A Gene-Expression Screen Identifies A Non-Toxic Sumoylation Inhibitor That Mimics SUMO-Less Human LRH-1 In Liver

Miyuki Suzawa¹,6, Diego A. Miranda¹,6, Karmela A. Ramos¹,6, Kenny K-H. Ang², Emily J. Faivre¹, Christopher G. Wilson², Laura Caboni³, Michelle R. Arkin², Yeong-Sang Kim⁵, Robert J. Fletterick³, Aaron Diaz⁴, John S. Schneekloth Jr.⁵ and Holly A. Ingraham¹,7

Affiliations

¹Department of Cellular and Molecular Pharmacology, ²Small Molecule Discovery Center, Department of Pharmaceutical Chemistry, ³Department of Biochemistry and Biophysics ⁴Department of Epidemiology and Biostatistics, University of California, San Francisco, San Francisco, California 94158, USA ⁵Chemical Biology Laboratory National Cancer Institute 376 Boyles Street Frederick, MD 21702, USA ⁶These authors contributed equally

⁷Correspondence to: holly.ingraham@ucsf.edu

Running Title: Phenotypic Screen Identifies TA As SUMO Inhibitor

Manuscript Word Count and Display Count

Title: 112 Characters
Running Title: 49
Abstract: 149/150
Display Items: 7
Tables 3
ABSTRACT

SUMO-modification of nuclear proteins has profound effects on gene expression. However, non-toxic chemical tools that modulate sumoylation in cells are lacking. Here, to identify small molecule sumoylation inhibitors we developed a cell-based screen that focused on the well-sumoylated substrate, human Liver Receptor Homolog-1 (hLRH-1, NR5A2). Our primary gene-expression screen assayed two SUMO-sensitive transcripts, APOC3 and MUC1 that are upregulated by SUMO-less hLRH-1 or by siUBC9 knockdown, respectively. A polyphenol, tannic acid (TA) emerged as a potent sumoylation inhibitor in vitro (IC50 = 12.8 µM) and in cells. TA also increased hLRH-1 occupancy on SUMO-sensitive transcripts. Most significantly, when tested in humanized mouse primary hepatocytes, TA inhibits hLRH-1 sumoylation and induces SUMO-sensitive genes, thereby recapitulating the effects of expressing SUMO-less hLRH-1 in mouse liver. Our findings underscore the benefits of phenotypic screening for targeting post-translational modifications, and illustrate the potential utility of TA for probing the cellular consequences of sumoylation.

Impact Statement

Discovery of the FDA-approved compound and plant extract, tannic acid as a non-toxic chemical tool for modulating sumoylation in multiple platforms.
SUMO-modification or sumoylation with the Small Ubiquitin-like Modifier (SUMO) is a prevalent post-translational modification of many transcription factors and is generally associated with transcriptional repression (Gill, 2005). Similar to other ubiquitin-like modifications, the sumoylation cycle is multi-stepped, as reviewed in (Gareau and Lima, 2010), and is initiated by E1 (SAE1), which forms a thioester bond with either SUMO-1, 2, or 3. The single E2 (Ubc9) facilitates the hand-off and covalent conjugation of SUMO on a given protein substrate. Although E3s are believed to guide substrate selectivity in cells, only E1 and E2 are required for in vitro sumoylation. Sumoylation is then reversed by sentrin-specific proteases (SENPs).

Genetic manipulations that either eliminate the sumoylation machinery or permanently disrupt the normal sumoylation cycle of a substrate can result in embryonic lethality or impaired organogenesis (Flotho and Melchior, 2013; Kang et al., 2010; Lee et al., 2011a; Nacerddine et al., 2005; Wang et al., 2014). Our lab and others find that sumoylation represents an important, ligand-independent mode to regulate the NR5A nuclear receptor subfamily that includes Steroidogenic Factor 1 (NR5A1/SF-1) (Campbell et al., 2008; Lee et al., 2011a; Lee et al., 2005) and Liver Receptor Homolog 1 (NR5A2/LRH-1) (Chalkiadaki and Tali­anidis, 2005; Stein et al., 2014; Venteclef et al., 2010; Ward et al., 2013; Yang et al., 2009). For instance, knocking-in an unsumoylatable or SUMO-less mutant SF-1 allele leads to profound changes in endocrine physiology and tissue development (Lee et al., 2011a). Importantly, even in the presence of one wild type allele, the phenotypic effects of SUMO-less SF-1 dominate. Mechanistically, we find that SUMO-less SF-1 can regulate select downstream targets, which we refer to as SUMO-sensitive (Campbell et al., 2008). Indeed, sonic hedgehog (SHH) signaling is ectopically activated after expressing SUMO-less SF-1 in both cells and tissues. This gain-of-function or dominance of the SUMO-less SF-1 mutant leads to expansion of select cell types and hormone imbalance, illustrating how disrupting the normal cycle of substrate sumoylation results in...
Phenotypic Screen Identifies TA As SUMO Inhibitor

disease states. Consistent with our findings, gain-of-function heterozygous missense mutations in the sumoylation site of the transcription factor MITF are tightly linked with some forms of human melanoma (Bertolotto et al., 2011). SUMO-less variants of the androgen (AR) and glucocorticoid hormone (GR) receptors are also reported to activate new transcriptional programs linked to cellular proliferation (Paakinaho et al., 2014; Sutinen et al., 2014). These studies suggest that successful efforts to chemically target substrate sumoylation could be used to alter transcription factor activity, to either promote or attenuate SUMO-sensitive genetic programs.

Thus far, efforts to drug sumoylation using in vitro target-based assays and in silico screens have identified different classes of SUMO inhibitors with IC\textsubscript{50}s in the micromolar range (Table 1). In vitro target-based screens rely exclusively on defined components (E1, UBC9 and a test substrate), and as such, exclude the contributions by E3s and other unidentified obligate cofactors on substrate sumoylation. An in situ cell-based screen in permeabilized, fixed cells partially overcomes this limitation, but still requires the addition of exogenous SUMO components for the assay (Hirohama et al., 2013). Currently, inhibitors that target E1 include ginkgolic acid (GA), davidiin, and kerriamycin B. Inhibitors of UBC9 include 2-D08, GSK145A and spectomycin B1. While GA remains the most widely used and commercially available chemical probe targeting general sumoylation, its efficacy as an inhibitor can vary greatly depending on the assay and substrate being tested (Bogachek et al., 2014; Kim et al., 2013; Tossidou et al., 2014).

Phenotypic cell-based screens offer an alternative approach to in vitro target-based screens for finding new molecular entities (Swinney and Anthony, 2011). Here, we set out to identify non-toxic chemical probes that would modulate substrate sumoylation using a cell-based gene-expression screen, which assayed two human LRH-1 (hLRH-1) SUMO-sensitive transcripts as the primary readout. Human LRH-1 is an ideal test substrate for evaluating any potential hits because one has the ability to test how candidate hits affect LRH-1 activity in both immortalized
Phenotypic Screen Identifies TA As SUMO Inhibitor

hepatocellular carcinoma cells and in primary mouse hepatocytes (Lee et al., 2011b; Mataki et al., 2007; Oosterveer et al., 2012). In addition, as shown in this study, hLRH-1 is well-sumoylated at multiple sites in vitro, in cells, and in vivo. From the initial phenotypic screen of the FDA- and European-approved Pharmakon 1600 drug library, the commercial plant extract, tannic acid (TA) was identified as an effective, non-toxic general sumoylation inhibitor in multiple platforms, including primary mouse hepatocytes.

RESULTS

A Phenotypic Screen Assaying SUMO-sensitive Genes Identifies TA As The Top Hit

Sumoylation of hLRH-1 occurs primarily in the flexible hinge domain on two major conserved acceptor lysines K192 and K270, with a minor site located in the N-terminal region at K44 (Figure 1A). Similar to our prior results obtained with SF-1, hLRH-1 is efficiently sumoylated (~30%) in human placental choriocarcinoma JEG3 and hepatocellular carcinoma HepG2 cells expressing Flag-hLRH-1 (Figure 1A). Importantly, multiple sumoylated hLRH-1 species are readily detected with only the endogenous SUMOylation machinery and without the need to add exogenous SUMO or Ubc9. Substituting lysines K192 and K270 with arginines (K192R/K270R or 2KR) eliminates nearly all sumoylated LRH-1 species, as previously noted (Chalkiadaki and Talianidis, 2005; Lee et al., 2005; Yang et al., 2009) and Figure 1 - figure supplement 1. In vivo sumoylation of hLRH-1 is also equally robust, as observed in mouse liver humanized to express wild type hLRH-1 (WT) (Figure 1B), following viral-mediated infection with adeno-associated virus serotype 8 (AAV8) (Cotugno et al., 2012; Ill et al., 1997). Impressively, the extent and pattern of hLRH-1 sumoylation in mouse liver are identical to those found in cultured cell lines, demonstrating for the first time that hLRH-1 is efficiently sumoylated at multiple lysines in vivo. By contrast, expressing SUMO-less hLRH-1 (2KR) eliminates nearly all hLRH1 sumoylation (Figure 1B), despite the fact that SUMO-less hLRH-1 transcripts and protein are efficiently expressed in mouse liver and in mouse hepatocytes (Figure 1C, D - figure supplement 2).
These collective data establish that hLRH-1 is robustly sumoylated in several platforms, making it an excellent test substrate for assessing both the biochemical and functional effects of small molecule inhibitors of sumoylation.

