In-depth Proteomic Characterization of Endogenous Nuclear Receptors in Mouse Liver

Qiongming Liu‡§, Chen Ding‡¶¶, Wanlin Liu‡§, Lei Song‡§, Mingwei Liu‡§, Liang Qi‡§, Tianyi Fu‡§, Anna Malovannaya¶¶, Yi Wang¶¶, Jun Qin‡§¶¶, and Bei Zhen‡§¶¶

Nuclear receptors (NRs) are a superfamily of transcription factors that, upon binding to ligands, bind specific DNA sequences and regulate a transcriptional program governing cell proliferation, differentiation, and metabolism. In the liver, by sensing lipid-soluble hormones and dietary lipids and governing the expression of key liver metabolic genes, NR proteins direct a large array of key hepatic functions that include lipid and glucose metabolism, bile secretion, and bile acid homeostasis. Although much has been learned about the physiology of NRs, little is known about their protein expression and DNA binding activity in the liver because of their low abundance and the lack of high-throughput methods for detection at the protein level. Here we report a method for profiling the DNA binding activity of the NR transcription factor superfamily in mouse liver. We use DNA constructs of hormone response elements (HREs) as affinity reagents to enrich NR proteins from nuclear extracts of mouse liver and then identify them using mass spectrometry. We evaluated 20 DNA constructs containing various combinations of HREs for their ability to enrich endogenous NR proteins and found that two different HREs are sufficient to achieve isolation and identification of nearly all endogenous NR proteins from one mouse liver. We have detected proteins for 35 members of the NR family out of 41 that are expressed in mouse liver at mRNA level. Thus, this method allows coverage of most of the whole NR proteome and establishes a practical assay for the investigation of NR actions in mouse liver. We anticipate that this method will find widespread use in future investigations of NR actions in liver biology and pathology. Furthermore, this workflow is a useful tool for NR biologists interested in measuring NR expression, DNA binding, post-translational modifications, cellular localization, and other functional aspects of NRs in organs under normal physiological and pathological conditions, as well as during pharmacological intervention. Molecular & Cellular Proteomics 12: 10.1074/mcp.M112.022319, 473–484, 2013.

The nuclear receptor (NR)1 superfamily represents a highly conserved, widely expressed group of ligand-regulated transcription factors found in metazoaos (1, 2). There are a total of 48 designated NRs in human and 49 NRs in mice that include classical receptors, adopted orphan receptors, and orphan receptors (3). Nearly all NRs contain two well-defined structural domains: an N-terminal or central DNA-binding domain and a ligand-binding domain positioned along the C-terminus (4). NRs respond to environmental and hormonal stimulation by binding a diverse set of lipophilic substances such as fat-soluble hormones, vitamins, and dietary lipids (5, 6). Upon activation by their ligands or ligand-independent activation through phosphorylation and other post-translational modifications, these receptors bind to DNA and recruit other co-regulator proteins to modulate the expression of specific genes that control development, homeostasis, reproduction, and metabolism of the organism. Consequently, NRs play important roles in normal metabolic and homeostatic processes, as well as in diseases such as cancer, inflammation, diabetes, obesity, cholestasis, and atherosclerosis (7–9). Furthermore, NRs are unique in their ligand dependence, which makes them alluring as pharmaceutical targets. The therapeutic impact of NR ligands is evident in treatments for a wide spectrum of diseases, ranging from diabetes (thiazolidenediones for PPARγ (Nr1c3)) and cancer (tamoxifen for ER-alpha (Esr1)) to heart disease (fenofibrates for PPARα (Nr1c1)) (10). Approximately 13% of the drugs approved by the U.S. Food and Drug Administration target NR proteins (11). Thus, NR proteins are now—and will continue to be—targets of intense research, with most of the efforts directed toward understanding NR functions, improving current NR drugs,

1 The abbreviations used are: CHS, consensus half-site; DR, direct repeat; HRE, hormone response element; iBAQ, intensity based absolute quantification; IR, inverted repeat; NE, nuclear extract; NR, nuclear receptor; QPCR, quantitative RT-PCR; PSM, Peptide-Spectrum Match.
and screening currently unexploited NRs for pharmaceutical aims.

