMO proteases*. Trends Biochem Sci, 2007. **32**(6): p. 286-95.
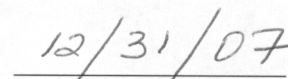
Publishing Agreement

It is the policy of the University to encourage the distribution of all theses and dissertations. Copies of all UCSF theses and dissertations will be routed to the library via the Graduate Division. The library will make all theses and dissertations accessible to the public and will preserve these to the best of their abilities, in perpetuity.

Please sign the following statement:

I hereby grant permission to the Graduate Division of the University of California, San Francisco to release copies of my thesis or dissertation to the Campus Library to provide access and preservation, in whole or in part, in perpetuity.

_____   12/31/07
Author Signature   Date

128