Rather than assessing substrate sumoylation directly, we used a cell-based screen that monitored two SUMO-sensitive transcripts as the endpoint (diagrammed in Figure 2A). The JEG3 cell line was initially chosen because it performed well in all steps of the primary screen and has been used previously to study NR5A activity (Campbell et al., 2008). Profiling was carried out on JEG3 cells stably expressing hLRH-1 or the SUMO-less hLRH-1 mutant, or after siRNA knock down of UBC9 (siUBC9) to identify the most robust SUMO-sensitive genes. To ensure that minor sumoylation on hLRH-1 was eliminated we used the 3KR mutant that disrupts K44, as well as the two major acceptor lysines in the hinge region. Two genes, APOC3 and MUC1 were identified by microarray as the readout transcripts for the screen. These genes are highly induced by either SUMO-less LRH-1 or by siRNA knock down of UBC9 (siUBC9) (Figure 2B), and were chosen as two SUMO-sensitive genes in the primary screen assay. Interestingly, APOC3 can be directly regulated by LRH-1 (Hwang-Verslues and Sladek, 2008), whereas MUC1 is regulated by the androgen receptor (Rajabi et al., 2011). In contrast, expression of a well-known NR5A downstream target gene, CYP11A1 is unaffected both by SUMO-less LRH-1 and by siUBC9 knockdown, and is thus designated as a SUMO-insensitive LRH-1 target (Figure 2B).

A gene-expression based screen adapted from (Arany et al., 2008) assayed APOC3 and MUC1 with a 1600-compound drug library (Pharmakon 1600). JEG3 hLRH-1 cells were cultured in a 384-well format, treated with 10 µM of each drug, and measured for APOC3 and MUC1 transcripts. Robust Z-scores for each drug treatment were obtained by normalizing the amplification cycle number (C$_T$) of APOC3 or MUC1 to TBP as an internal control (ΔC$_T$), and then to the DMSO external control (ΔΔC$_T$). A scatter plot of Z-scores ± 2SD from the primary
Phenotypic Screen Identifies TA As SUMO Inhibitor

screen shows 13 drugs producing significant changes in MUC1 and APOC3 expression (up and down), representing a 0.8 % hit rate (Figure 2C and Source data 1). Drugs resulting in cytotoxicity or changing TBP expression were then filtered out leaving six potential hits listed in Table 2.

TA is a Potent Inhibitor of Substrate Sumoylation In Cells and In Vitro

All potential non-toxic candidates were repurchased and tested in a dose response for induction of the two SUMO-sensitive genes used in the primary assay. Only tannic acid (TA) emerged as a valid hit from the primary screen, showing significant induction of APOC3 and MUC1, but not CYP11A1 (Figure 3A). We then asked if hLRH-1 sumoylation is required to observe the stimulatory effects of TA on APOC3 and MUC1. As expected, TA had no effect on CYP11A1 regardless of the hLRH-1 variant tested (Figure 3B). On the other hand, the dose-dependent effects of TA on APOC3 are lost when tested with the SUMO-less hLRH-1 (3KR), demonstrating that activation of APOC3 by TA depends on the ability of hLRH-1 to be sumoylatable. Further, APOC3 upregulation by TA depends on hLRH-1; TA failed to change APOC3 alone (Figure 3 - figure supplement 1). On the other hand, the effects of TA on MUC1 levels are largely independent of hLRH-1, implying that TA can act more broadly on non-hLRH-1 SUMO-sensitive targets.

The transcriptional effects of TA correlated well with the dose-dependent inhibition of hLRH-1 sumoylation in cells, as evidenced by a ~40-50 % decrease in all hLRH-1 sumoylated species at 10 µM TA (Figure 3C). Diminished hLRH-1 sumoylation is observed only after TA and GA treatment; the other published sumoylation inhibitor, 2’,3’,4’-trihydroxyflavone (2-D08) (Kim et al., 2013), and another top screening hit, trifluridine (Tri) failed to show similar results (Figure 3D). However, while GA appears more effective than TA at reducing hLRH-1 sumoylation in JEG3 cells (Figure 3D), this compound leads to significant cytotoxicity beginning
Phenotypic Screen Identifies TA As SUMO Inhibitor

at 10 µM after 5 h (Figure 3E) and 24 h exposure (Figure 3 - figure supplement 2). By contrast, TA is non-toxic even at higher concentrations.

Next, we asked whether TA would also inhibit in vitro sumoylation (IVS) of recombinant full-length human LRH-1 (FL-hLRH-1) protein. In our assay conditions, the pattern of sumoylated FL-hLRH-1 in vitro is identical as that found in cells and in vivo, and collapses down to a single unmodified band after addition of SENP1 (Figure 4A and refer back to Figure 1). Using IVS conditions that achieve ~50% sumoylation of FL-hLRH-1, TA is the most effective sumoylation inhibitor when compared to 2-D08 and GA, as well as other candidate hits, (Figure 4B), with an apparent IC₅₀ of 12.8 µM (Figure 4C and Figure 4 - figure supplement 1). As predicted from our initial cellular data, TA inhibits and impairs the rate of sumoylation of other substrates, including recombinant hinge-LBD SF-1, full-length IκBα, and an AR peptide (Figure 4 - figure supplement 2). As found for davidiin, another tannin sumoylation inhibitor (Takemoto et al., 2014), TA impairs E1 thioester formation in non-reducing conditions (Figure 4D).

Given that TA and other polyphenols are prone to aggregate formation which can lead to non-specific or promiscuous inhibition (Feng and Shoichet, 2006), we assessed the performance of TA in the presence of a non-ionic detergent (Triton X-100), which limits aggregate formation (Pohjala and Tammela, 2012). Even in the presence of 0.01% Triton X-100, FL-LRH-1 sumoylation was inhibited by TA treatment, as shown in Figure 4E. That TA maintains its inhibitory activity with this detergent suggests that in these assay conditions, the inhibitory effects of TA are not due to non-specific aggregate formation.

Next, we tested whether TA is also effective in hepatocellular carcinoma HepG2 cells, a relevant cell line for studying hLRH-1, (Figure 5A). When compared to a non-toxic dose of GA (10 µM) or 2-D08, TA is much more efficient at decreasing levels of sumoylated hLRH-1, and shows little cytotoxicity in HepG2 cells after even at longer exposure times (Figure 5B). We also directly compared the effects of TA versus knockdown of UBC9 (siUBC9) on hLRH-1 and other sumoylated proteins in HepG2 cells. Surprisingly, despite a substantial loss of UBC9 transcripts
Phenotypic Screen Identifies TA As SUMO Inhibitor

(95%) and protein (60%) following siRNA-mediated knockdown for 72 h (Figure 5 - figure supplement 1), hLRH-1 remained fully sumoylated (Figure 5C). On the other hand, TA levels of hLRH-1 sumoylation decreased in a dose-dependent manner (Figure 5C); higher migrating sumoylated hLRH-1 species in HepG2 cells that diminish with TA were authenticated as SUMO-1 or SUMO-2 species (Figure 5D). Interestingly, 1x-Su-hLRH-1 appears to be exclusively modified by SUMO1, whereas higher Su-hLRH-1 species are modified by SUMO2. Both siUBC9 and TA inhibit global sumoylation, but in this setting, siUBC9 is slightly more effective (Figure 5E). As expected, neither TA nor siUBC9 changes the pool of ubiquitinated proteins. Lastly, we wanted to know if TA similarly modulates sumoylation of endogenously expressed LRH-1. Unfortunately, our ability to cleanly detect or efficiently pulldown endogenous sumoylated LRH-1 species in cells/tissues is difficult with available anti-LRH-1 antibodies, including both commercial and non-commercial sources. Instead, we used human adrenal carcinoma H295R cells that express high levels of endogenous hSF-1, and found that similar to exogenously expressed hLRH-1, TA decreases levels of sumoylated hSF-1 in H295S cells (Figure 5F). Similar results were also obtained for endogenous RanGap (Figure 5 - figure supplement 2).

Taken together, these data show that TA is a non-toxic potent global sumoylation inhibitor.

TA Enhances hLRH-1 Activity and Alters SUMO-Sensitive Targets in HepG2 cells.