Because the expression pattern of NR genes in different tissues may suggest their relative functional importance in different tissues, expression profiling of NRs has greatly facilitated our understanding of the physiological functions of these receptors (12–18). The expression of the entire NR transcriptome was characterized by Mangelsdorf and colleagues; in their work, all 49 mouse NR mRNAs were measured via quantitative RT-PCR (QPCR) in 39 tissues (17). The resulting dataset uncovered several NR clades whose patterns of expression indicate their ability to coordinate the transcriptional programs necessary to affect distinct physiological pathways. That study also revealed the existence of a hierarchical transcriptional network tying NR functions to reproduction, development, metabolism, and energy homeostasis (17). Importantly, surveying the diurnal expression profiles of all 49 mouse NRs in four metabolically relevant tissues revealed that of the 45 NRs expressed, 25 are in a rhythmic cycle and 3 exhibit a single transient pulse expression (18). The rhythmicity of NR expression offers a logical explanation for the known cyclical behavior of lipid and glucose metabolism and suggests that NRs play important roles in orchestrating the peripheral circadian clock to provide divergent metabolic outputs (18).

Despite the importance of transcription profiling, it serves only as a surrogate for the expression profile of NR proteins. As proteins directly facilitate biological functions in the cell, the protein levels of NRs and, better yet, their DNA binding and transcriptional activity are more directly correlated with the NR functions. This information is currently unavailable, in part because of a lack of suitable proteomic assays with which to survey the whole family of NRs. Unlike RNA/DNA, which can be amplified through PCR-based techniques to create samples for profiling, proteins cannot be amplified. Thus, proteins of low abundance, such as NRs, remain a major challenge to measure. Although antibodies are undoubtedly the best affinity reagents for detecting proteins (19, 20), the generation of usable antibodies for detecting endogenous proteins is a process of trial and error. In fact, generating reliable antibodies against the whole collection of NR proteins has proven to be difficult. Therefore, a method for the detection of the entire family of endogenous NR proteins is greatly needed.

NRs as transcription factors regulate transcription by binding to hormone response elements (HREs) localized in the physiological functions of these receptors (12–18). The expression of the entire NR transcriptome was characterized by Mangelsdorf and colleagues; in their work, all 49 mouse NR mRNAs were measured via quantitative RT-PCR (QPCR) in 39 tissues (17). The resulting dataset uncovered several NR clades whose patterns of expression indicate their ability to coordinate the transcriptional programs necessary to affect distinct physiological pathways. That study also revealed the existence of a hierarchical transcriptional network tying NR functions to reproduction, development, metabolism, and energy homeostasis (17). Importantly, surveying the diurnal expression profiles of all 49 mouse NRs in four metabolically relevant tissues revealed that of the 45 NRs expressed, 25 are in a rhythmic cycle and 3 exhibit a single transient pulse expression (18). The rhythmicity of NR expression offers a logical explanation for the known cyclical behavior of lipid and glucose metabolism and suggests that NRs play important roles in orchestrating the peripheral circadian clock to provide divergent metabolic outputs (18).

Despite the importance of transcription profiling, it serves only as a surrogate for the expression profile of NR proteins. As proteins directly facilitate biological functions in the cell, the protein levels of NRs and, better yet, their DNA binding and transcriptional activity are more directly correlated with the NR functions. This information is currently unavailable, in part because of a lack of suitable proteomic assays with which to survey the whole family of NRs. Unlike RNA/DNA, which can be amplified through PCR-based techniques to create samples for profiling, proteins cannot be amplified. Thus, proteins of low abundance, such as NRs, remain a major challenge to measure. Although antibodies are undoubtedly the best affinity reagents for detecting proteins (19, 20), the generation of usable antibodies for detecting endogenous proteins is a process of trial and error. In fact, generating reliable antibodies against the whole collection of NR proteins has proven to be difficult. Therefore, a method for the detection of the entire family of endogenous NR proteins is greatly needed.

NRs as transcription factors regulate transcription by binding to hormone response elements (HREs) localized in the regulatory region of target genes (21, 22). Typically, the HREs are composed of two hexameric “consensus half-sites” (CHSs), often with the 5′-AGGTCA-3′ sequence as the core recognition motif (23). Although some monomeric receptors can bind to a single half-site, most receptors bind as homodimers or heterodimers to binary HREs (Fig. 1A). For these dimeric motifs, the half-sites can be conjoined as direct repeats (DRs) or inverted repeats (IRs) (24). The number of spacer nucleotides between the half-sites varies from 1 to 5, referred to as the “1–5 rule ” for DR1 to DR5 and IR1 to IR5 (25–28).