To ask if TA alters SUMO-sensitive targets in HepG2 cells, we identified a small subset of target genes that is 1) upregulated by hLRH-1, 2) bound by LRH-1 as detected by chromatin-immunoprecipitation high-throughput sequencing (ChIP-Seq) and 3) altered by TA in the presence of hLRH-1 after profiling HepG2 cells (Source data 2). Of the 42 genes in this small subset, three genes that were considered SUMO-sensitive and also found to be putative LRH-1 target genes by ChIP-Seq analyses (Source data 2) were assayed in more detail; they include CYP24A1, PFKFB3 and SERPINE1 (Figure 6A, B). TA regulates all three genes in a dose-dependent manner (Figure 6C). Moreover, using ChIP-qPCR we find a significant increase in
hLRH-1 occupancy on these LRH-1 binding sites, as shown for a site in the SERPINE1 proximal
promoter region and also for an intronic site in CYP24A1 (Figure 6D). These cellular data
suggest that TA modulates and increases recruitment to SUMO-sensitive hLRH-1 target genes
in a relevant HepG2 cellular model system.

**TA Functionally Mimics SUMO-less LRH-1 In Vivo.**

We then determined if TA functions as a non-toxic sumoylation inhibitor in primary mouse
hepatocytes and is able to recapitulate gene expression changes observed with SUMO-less
hLRH-1. Limited profiling was carried out to identify SUMO-sensitive genes in infected mouse
liver overexpressing wild type or SUMO-less hLRH-1 (2KR). Both wild type and SUMO-less
hLRH-1 were expressed equivalently with levels ~10-fold higher than that of endogenous
mLRH-1, as judged by immunoblots using an anti-LRH-1 antibody that detects both mouse and
human LRH-1 (Figure 7A). Expression of endogenous *mLrh*-1 is unaffected after infecting mice
with AAV8-hLRH-1 (refer back to Figure 1 - figure supplement 2). Although wild type and
SUMO-less hLRH-1 upregulate the classic LRH-1 target *Cyp8b1* (Lee et al., 2008), expression
of SUMO-less hLRH-1 leads to robust activation of *adiponectin* (*Adipoq*) and *sonic hedgehog*
(*Shh*), in mouse liver (Figure 7B). The pattern of hLRH-1 sumoylation is preserved in cultured
hepatocytes, but lost in primary cultures expressing the SUMO-less hLRH-1 mutant (Figure 7C,
left panel). Importantly, when tested in primary mouse hepatocytes TA (10 µM, 5 h) diminishes
levels of sumoylated wild type hLRH-1, with nearly all sumoylated species absent including 1x-
SUMO-hLRH-1 (Figure 7C, right panel); thus closely resembling the SUMO-less 2KR mutant.

SUMO-sensitive genes identified in whole liver were then tested with TA in mouse primary
hepatocytes that express only endogenous mLRH-1. As found in the liver, *Shh*, its downstream
effector *Gli2*, and *Adipoq* are essentially switched-on by TA (Figure 7D). This trigger-like effect
of TA and SUMO-less hLRH-1 on *Shh* and *Gli2*, is consistent with our previous findings that
hedgehog signaling can be ectopically activated in mouse organs after knocking-in SUMO-less
SF-1 (Lee et al., 2011a). Importantly, cell viability was unchanged by TA after 24 h exposure (Figure 7 - figure supplement 1). These data establish that in primary cells, TA abrogates hLRH-1 sumoylation and induces hLRH-1 SUMO-sensitive downstream target genes.

DISCUSSION

Here, using a phenotypic, gene-expression based screen we identify the polyphenol, tannic acid as a potent inhibitor of hLRH-1 sumoylation in multiple platforms. In vitro assays confirm that TA impairs substrate sumoylation and in cellular model systems, an acute, non-toxic dose of TA treatment markedly reduces levels of sumoylated hLRH-1, as well as hSF-1, and modulates expression of hLRH-1 target genes. Impressively, in primary cultures of mouse hepatocytes, TA inhibits hLRH-1 sumoylation and is able to mimic the transcriptional output of the SUMO-less LRH-1 mutant by switching on genes that have been associated with hepatic injury.

Our gene-expression based screen was designed to sample the transcriptional output of substrate sumoylation, which has not been possible with in vitro target-based screens. Two factors make hLRH-1 an excellent test substrate in our initial and follow-up assays. First, our earlier work established that blocking the NR5A sumoylation cycle results in robust transcriptional changes that could be leveraged in a phenotypic screen. We reasoned that a transcriptional endpoint makes it possible to integrate the influence of other factors that promote substrate-specific sumoylation or desumoylation, including possible synergistic effects of multiple transcription factor sumoylation (Holmstrom et al., 2008; Komatsu et al., 2004). Second, hLRH-1 is well-sumoylated with a stereotypic pattern of multiple sumoylated species when expressed in multiple cell types and in liver. This pattern is readily duplicated in vitro, allowing one to also test candidate sumoylation modulators in a defined cocktail. In our hands, multiple sumoylated species of Flag-tagged hLRH-1 are readily detected in HEK293S, JEG3 and HepG2 cells, in primary hepatocytes and in liver, with only the endogenous sumoylation machinery;
Phenotypic Screen Identifies TA As SUMO Inhibitor

these findings differ from the need for exogenous SUMO1/2 and E3s required to detect 1X-sumoylated V5-tagged mLRH-1 in HEK293 cells or primary hepatocytes (Stein et al., 2014). Finally, the number and intensity of hLRH-1 sumoylated species observed in cells or in AAV-infected liver tissue far exceed the faint, single FXR sumoylated band detected when human FXR is expressed in cells and in mouse liver (Kim et al., 2015). We posit that NR5As might be particularly good substrates for sumoylation because their major acceptor lysines reside in the flexible hinge domain, possibly promoting strong protein-protein interactions previously observed between Ubc9 and both SF-1 and LRH-1 (Lee et al., 2005). Indeed, the residual UBC9 protein following siUBC9 likely accounts for the persistent levels of 1xSU-LRH-1 and suggests that once UBC9 is charged with SUMO1/2, sumoylation of LRH-1 proceeds efficiently. Interestingly, as with siUBC9, targeting E2 by 2-D08 is also ineffective at blocking hLRH-1 sumoylation. Regardless of the underlying signals that confer relatively high basal levels of NR5A sumoylation in cells and in tissues, the consistent and robust sumoylation observed for hLRH-1 were instrumental in facilitating follow-up studies on candidate small molecule hits from our primary screen.

While polyphenols such as TA are easily dismissed as promiscuous inhibitors and false positives in high-throughput screens (Feng and Shoichet, 2006; Pohjala and Tammela, 2012), the effects of TA in our sumoylation assays are extremely reproducible in multiple cell lines and in primary hepatocytes. Interestingly, bergapten and coumarin 153, two well-characterized colloid-forming phenolic compounds (Pohjala and Tammela, 2012) present in the Pharmakon library, failed to emerge a hits in the primary screen. In our IVS assay conditions, TA is more effective than 2-D08 and GA and can inhibit sumoylation of multiple substrates in vitro, including SF-1 and a fragment of the androgen receptor. Consistent with GA-mediated decrease in HepG2 cell viability noted by (Liu and Zeng, 2009), we also noted that GA leads to significant cell toxicity. Hence, while GA might decrease sumoylation as an adaptive response to cell death, the utility of GA in assessing the transcriptional responses of substrate sumoylation is
Phenotypic Screen Identifies TA As SUMO Inhibitor

The fact that 2-D08 does not inhibit hLRH-1 sumoylation but is effective on other substrates (AR peptide and FL-IκBα) might reflect the fact that 2-D08 fails to block the strong interactions between Ubc9 and NR5As. To date, the SUMO-interacting motifs in hLRH-1 that mediate this interaction have yet to be identified. Nonetheless, these data imply that mechanistically distinct sumoylation inhibitors act on different classes of substrates.

Our data suggest strongly that TA inhibits E1 thioesterization, as found for the ellagitannin, Davidiin (Takemoto et al., 2014). The known aggregate formation and anti-oxidant properties of TA appear to be less important in inhibiting substrate sumoylation. Indeed, TA inhibits FL-hLRH-1 sumoylation even in the presence of detergent. Polyphenols, including TA are also antioxidants and can scavenge reactive oxygen species (ROS) during oxidative stress (Chen et al., 2007; Yazawa et al., 2006), which might also directly affect the equilibrium between sumoylation-desumoylation (Bossis and Melchior, 2006). In this regard, we find that two other antioxidants, ellagic acid and EGCG are ineffective at inhibiting hLRH-1 sumoylation (data not shown). Furthermore, conditions in our IVS assays are highly reducing making it unlikely that TA inhibits LRH-1 sumoylation via its antioxidant properties in this setting.

That TA is effective at blocking hLRH-1 sumoylation in humanized primary hepatocytes greatly strengthens the validity of TA as a useful chemical tool to assess the cellular effects of sumoylation. Interestingly, TA is more effective at blocking hLRH-1 sumoylation in primary hepatocytes as compared to HepG2 cells where 1X SUMO-LRH-1 persists even at the highest dose of TA; a similar trend was noted for endogenous hSF-1 in H295R cells. The lower efficacy of TA in immortalized cell lines may reflect an increase in the general sumoylation machinery in immortalized versus primary cells, as noted by (Bellail et al., 2014). The use of humanized mouse hepatocytes and the dramatic changes in adiponectin and sonic hedgehog transcripts observed may begin to provide new insights into the in vivo function of LRH-1 sumoylation. The ectopic activation of SHH signaling observed here with SUMO-less hLRH-1 and TA in primary hepatocytes confirms our earlier work showing that eliminating SF-1 sumoylation activates...
hedgehog signaling in endocrine tissues (Lee et al., 2011a). Interestingly, hyperactivation of hedgehog signaling in liver is associated with non-alcoholic steatohepatitis (NASH) progression and responses to liver injury (Grzelak et al., 2014; Guy et al., 2012; Hirsova and Gores, 2014). Adiponectin is an adipocyte-derived protein that reduces fatty liver (Xu et al., 2003) and appears protective against NASH (Asano et al., 2009). Indeed, while adiponectin is normally never expressed in liver, hepatic adiponectin transcripts are observed in rats after chemically induced hepatotoxicity (Yoda-Murakami et al., 2001) and in patients with fatty liver or fully progressed NASH (Uribe et al., 2008). That both SUMO-less LRH-1 and TA switch on hepatic adiponectin and hedgehog signaling leads us to speculate that tipping the balance of the hLRH-1 sumoylation cycle towards desumoylation initiates adaptive responses to liver injury, and eventually a pro-inflammatory response, as suggested by others (Venteclef et al., 2010).

Interestingly, a global knock-in of a single SUMO mutation (K289R) in mouse LRH-1, which is equivalent to K270R in human LRH-1, has no phenotype on its own, but mitigates aortic plaque formation in Ldlr\(^{-/-}\) arteriosclerosis-prone mice (Stein et al., 2014). Hence, revealing the full physiological consequences of LRH-1 sumoylation may require the elimination of both major sumoylation sites in the flexible hinge domain and the use of conditional knock-in strategies that are specific for the adult liver.

In summary, using a novel cell-based assay, we report that the commercially derived, plant extract tannic acid is a useful, non-toxic chemical tool for assessing the transcriptional and cellular effects of sumoylation in both immortalized and primary cell cultures. Based on our collective studies that have focused on the sumoylation of NR5As, we propose that the ratio of sumoylated to desumoylated substrate can be chemically manipulated to switch on and off sumo-sensitive transcriptional programs. Clearly, continued efforts are needed to determine whether more selective chemical tools can be found that promote or block sumoylation of a given substrate.
MATERIALS AND METHODS

Cell lines and Transfections

To generate tetracycline (TET)-inducible Flp-In T-REx stable JEG3 cells, 3x Flag-tagged wild type (WT) and 3KR (K44/192/270R) hLRH-1 in pcDNA5/FRT/TO (Life technologies) were cloned into pcDNA5/FRT/TO expression vectors. Flp-In T-REx JEG3 cell lines were transfected with WT and 3KR hLRH-1 in pcDNA5/FRT/TO and selected with 100 or 125 μg/ml Hygromycin B (Gemini Bio-Products), respectively. JEG3-hLRH-1 cells were treated with tetracycline (100 ng/mL, Teknova Laboratory) for 6 h to induce WT or SUMO-less LRH-1 proteins.

For making doxycycline (Dox)-inducible HepG2 3G stable cells, we cloned 3x Flag-tagged WT and 2KR (K192/270R) hLRH-1 into pTRE 3G vectors (Clontech). The TET-On 3G HepG2 parental cell line was a generous gift from Dr. Stephen Hand (Li et al., 2012). Linearized expression vectors were transfected in TET-On 3G parental HepG2 cells and selected with 250 μg/mL Hygromycin B (Gemini Biosciences). For detecting wild type or SUMO-less LRH-1 expression, HepG2 3G cells were treated with 200 ng/mL Dox (Sigma-Aldrich) for 6 h. For siUBC9 knockdowns, Ubc9 (SI04185937, SI04368420) and non-silencing control (SI03650318) siRNA were purchased from Qiagen. SiRNA at 5 nM final concentration was reverse-transfected into JEG3 or HepG2 WT hLRH-1 stable cells by RNAiMax (Life Technologies) for 72 h followed by induction of hLRH-1 expression by addition of 100 ng/mL TET for 24 h to JEG3 cells with or 250 ng/mL Dox for 6 h for HepG2 cells.

Cell Viability Assay

For cell viability assays, JEG3 hLRH-1 or HEPG2 hLRH-1 cells were plated on 24-well plates in 0.5 ml of media. Primary hepatocytes were seeded on 96-well plates in 0.1 ml of media. The following day, fresh media was applied with compounds or DMSO control. After 5 or 24 h treatment, cell viability was assayed using CelltiterGlo (Promega) according to manufacturers
instructions. Relative luminescence was measured with Veritas Microplate Luminometer (Turner BioSystems) with an integration time of 1.0 s and normalized to the DMSO control.

**Immunoprecipitation and Western Blotting**

Cells were lysed in RIPA buffer (6 mM Na$_2$HPO$_4$, 4 mM NaH$_2$PO$_4$, 2 mM EDTA pH 8.0, 150 mM NaCl, 1% NaDOC, 0.1% SDS and 1% NP-40) and tissues were lysed in Tris-SDS buffer (2% SDS, 0.6 M Tris-Cl pH 8.0 and 0.1 M DTT) supplemented with protease inhibitors (Calbiochem) and 20 mM N-Ethylmaleimide (NEM; Sigma-Aldrich) and sonicated using the Diagenode Bioruptor. Lysates were clarified by centrifugation and protein concentration was measured using Protein assay Dye reagent concentrate (Bio-Rad) according to the manufacturer's protocol. The following antibodies and concentrations were used: anti-Flag M2; 1:7500 (Sigma-Aldrich), anti-LRH-1; 1:3000 for mouse liver and 1:10,000 for in vitro assay (R&D), anti-SF-1; 1:1000 (Upstate, EMD Millipore), anti-UBC9; 1:1000 (Cell Signaling), anti-SUMO1; 1:1000 (Developmental Studies Hybridoma Bank), anti-SUMO2; 1:1000 (Life technologies) and Ubiquitin monoclonal P4D1 antibody; 1:1000 (Cell Signaling), HRP-conjugated anti-βactin 1:2500 (Cell Signaling) and anti-GAPDH 1:5000 (Santa Cruz).

**Quantitative Real-time Polymerase Chain Reaction (qPCR)**

Total RNA from cells and tissues were isolated using Trizol Reagent (Life Technologies) and PureLink RNA mini kit (Life Technologies), respectively. DNase-treated 1 µg total RNA was used to generate cDNA using High-Capacity cDNA Reverse Transcription kits (Life Technologies). RT-qPCR was performed using SYBR Green, High ROX (Biotool or Quanta) and data analyzed essentially as described (Kurrasch et al., 2007). Sequences for all primer pairs used for qPCR reactions are listed in Supplementary File 1.

**Primary Screening Workflow and Gene-Expression Based qPCR Assays**
Phenotypic Screen Identifies TA As SUMO Inhibitor

JEG3 WT hLRH-1 cells (5000 cells per well) were plated into 384-well by cell dispenser Wellmate (Thermo Scientific) for 24 h. Using an FDA- and European-approved Pharmakon library of 1600 compounds (MicroSource Discovery Systems, Gaylordsville, CT) drugs were pinned at a concentration of 10 µM in 0.1% DMSO using a Biomek FXP (Beckman). At the same time, cells were treated with Tet (100 ng/ml) for inducing wild type hLRH-1. Twenty-four hours later, cells were washed once in PBS and then lysed in 25 µl of lysis buffer provided in the TurboCapture 384 mRNA Kit (Qiagen) using EL406 microplate washer (BioTek). After a 10 min incubation at 37°C, 20 µl of cell lysate was transferred to 384-well oligo (dT)-coated plate (Qiagen) using Biomek FXP and incubated at room temperature for 90 min with shaking. Plates were washed three times with washing buffer and reverse transcription was performed in the same well using High-Capacity cDNA Reverse Transfection kits (Life Technologies), according to the manufacturer’s instruction with a total volume of 20 µl. Aliquots of cDNA was delivered to 384-well qPCR plates using a Biomek FXP Liquid handler and stored at -20°C for subsequent qPCR assays. See Supplementary File 2 for further details. Retesting of top candidates was carried out with repurchased drugs from sources listed in Supplementary File 3.