HREs have been used as affinity reagents to isolate NR proteins and associated protein complexes in the past. Ivan Nalvarite and colleagues (29) attempted to isolate the ERα interactors on the estrogen response elements. This group used recombinant ERα and MCF-7 cell nuclear extracts to assemble ERα complexes and followed with mass spectrometry to identify the interacting proteins. Here, we report a method for profiling DNA binding activity for the whole panel of endogenous NR proteins. In other words, we developed a “pan-NR” assay for the detection of all NR superfamily members in a single experiment. Because our ultimate goal is to be able to survey NR proteins in tissues, we used mouse liver to show the practicality of this approach. We first adapted a cell culture nuclear extraction protocol for mouse liver tissue and fitted the DNA pull-down protocol for optimal NR protein binding to HREs. We then used this assay to successfully detect 35 different (85% of all expressed) DNA-bound, or “active,” NRs in mouse liver, providing a path for the elucidation of NR actions in mouse liver and other tissues.

**EXPERIMENTAL PROCEDURES**

**Materials and Reagents—Single-stranded DNA sequences and biotinylated primers were synthesized by Generay Biotech Co. (Shanghai, China). XhoI, Sall, alkaline phosphatase, T4 polynucleotide kinase, and a DNA ligation kit were purchased from Takara Bio (Dalian, China). 2× HiFiTaq PCR StarMix was purchased from GenStar (Beijing, China). A TIANgel Midi purification kit and a TIANprep Mini plasmid kit were purchased from TIANGEN (Beijing, China). Dynabeads M-280 Streptavidin was purchased from Invitrogen (Carlsbad, CA). Sequencing grade modified trypsin was purchased from Promega (Madison, WI). The following buffers were prepared: hypotonic buffer (10 mM Tris pH 7.3, 1.5 mM MgCl2, 10 mM KCl), low salt buffer (20 mM Tris pH 7.3, 1.5 mM MgCl2, 20 mM KCl, 0.2 mM EDTA, 25% glycerol), high salt buffer (20 mM Tris pH 7.3, 1.5 mM MgCl2, 1.2 mM KCl, 0.2 mM EDTA, 25% glycerol), BC-150 (20 mM Tris pH 7.3, 150 mM KCl, 0.2 mM EDTA, 20% glycerol), NETN buffer (20 mM Tris pH 8.0, 50 mM NaCl, 1 mM EDTA, 0.5% Nonidet P-40), and DNA binding buffer (5 mM Tris pH 7.5, 0.5 mM EDTA, 1 mM NaCl). All buffers for protein extraction were supplemented with 10 mM β-mercaptoethanol and 1 mM phe- nylmethyl sulfonylfluoride immediately before use.

**Preparation of HRE-containing DNA Constructs—** Sequences of the single-stranded DNAs used in this study are listed in supplemental Table S1. Double-stranded DNA oligos carrying three repeats of HREs were prepared by annealing complementary oligonucleotides into a 10-mmol Tris-EDTA buffer (pH 8.0) containing 100 mM NaCl. Each double-stranded DNA sequence was designed to contain sticky ends with 5′-end (Xhol) and 3′-end (SalI) restriction sites to enable self-ligation. In order to make the pET24a-based DNA baits carrying multiple tandems of XH HRE repeats (6X, 9X, and so on), 5′-phosphates were added to HRE oligonucleotides by T4 polynucleotide kinase. The pET24a+ vector was treated with SalI and Xhol followed by dephosphorylation with alkaline phosphatase to prevent self-ligation. In vitro ligation reaction was performed with >8 parts oligo DNA to 1 part linear vector. The sequence and copy number of HRE were confirmed by sequencing. HRE DNA baits were amplified from these constructs using 2× HiFiTaq PCR StarMix with biotinylated forward and reverse primers as follows: BGPf- 5′-biotin-gtcagcggccgcagcatgagc-3′ and BGPr-5′-biotin-gcggcgcctaacccaccc-3′. These primers anneal to...
vector 200nt upstream of Sall (forward primer) and 200nt downstream of Xhol site (reverse primer), creating extended arms on both sides of the inserted HRE. The PCR products were electrophoresed in 1% agarose gel, and the target biotin-labeled double-stranded DNA was excised from the gel and purified using a TIANgel Midi purification kit.