For qPCR assays, the master-mix buffer (8 µl) containing PCR oligos and Quanta qScript cDNA SYBR Mix (Quanta Biosciences) was added to cDNA (2 µl). All qPCR assays were performed using an ABI 7900HT instrument. Data were analyzed using the ΔΔC_T method. Average of ΔC_T from DMSO-treated samples (N=58) was used as external controls. MUC1 and APOC3 genes were used as the end-point read-outs for the SUMO-sensitive genes.

Calculations:

ΔC_T = Gene of interest (APOC3 or MUC1) - C_T House Keeping gene (TBP) C_T

ΔΔC_T = ΔC_T (Drug) - ΔC_T (DMSO)

For Reference: 1C_T change = 2 fold change

Z-score: -(Drug ΔΔC_T – Average ΔΔC_T (for all 1600 Drugs))/SD ΔΔC_T (for all 1600 Drugs)

For Reference:  Positive Z-score = Upregulation of MUC1/APOC3

Negative Z-score = Downregulation of MUC1/APOC3

Page 17
In Vitro Sumoylation and Thioester Assays

Full length LRH-1 (aa1-541; UniprotKB entry: O00482) was subcloned into pRSF-2 vector (Novagen™) and grown in E. Coli BL21Star (DE3) cells (Invitrogen) at 16°C for 16-18 h to and OD 0.6-0.8 and induced with IPTG (0.2 mM). Cells were resuspended in lysis buffer A (20 mM Tris HCl pH 8, 300 mM NaCl, 10% glycerol, 40 mM Imidazole, 5 mM β-mercaptoethanol (BME), 1 mM CHAPS) supplemented with protease inhibitors (Roche). HLRH-1 protein was purified using Ni-nitrilotriacetate beads (Qiagen) and eluted with Buffer A and 300 mM imidazole. Eluted hLRH-1 was bound to 24 bp duplex region of the Inhibin-A promoter (5’-GGAGATAAGGCTCATGGCCACAGA-3’) to stabilize protein and was further purified by size exclusion chromatography in Superdex 200 (16/60). Native gel electrophoresis confirmed that the hLRH-1/DNA complex eluted as a monomer.

His-tagged components of sumoylation reactions including hE1, hUBC9 and hSUMO1 were grown to an OD 0.3-0.7 and induced with IPTG (0.35 mM) for 5 h at 22°C. Cells were lysed in 50 mM Tris-HCl, pH 8.0, 150 mM NaCl and 5% glycerol supplemented with protease inhibitors and proteins purified as described (Reverter and Lima, 2009; Yunus and Lima, 2009). Full length human LRH-1 was prepared and lysed in 20 mM Tris-HCl pH 8.0, 1 mM CHAPS, 10% glycerol, 5 mM BME, 20 mM imidazole, and 300 mM NaCl supplemented with protease inhibitors and then eluted with lysis buffer with 300 mM Imidazole. IVS reactions were performed in at 37°C for 1 h using 0.1 µM E1, 10 µM UBC9, 30 µM SUMO1 and 1 µM full-length hLRH-1 substrate (Ward et al., 2013) in 50 mM Tris-HCl, 100 mM NaCl, 10 mM MgCl2, 2 mM DTT and initiated by addition of freshly made 10 mM ATP. Aggregation assays used 0.01% Triton X-100 (Sigma). IVS reactions were quenched with 4x Laemmli Buffer with BME, boiled for 5 min and loaded onto a Novex Nupage 4-12% Bis-Tris gel and transferred to nitrocellulose membranes followed by incubation with mouse anti-LRH-1 (1:7500, R&D) or mouse anti-SUMO1 (1:325, DSHB). Proteins were visualized using LiCor Odyssey system and goat anti-
mouse 800 (1:20,000 LiCor, Pierce) and quantitated by Image Studio Lite. Percent conversion was calculated by the ratio of sumoylated protein over total signal per reaction normalized to DMSO control. Concentration curves were derived from at least three independent reactions and fit with nonlinear fitting of $\log_{10} [\mu M \text{TA}]$ versus variable slope using Prism graphing software (GraphPad). In vitro sumoylation of full-length IκBα and fluorescent AR peptide was performed as previously described (Kim et al., 2013). Conditions for the thioester assay were as described above, but with only E1 and SUMO1 proteins added to IVS reactions.

**Microarrays**

Human Exonic Evidence Based Open-source (HEEBO) arrays were printed at the UCSF Center for Advanced Technology (CAT). Hybridization conditions were carried out in Flp-In T-REx JEG3 cells as previously described (Kurrasch et al., 2007) to identify top genes upregulated by expression of SUMOless hLRH-1 versus WT hLRH-1, or after Ubc9 knockdown as described above. For siRNA experiments, HepG2 hLRH-1 cells were reverse-transfected with 5 nM of pooled siRNA directed against human siUBC9 or siRNA control from Qiagen, with RNAiMAX transfection reagent (Life technologies) according to the manufacturer’s protocol. Seventy-two hours after siRNA transfection, wild type hLRH-1 was induced with DOX for 6 h. Total RNA was purified using RNAeasy kit (Life technologies) according to the manufacturer’s protocol. Hybridizations were performed at 65°C for 16 h using mixers compatible with the MAUI hybridization systems (BioMicro Systems). Arrays were scanned using an Axon Scanner 4000B, and data analyzed by GenePix 6.0 software (Molecular Devices). Heat maps were generated using TreeView software.

**Chromatin Immunoprecipitation – DNA Sequencing**

HepG2 hLRH-1 cells were seeded (4 x 10⁵) on 10 cm plates overnight, induced with 250 ng/mL DOX and treated with DMSO or 50 μM Tannic acid for 6 h. Cells were cross-linked with
Phenotypic Screen Identifies TA As SUMO Inhibitor

1% formaldehyde for 3 min at room temperature and quenched by addition of 400 mM glycine. Cells were harvested in 50 mM HEPES-KOH pH 7.4, 1 mM EDTA, 150 mM NaCl, 10% glycerol and 0.5 % Triton X-100, swelled for 40 min at 4°C, then nuclei were pelleted at 600 x G for 5 min and resuspended in RIPA buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA, 150 mM NaCl, 5% glycerol, 0.1% sodium deoxycholate, 0.1% SDS, 1% Triton X-100)(Watson et al., 2013). Lysates were sonicated for a total of 30 min (30 s on, 30 s off, 5 min intervals) with a Diagenode Biorupter UD-200 on High setting at 4°C. Sonicated chromatin was clarified by centrifugation then IP’d with 1 µg anti-Flag M2 antibody pre-conjugated to 10 µl Protein G Dynabeads (Invitrogen) for 2 h at 4°C. Bound protein were washed with 500 mM NaCl and LiCl buffer before reverse cross-linking and proteinase K digested overnight. DNA was isolated using Zymogen ChIP DNA Clean and Concentrator columns and pooled for deep sequencing. ChIP DNA was sent to Hudson Alpha Genomic Services Laboratory for library preparation using Illumina TruSeq Kit.

Bioinformatics Analysis

Triplicates of hLRH1 ChIP-Seq (WT hLRH-1) and a control (Input), were sequenced on the Illumina HiSeq 2000 platform using 50 bp, single-end reads. Reads were mapped to the hg19 human reference using bowtie and de-duplicated using Samtools. Final data compilation includes a total of 3.36, 10.77 and 7.7 million aligned sequence reads for WT hLRH-1 and 2.67 million reads from Input. Quality control and ChIP-signal strength assessment was performed via CHANCE (Diaz et al., 2012). CHANCE called both experiments as successful (via a comparison with the distribution of ChIP-strengths observed in the ENCODE repository) at a combined, positive false discovery rate (FDR) of FDR=2.1X10^-4 for WT hLRH1. Reads from replicates were then pooled, and peaks were called via MACS (Zhang et al., 2008), using the default parameter settings. This generated 18884 peaks from the wthLRH-1 samples. Genes
Phenotypic Screen Identifies TA As SUMO Inhibitor

up-regulated and down-regulated in response to TA was called by setting a probe-intensity threshold at the 95th percentile or 5th percentile, respectively, of array-wide probe intensities.

Motif searches were done by MEME-chip (Machanick and Bailey, 2011) and NR5A binding sequences were discovered by PROMO (Farre et al., 2003; Messeguer et al., 2002).

AAV8 Virus Generation and Retro-Orbital Infection Protocol

Expression of 3X Flag-tagged WT or 2KR hLRH-1 in mouse liver was achieved using adeno-associated virus serotype-8 (AAV8) by cloning each respective cDNA into the viral gateway vector pAAV2.1-TBG (Penn Vector Core) downstream of the liver-specific promoter thyroxine binding globulin (TBG). AAV8 virus expressing wild type hLRH-1, 2KR or enhanced green fluorescent protein (eGFP) (AAV8-hLRH1, AAV8-hLRH1-2KR or AAV8-eGFP) was amplified at the University of Pennsylvania Gene Therapy Vector Core. Tissue specificity and efficiency of infection was assessed in 8-week old C57BL/6 male mice (JAX labs) infected via retro-orbital injection with AAV8-eGFP at concentration of $1 \times 10^{11}$ gene copies per mL (GC/ml), this concentration was used for subsequent experiments using AAV8-hLRH-1 and AAV8-2KR. Tissues were collected 14 days post-infection, as previously described (Lu et al., 2012), and analyzed for eGFP fluorescence by fluorescent microscopy or extracted for mRNA as described below. Mice were euthanized in accordance with the UCSF Institutional Animal Care and Use Committee under Ingraham lab protocol. Mice were perfused with PBS prior to collection of liver tissue for all subsequent biochemical and gene expression studies.

Primary Hepatocyte Isolation

Primary hepatocytes were isolated from mice as previously described (Silver et al., 2000). Briefly, mice were anesthetized with AVERTIN (250 mg/kg) and perfused with pre-warmed perfusion buffer (HBSS supplemented with HEPES, pH 7.4) followed by perfusion with 75 mL digest buffer (HBSS/HEPES, pH 7.4, 0.03 mg/mL Collagenase Type 1, EDTA-free protease
Phenotypic Screen Identifies TA As SUMO Inhibitor

Digested liver was then dispersed on a 100 mm cell culture dish containing 25 mL cold DMEM (DMEM, 10% FBS, Penicillin/Streptomycin) and filtered through a 200 um-nylon membrane into a 50 ml-Falcon tube. Hepatocytes were isolated by mixing filtrate with 24 ml Percoll solution (1x HBSS, 4 mM NaHCO3, pH 2.2) and centrifuged at 100 x g for 7 min at 4 °C. Pellet containing hepatocytes was then washed with 40 mL DMEM and re-suspended in desired media volume. Cells were plated on 6-well plates coated with collagen-1 and were allowed to attach overnight. The following morning DMEM was replaced with fresh DMEM supplemented with TA at the indicated concentration.

Statistics

Data are represented as mean + SEM (or SD, as indicated): *p < 0.05; **p < 0.005; ***p < 0.001; ****p < 0.0001. Statistical analyses were performed using Prism 5 (GraphPad) software. Statistical significance was determined by unpaired Student’s T-test unless otherwise indicated.
ACKNOWLEDGEMENTS

We wish to thank Dr. S. Hand for valuable reagents. We would also like to acknowledge Drs. J. Ward, M. Asahino, B. Shoichet, A. Pierce, D. Silver, and R. Blind for experimental advice and discussion, as well as for critical reading of this manuscript. Funding sources that supported this work include an Innovation Grant from UCSF-PBBR/Roche grant, an R01-DK063592, and an R01-DK099722 and ADA 1-15-MI-08 to H.A.I., an NIDDK Supplemental Award DK063592-S1 to H.A.I to support K.A.R., T32HD726330 and AHA 14POST2013-0048 to D.A.M., the QB3-Malaysia Program to support K.K.H.A., and the Intramural Research Program at NIH, NIH National Cancer Institute, Center for Cancer Research to support J.S.S.

COMPETING FINANCIAL INTERESTS STATEMENT

Nothing to Disclose.
Table 1. List of Sumoylation Inhibitors Identified by Screen Type and Reported IC50 Values.

<table>
<thead>
<tr>
<th>Compound</th>
<th>Class</th>
<th>Screen</th>
<th>Library</th>
<th>Assay</th>
<th>Substrate</th>
<th>Target</th>
<th>IC50 (µM)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-D08</td>
<td>Flavonoid</td>
<td>Target</td>
<td>500 Flavones</td>
<td>IVS</td>
<td>AR Peptide</td>
<td>UBC9</td>
<td>6.0</td>
<td>(Kim et al., 2013)</td>
</tr>
<tr>
<td>Davidiin</td>
<td>Ellagitannin</td>
<td>Target</td>
<td>750 Extracts</td>
<td>In Situ</td>
<td>RanGap1</td>
<td>E1</td>
<td>0.15</td>
<td>(Takemoto et al., 2014)</td>
</tr>
<tr>
<td>Ginkgolic Acid</td>
<td>Alkylphenol</td>
<td>Target</td>
<td>500 Extracts</td>
<td>In Situ</td>
<td>RanGap1</td>
<td>E1</td>
<td>3.0</td>
<td>(Fukuda et al., 2009a)</td>
</tr>
<tr>
<td>GSK145A</td>
<td>Diamino-pyrimidine</td>
<td>Target</td>
<td>GSK Library</td>
<td>IVS</td>
<td>TRPS1 Peptide</td>
<td>UBC9</td>
<td>12.5</td>
<td>(Brandt et al., 2013)</td>
</tr>
<tr>
<td>Kerriamycin B</td>
<td>Antibiotic</td>
<td>Target</td>
<td>1800 Broths</td>
<td>In Situ</td>
<td>RanGap1</td>
<td>E1</td>
<td>11.7</td>
<td>(Fukuda et al., 2009b)</td>
</tr>
<tr>
<td>Spectomycin B1</td>
<td>Antibiotic</td>
<td>Target</td>
<td>Chemical Library</td>
<td>In Situ</td>
<td>RanGap1</td>
<td>UBC9</td>
<td>4.4</td>
<td>(Hirohama et al., 2013)</td>
</tr>
<tr>
<td>C#21</td>
<td>Phenyl Urea</td>
<td>Virtual</td>
<td>Maybridge</td>
<td>Docking</td>
<td>RanGap1</td>
<td>E1</td>
<td>14.4</td>
<td>(Kumar et al., 2013)</td>
</tr>
<tr>
<td>Tannic Acid</td>
<td>Gallotannin</td>
<td>Phenotypic</td>
<td>Pharmakon</td>
<td>qPCR</td>
<td>LRH-1</td>
<td>E1</td>
<td>12.8</td>
<td>(This Study)</td>
</tr>
</tbody>
</table>

Table 2. List of Top 6 hits from Primary Screen (Z-Scores).

<table>
<thead>
<tr>
<th>Drug</th>
<th>APOC3</th>
<th>MUC1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tannic Acid</td>
<td>4.87</td>
<td>4.31</td>
</tr>
<tr>
<td>Trifluridine</td>
<td>4.01</td>
<td>2.71</td>
</tr>
<tr>
<td>Taxol</td>
<td>2.93</td>
<td>-0.26</td>
</tr>
<tr>
<td>Vincristine</td>
<td>-3.61</td>
<td>-1.01</td>
</tr>
<tr>
<td>Colforsin</td>
<td>1.24</td>
<td>-5.70</td>
</tr>
<tr>
<td>Ouabain</td>
<td>-5.14</td>
<td>-0.38</td>
</tr>
</tbody>
</table>
FIGURE LEGENDS

Figure 1. Human LRH-1 is Efficiently Sumoylated in Cells and In Vivo

(A) Schematic of hLRH-1 protein (NR5A2 isoform 1) showing the location of major sumoylation sites at K192 and K270, and the minor K44 site (top panel). WT and SUMO-less forms of hLRH-1 (3KR and 2KR) expressed in JEG3 and HepG2 cells are indicated as detected with anti-Flag antibody. Unsumoylated (hLRH-1) as well as sumoylated hLRH-1 species (1X, 2X, and 3X) are indicated in bottom panel by arrows. Additional bands observed in HepG2 cells that persist after mutating K44, K192, and K270 are indicated with asterisk (*). Strategy used to humanize mouse liver for expression of wild type or SUMO-less (2KR) hLRH-1. (B) Sumoylated hLRH-1 species detected by anti-Flag in harvested livers after first infecting with AAV8-virus expressing eGFP, WT or 2KR. (C) Relative transcripts levels of hLRH-1 transcripts in mouse livers infected with AAV8-virus expressing eGFP, wild type hLRH-1 (WT) or SUMO-less hLRH-1 (2KR). (D) Staining for tagged-hLRH-1 as detected by immunofluorescence using anti-FLAG (white arrows). Hepatocytes are prepared as described in Materials and Methods from harvested, perfused livers 2 weeks following retro-orbital viral-mediated infection.

Figure 1 - figure supplement 1. Mutating individual acceptor lysines in hLRH-1 establishes the importance of K192 and K270 in SUMO modification of hLRH-1. Schematic of hLRH-1 with position of 3X-Flag epitope tag and three sumoylation sites in human LRH-1 (upper panel). Sumoylation pattern in HepG2 cells expressing WT or mutant forms (K192R, K270R, or 2KR) of Flag-Tagged hLRH-1. * = non-specific bands observed in all mutants forms of hLRH-1 in HepG2 cells (lower panel).