**Isolation of Mouse Hepatocytes**—C57BL/6J mice 6 to 8 weeks old were purchased from HFK Biotech Co. (Beijing, China). Mice were maintained at 21 °C to 24 °C with a 12 h/12 h light/dark cycle (lights on 06:00–18:00) and were provided water and standard rodent chow ad libitum. All animal care and use procedures were in accordance with institutional Animal Care and Use Committee guidelines. Hepatocytes were prepared from mouse livers as previously described by Li and colleagues, with few minor modifications (30). Mice were anesthetized via intraperitoneal injection of pentobarbital. Liver tissues were perfused for 5 min with pre-perfusion solution (8g/l NaCl, 400 mg/l KCl, 88.17 mg/l NaH₂PO₄, 2H₂O, 120.45 mg/l Na₂HPO₄, 2.38g/l HEPES, 350 mg/l NaHCO₃, 190 mg/l EGTA, 900 mg/l glucose, pH 7.3, and 0.75 U/ml heparin added before use). Once the liver turned a khaki color, it was perfused with 0.05% collagenase in DMEM for 10 min at 37 °C. After perfusion, liver cells were dispersed in DMEM mechanically with forceps and filtered through a 40-μm nylon filter. The cell suspension was centrifuged at 1000 × g for 5 min at room temperature, and the cell pellet was washed and collected via centrifugation.

**Preparation of Nuclear Extract from Mouse Liver Hepatocytes**—The collected hepatocytes were suspended in 10X volumes of hypotonic buffer for 10 min on ice. The sample was then centrifuged at 1000 × g for 15 min at 4 °C, and the resulting pellet was dounced 15 times to break cell membranes. Homogenate was centrifuged at 4000 × g to separate the nuclear and non-nuclear fractions. The nuclear pellet was re-suspended with 0.5 volumes of low salt buffer and dounced an additional 10 times to re-suspend the nuclei. Another half volume of high salt buffer was added dropwise while the solution was stirred to extract nuclear protein. The solution was stirred at 4 °C for an additional 30 min after high salt addition and then centrifuged at 25,000 × g for 20 min at 4 °C. The supernatant containing nuclear extract (NE) was dialyzed against BC-150 for 1 h at 4 °C. The resulting NE was aliquoted, snap-frozen in liquid nitrogen, and stored at -80 °C.

**HRE Pull-down and Trypsin Digestion**—Biotinylated DNA was immobilized on Dynabeads with DNA binding buffer and washed with BC-150 buffer once. NE was spun down at 100,000 × g in a Beckman Optima ultracentrifuge (TLA100 rotor) for 20 min at 4 °C. The supernatant was transferred to a clean 1.5-mL tube and kept on ice. The concentration of the supernatant was measured using the Bradford method. Before being mixed with HRE-Dynabeads, the NE was adjusted to between 200 and 250 μM total salt with NaCl and 1 mM final EDTA. HRE-Dynabeads were incubated with NE with rotation for 2 h at 4 °C. The supernatant was discarded and the beads were washed twice with NETN and then three times in PBS. In in-solution digest, the beads were re-suspended with 45 μl of 50 mM NH₄HCO₃ and digested overnight with trypsin at an enzyme-to-protein ratio of 1:100. Peptides were extracted with 200 μl of 50% acetonitrile/0.1% formic acid. For isoelectric focusing (IEF) samples, tryptic peptides were separated on pH 3–10 IPG strips (GE Healthcare) using a 3100 OFFGEL Fractionator (Agilent Technologies, USA) according to the manufacturer’s instructions. For SDS-PAGE samples, HRE pull-down beads were re-suspended with 20 μl of SDS sample buffer and boiled for 5 min. The samples were then loaded on 10-cm 10% SDS-PAGE and run to 1/3 length. The gel was minimally stained with Coomassie Brilliant Blue and briefly washed in 5% ethanol/10% acetic acid solution. It was sliced into six bands to separate the protein content by molecular weight ranges. The gel slices were destained with 500 μl of 40% methanol and 50 mM ammonium bicarbonate, followed by dehydration with 200 μl of 75% HPLC-grade acetonitrile. After removal of the acetonitrile, the gel slices were rehydrated with HPLC-grade water and 50 mM ammonium bicarbonate and crushed. The gel slices were then incubated in 20 μl of 50 mM ammonium bicarbonate containing 1 μg/ml trypsin overnight (~6 h) at 37 °C. The peptides were extracted with 200 μl acetonitrile by being vortexed vigorously for 5 min; gel pieces were re-suspended in 30 μl of 0.1% formic acid for 5 min and re-extracted with 200 μl acetonitrile again. The peptide solutions were combined and dried completely in a vacuum concentrator (SpeedVac, Thermo Scientific, Germany).