Figure 1 - figure supplement 2. Human LRH-1 transcripts and protein are expressed in liver after AAV8-TBG viral infection. a. Relative expression of hLRH-1 or endogenous mLrh-1 transcripts in mouse liver 14 days post-infection with either AAV8-eGFP (GFP) or AAV8-
Phenotypic Screen Identifies TA As SUMO Inhibitor

hLRH-1 (hLRH-1) at MOI of 1 x 10^{11} (genome copies/ml). **b.** Human LRH-1 detected by anti-FLAG antibody in heart and liver tissue collected from mice expressing Flag-hLRH-1 (hLRH-1) or eGFP as described in Materials and Methods. Human flag-tagged LRH-1 protein expressed in HepG2 cells is indicated with arrow (far left lane) and loading controls for each sample (Gapdh).

**Figure 2. A Phenotypic Screen Identifies TA as A Small-Molecule Sumoylation Modulator.**

(A) Schematic outline of the primary screen using JEG3 cells expressing wild type hLRH-1 and downstream filtering steps to identify small molecules that modulate sumoylation. Individual amplification profiles for MUC1 transcripts are shown for each drug treatment using the Pharmakon 1600 library (upper right panel). Highlighted in red is the amplification curve of MUC1 obtained with tannic acid, the top hit from the primary screen. (B) Relative levels of SUMO-sensitive transcripts APOC3, MUC1 and the SUMO-insensitive transcript CYP11A1 in JEG3 cells expressing wild type hLRH-1 or SUMO-less hLRH-1 (3KR) (top panel). Relative levels of transcripts as above are shown after 72 h siControl (Con) or siUBC9 (UBC9) treatment in JEG3 cells expressing hLRH-1. Results represent values obtained for triplicate samples. Statistical significance: ****p<0.0001, **p<0.01, *p<0.05. (C) Scatter plot from the primary screen showing normalized Z-scores for APOC3 and MUC1, calculated as described in Materials and Methods. All compounds yielding Z-scores greater than +2 or less than -2 are shown within red dashed boxes. Positive Z-scores correspond to increased expression of transcripts relative to the control housekeeping gene, TBP. The Z-score obtained for tannic acid (TA) is indicated as red dot.

**Figure 3. TA Enhances SUMO-sensitive Gene Expression in Cells.**

(A) Relative expression of CYP11A1, APOC3, and MUC1 in JEG3 cells expressing human wild type LRH-1 following 24 h treatment with increasing TA. (B) Relative expression of CYP11A1,
APOC3, and MUC1 in JEG3 cells transiently expressing wild type (WT) or SUMO-less hLRH-1 (3KR) following 6 h treatment with increasing concentrations of TA as indicated. Vehicle DMSO control is shown (-). (C) Levels of unsumoylated (hLRH-1) and sumoylated hLRH-1 species in JEG3 cells detected with anti-Flag following 5 h treatment with increasing TA concentrations (1-50 µM). Shorter exposure of 1x-sumoylated hLRH-1 is shown in bottom panel (1x). Vehicle control (DMSO) is shown (-). (D) Levels of unsumoylated (hLRH-1) and sumoylated hLRH-1 species in JEG3 cells detected with anti-Flag after tannic acid (TA), trifluidine (Tri.), 2D08, and ginkgolic acid (GA) treatment (30 µM each) following 5 h treatment. Shorter exposure of unsumoylated hLRH-1 (hLRH-1) is shown in bottom panel. (E) Cell viability is shown for JEG3 cells expressing wild type hLRH-1 following 5 h treatment with increasing ginkgolic acid (GA) or tannic acid (TA). Statistical significance: ****p<0.0001, *p<0.05.

Figure 3 - figure supplement 1. Upregulation of APOC3 by TA in JEG3 cells depends on presence of hLRH-1. Relative levels of endogenous APOC3 in JEG3 cell lines expressing either a control vector (iEV) or human wild type LRH-1 (WT) after treatment with no (-) or increasing concentrations of TA for 6 h.

Figure 3 - figure supplement 2. Significant cell toxicity in JEG3 wtLRH-1 cells after 24 h treatment with GA but not TA. Percent cell viability in JEG3 cells expressing wild type hLRH-1 (JEG3 + hLRH-1) following 24 h treatment with increasing concentrations of either TA (red) or GA (blue).

Figure 4. TA is a Detergent-Resistant Inhibitor of Substrate Sumoylation In Vitro.
(A) In vitro sumoylation (IVS) of recombinant full length (FL)-hLRH-1, without ATP, with ATP or with ATP and recombinant SENP1 added to IVS reactions as described in Materials and Methods. (B) IVS assays with increasing tannic acid (TA), trifluidine (Tri), 2-D08 and ginkgolic
acid (GA). Sumoylated and unsumoylated FL-hLRH-1 are indicated with arrows as detected with anti-LRH-1 antibody. (C) IC\textsubscript{50} of TA in FL-hLRH-1 IVS assay. Data are represented as mean ± SEM from at least three independent replicates. (D) Formation of E1 thioester with or without TA, in non-reducing conditions (-DTT). Effects of TA (10 µM) on formation of SUMO-E1 complex (SUMO-SAE1, top band) compared to reducing conditions without TA (-DTT, last lane). SUMO1 dimers are formed in non-reducing conditions (SUMO1 Dimer). E1 thioester formation assays are initiated by addition of freshly prepared ATP (10 mM) and described in Materials and Methods. Anti-SUMO1 antibody was used to detect SUMO1 species. (E) Levels of sumoylated and unsumoylated FL-LRH-1 in IVS assay with TA (15 and 30 µM) and in the presence or absence of Triton X-100 in vitro (left panel). Bar graph of quantified data showing percent inhibition of hLRH-1 sumoylation by TA and with or without Triton X-100 (right panel).

**Figure 4 - figure supplement 1. Effects of TA, other candidate hits, and published sumoylation inhibitors in an IVS assay of full-length hLRH1.**

a. Dose-dependent inhibition of full-length (FL)-hLRH-1 by TA compared to other top candidate hits from primary screen as assayed by IVS. IVS assay and immunoblotting conditions used to detect hLRH-1 species are described in Materials and Methods. Sumoylated hLRH-1 (1x, 2x, 3x) and unmodified LRH-1 (LRH-1) species are indicated by arrows. b. IVS assays of FL-hLRH-1 were performed with increasing concentrations of TA are shown (left panel) and plotted as normalized values in graph (right panel). Effects of two other published sumoylation inhibitors, 2-D08 and GA are also shown in graph. IVS data was normalized to DMSO control for each compound, and then plotted as percent conversion per log\textsubscript{10} [µM] concentration. Curve fitting of data is described in Materials and Methods.

**Figure 4 – figure supplement 2. IVS of Multiple Substrates inhibited by TA.**

a. Coomassie staining of recombinant hinge-LBD mSF-1 protein (aa178-462) with DMSO (0) or with 10 µM TA.
This protein fragment of mSF-1 contains only one of the two conserved sumoylation consensus
sites at K194. Migration of SUMO1-SF-1 and unmodified SF-1 hinge-LBD (SF-1) are indicated
by arrows.  b. IVS of full length IκBα without (0) or with increasing concentrations of TA, as
indicated, with SUMO-IκBα and unmodified IκBα indicated by arrows.  c. Dose dependent
inhibition of IVS of fluorescently labeled AR peptide by TA with the IC_{50} provided. Data are
plotted as percent conversion versus TA concentration (log_{10} [µM], left panel). Real-time
sumoylation of AR peptide (% Conversion) are plotted for at different concentrations of TA (right
panel). IVS conditions and detection of AR peptide sumoylation by electrophoretic mobility shift
assay are previously described in Kim YS, Nagy K, Keyser S, Schneekloth JS, Jr. Chem Biol 20,
604-613 (2013).

Figure 5. TA Inhibits Exogenous and Endogenous NR5A Sumoylation, As Well As
General Sumoylation in Cells.

(A) Levels of unsumoylated (LRH-1) and sumoylated hLRH-1 species in HepG2 cells detected
with anti-Flag following TA, 2D08, or GA at specified concentrations after 24 h treatment.
Vehicle control (DMSO) is shown (-).  (B) Cell viability for HEPG2 cells expressing wild type
hLRH-1 following 24 h treatment with increasing concentrations of GA or TA.  (C) Sumoylation of
wild type hLRH-1 expressed in HepG2 cells and detected with anti-FLAG after siCont or siUBC9
knockdown for 72 h (+), and with increasing TA (6 h). Shorter exposure of 1x-sumoylated
hLRH-1 is shown in panel below (1x), as well as loading control (GAPDH). Vehicle control
(DMSO) is shown (-).  (D) Flag-tagged hLRH-1 protein immunoprecipitated by anti-Flag in
HepG2 cells treated with vehicle or TA (30 or 50 µM) followed by immunoblot with either anti-
SUMO1 (top panel) or anti-SUMO2 (bottom panel). Arrows indicated migration of 1X, 2X, and
3X sumoylated hLRH-1 species.  (E) Levels of total sumoylated or ubiquitinated proteins in
HepG2 cells following siControl and siUBC9 (72 hr) or TA treatment (6 h), as detected by anti-
SUMO1, -SUMO2 or -ubiquitin. (F) Effects of TA (6 h) on endogenous SF-1 sumoylation in H295R cells and detected anti-SF-1 antibody.

**Figure 5 - figure supplement 1. UBC9 transcripts and protein levels following siUBC9 knockdown in HepG2-hLRH-1 cells.** a. Relative expression of UBC9 (left panel) and hLRH-1 (right panel) in HepG2 cells following 72 h exposure to siControl (siCont) or siUBC9 (siUbc9) as described in Materials and Methods. b. Immunoblot of UBC9 protein following 72 h exposure to siCont or siUbc9 with percentage decrease of UBC9 protein shown in bar graph (right panel).

**Figure 5 - figure supplement 2. TA attenuates endogenous RanGap sumoylation.** Endogenous sumoylated (Su-RanGap 1x) and unmodified RanGap (RanGap) in HepG2-hLRH-1 cells after siCont, siUBC9, or after TA treatment (6 h) with concentrations indicated. Note that the α-RanGap rabbit monoclonal antibody from GeneTex recognizes both 1x Sumoylated RanGap and unmodified RanGap.

**Figure 6. TA Increases Expression and Promotes hLRH-1 Occupancy on Target Genes.** (A) Venn diagram representing the overlap between transcripts changed by induction of hLRH-1 (+Dox) (Blue) and hLRH-1+TA (+Dox, +TA 30 µM) (Yellow), as well as hLRH-1 binding sites identified by ChIP-Seq in HepG2 cells (+Dox, Orange). Heat map of top three genes from overlapping set of 42 genes: PFKFB3, SERPINE1 (PAI1) and CYP24A1 showing changes after induction of hLRH-1 (Dox/EtOH) and then after TA treatment (TA/DMSO). (B) ChIP-Seq binding profiles of the three hLRH-1 targets from panel A. Representative views for ChIP-Seq peaks called by MACS are shown along with genomic location and consensus sequence of putative LRH-1 binding sites (red text). (C) Relative expression of three hLRH-1 targets in HepG2 cells from (A and B) before (-Dox) and after induction of hLRH-1 (+Dox) and following treatment with TA for 6 h with. (D) ChIP-qPCR results in HepG2 cells expressing hLRH-1 for regions identified...
Phenotypic Screen Identifies TA As SUMO Inhibitor

Figure 6 - figure supplement 1. Quantification of transcriptional changes in HepG2 cells after TA or siUBC9 treatment and in presence of hLRH-1.

a. Venn diagram representing overlap between transcriptionally responsive genes by siUBC9 (Blue) or TA (Red); data analyses are described in Materials and Methods. b. Motif search of HepG2 WT hLRH-1 binding sites as described in Materials and Methods yields a consensus sequence with the frequency distribution quantified as bits and the E-value (probability). The 8 bp consensus site obtained after Motif search is shown after induction of WT hLRH-1 (+Dox).

Figure 7. TA Mimics SUMO-less hLRH-1 When Expressed In Humanized Mouse Primary Hepatocytes.

(A) Relative expression of hLRH-1 in mouse livers two weeks post-infection with MOI of virus indicated. Hepatic expression of endogenous mouse mLRH-1 (red arrow) in control mLRH-1\(^{ff}\) mice, in LKO mice (mLRH-1\(^{ff};Alb-Cre\)), or in mLRH-1\(^{ff}\) mice after 2 weeks virus infection with AAV8-eGFP (eGFP), AAV8-WT-hLRH-1 (WT) or AAV8-2KR-hLRH-1 (2KR) (black arrow).

Endogenous mLRH-1 and exogenous hLRH-1 are detected with an anti-LRH-1 antibody. Non-specific bands detected with the anti-LRH-1 antibody obscure sumoylated LRH-1 species. (B) Relative expression in mouse liver of the classic LRH-1 target, Cyp8b1 or SUMO-sensitive LRH-1 targets Adipoq and Shh (and its downstream target Gli2) following infection with either AAV8-eGFP, AAV8-hLRH-1 or a SUMO-less AAV8-2KR-hLRH-1. Each bar represents value obtained from three livers. Values below the threshold of detection (qPCR >40 cycles) are indicated as ND. (C) Sumoylation pattern of AAV8-hLRH-1 (WT) and AAV8-2KR hLRH-1 (2KR) in cultured mouse primary hepatocytes (left panel). Sumoylation of hLRH-1 in infected primary mouse hepatocytes treated with increasing TA for 5 h (right panel). Sumoylated species and
unsumoylated Flag-hLRH-1 are indicated by arrows and detected by anti-Flag antibody. (D)

Relative expression of genes shown in Panel B measured in uninfected primary mouse hepatocytes treated for 5 h with increasing concentrations of TA. Statistical significance for panels B and D: **p<0.01, *p<0.01.

Figure 7 - figure supplement 1. No cellular toxicity in primary hepatocytes by TA. Cell viability (%) is shown for primary hepatocytes treated with increasing TA concentrations as indicated for 24 h.

Legends for Supplementary Files and Source data

Supplementary File 1. List of forward and reverse qPCR primers used in our study.

Supplementary File 2. Table providing details and work flow of the primary and follow-up secondary assays.

Supplementary File 3. List of vendor for repurchasing candidate drug hits used in secondary filtering step and in immortalized cells, primary hepatocytes, and IVS assays.

Source data 1. Z-scores from primary screen are listed for each drug tested (10 µM, 24 h) after assaying for APOC3 or MUC1 transcripts in JEG3 cells stably expressing hLRH-1. Primary screening conditions and calculations for obtaining Z-scores are provided in the Material and Methods as well as in Supplementary File 2.

Source data 2. List of genes obtained after profiling (Columns 1-3) and after ChIP-seq (Column 4) in HepG2-hLRH-1 cells; data are also represented in Venn diagrams in either Figure 6 or Figure 6 - figure supplement 1. Experimental conditions are summarized in each column header. All listed genes in profiling experiments (Columns 1-3) were changed up or down by 2-fold (log2 ≥1.0 or ≥-1.0) after normalization of data and statistical significance were determined by comparing datasets, as described in Materials and Methods.
REFERENCES


Phenotypic Screen Identifies TA As SUMO Inhibitor


Phenotypic Screen Identifies TA As SUMO Inhibitor


Phenotypic Screen Identifies TA As SUMO Inhibitor


Phenotypic Screen Identifies TA As SUMO Inhibitor


Suzawa, Ramos et al., Figure 1

Infect with AAV8-TGB Flag-WT or 2KR hLRH-1

Infected Liver Expressing hLRH-1

Infect with AAV8-TGB Flag-WT or 2KR hLRH-1

Infected Liver Expressing hLRH-1

Suzawa, Ramos et al., Figure 1
JEG3 hLRH-1 Cells

A

PHARMAKON

DRUG LIBRARY

Filtering Assays

Cell-Based
Sumoylation
& Expression

In Vivo
Sumoylation
& Expression

B

Sumo-Sensitive

APOC3
MUC1
CYP11A1

Relative Expression

WT 3KR
WT 3KR
WT 3KR

**

C

Gene-Expression Screen

APOC3 (Z-score)

MUC1 (Z-score)

TA

Suzawa, Ramos et al., Figure 2
JEG3 hLRH-1 Cells

A

CYP11A1
APOC3
MUC1

Relative Expression

TA [μM] - 24 h

B

CYP11A1
APOC3
MUC1

Relative Expression

TA [μM] - 24 h

C

αFlag

TA (μM) - 5 h

D

Drug (30 μM) - 5 h

E

Viable Cells (%)

Compound [μM] - 5 h

Suzawa, Ramos et al., Figure 3
Suzawa, Ramos et al., Figure 4
A. HepG2 hLRH-1 Cells

IB: αFlag

LRH-1

IB: αSumo1

IB: αSumo2

IB: αUb

IB: αSF-1

H295R Cells

Suzawa, Ramos et al., Figure 5
HepG2 hLRH-1 Cells

A

B

C

D

Suzawa, Ramos et al., Figure 6
Cyp8b1  Adipoq  Shh  Gli2

**Infected AAV8 Virus (1x10^{11})**

**Uninfected Hepatocytes + TA [μM] - 5 h**

**Suzawa, Ramos et al., Figure 7**