**Nano-LC/MS Analysis**—Tryptic peptides were dissolved with loading buffer (5% methanol, 0.1% formic acid) and then separated on an on-line C18 column (75 μm inner diameter, 360 μm outer diameter, 10 cm, 3 μm C18). Mobile phase A consisted of 0.1% formic acid in water solution, and mobile phase B consisted of 0.1% formic acid in acetonitrile solution; a linear gradient from 3% to 100% B over a 75-min period at a flow rate of 350 nl/min was applied. For identification, peptides were fragmented via collision-induced dissociation and analyzed using an LTQ-Orbitrap Velos (Thermo Scientific, Germany). The survey scan was limited to 375–1600 m/z. Proteins were identified using the Proteome Discoverer 1.3 with the MASCOT search engine and the mouse reference sequence protein database from NCBI (downloaded November 2011). A total of 34,297 entries were actually searched. The threshold score/expected value for accepting individual spectra was set to an ion score of 10. The PSM false positive rate was set to 1% strict/5% relaxed cutoff. The mass tolerance was set to 20 ppm for precursors and 0.5 Da for product ions. Dynamic modifications of oxidation (Met), acetylation (protein N terminus), phosphorylation (ST), and Destreak (C) were chosen. The maximum number of missed cleavage sites was set at 2. The experimental m/z and retention times were recorded for precursor area quantification. The Xcalibur Qual Browser (v. 1.0.042) was used to produce the peak list from raw data. Peptide amounts and protein quantities were estimated based on the precursor area under the curve (AUC), which was calculated by Proteome Discoverer 1.3 using a precursor ion area detector with 20 ppm mass precision settings. The quantities of identified proteins were estimated by the intensity-based absolute quantification (IBAQ) method as described in Ref. 31. Briefly, the sum of peptide peak areas (or AUCs) of each protein was normalized by its theoretical number of all possible full tryptic peptides. For peptides that are shared by non-unique gene products (GP), we assumed that all GPs were present and divided the peak area evenly to all the possible GPs. The normalized AUC was referred to as the iBAQ and, for better visual comprehension, was displayed in 105 units. For analytical and biological reliability, we carried out many technical and biological replicates (a total of more than 180 HRE pull-down experiments from different mice were done). We reported NRs that were observed consistently.

**RESULTS**

**Preparation of HRE Affinity Reagents for the Enrichment of NRs**—We utilized HRE sequences to develop DNA reagents with high affinity for NR proteins. To systematically evaluate the effects of CHS orientation and spacing on NR binding, we synthesized 20 different HREs belonging to five different classes (see Figs. 1B–1F and supplemental Table S1): (1) DR-HREs, DRs (head to tail) with CHSs separated by one to five base pairs; (2) IR-HREs, IRs (head to head) with CHSs separated by one to five base pairs; (3) specific HREs for NR proteins that bind the CHS motif; (4) specific HREs for Gr, Ar, and Vdr that have CHS-like recognition motifs that deviate...
slightly from the CHS; and (5) DRm/IRm tandem mixed DRn/IRn sequences. To find the optimal number of HRE copies that enhance the binding affinity of the NR proteins, we also generated DNA constructs containing multiple copies of HREs described above. To make these constructs, synthetic double-stranded DNA oligos containing three copies of HREs and restriction site sticky ends were self-ligated and cloned into the pET24a+ vector. The ligation products were screened for plasmids that contained 3, 6, 9, or 12 tandem repeats of HREs (Fig. 2). These plasmids were then used as templates for PCR amplification of HREs with biotin-labeled primers. PCR products that contained multiple copies of HREs with biotins at both ends were immobilized to streptavidin-coated beads for the isolation of NR proteins from...
mouse liver nuclear extracts. The bound proteins were digested with trypsin and identified via mass spectrometry. The quantities of identified proteins were estimated using the iBAQ method (31), in which each protein quantity is calculated as the sum of peptide peak areas (or AUCs) normalized by the theoretical number of all possible full tryptic peptides for a given gene product (displayed in $10^3$ units, or iBAQ3, for better visual comprehension) or as a fraction of total iBAQ (iFOT5, displayed in $10^5$ units). Although the total number of proteins recovered in the DR1(12X) was comparable to that in the DNA control, the NR proteins were dramatically enriched on the DR1(12X) sequence (Fig. 3A). Seventeen members of NR proteins were detected by the DR1(12X) sequence, whereas the control DNA could detect only HNF4A, which was also significantly enriched on the DR1(12X) sequence (>200 folds).

Next, we tested whether an increase in the protein concentration can further increase the amounts of NR proteins detected. We made dilutions of NEs with a concentration range of 2.5–15 mg/ml and used these extracts in the DNA pull-downs with the same amount of DR1(12X) DNA. The number of detected NR proteins increased from 13 to 21 with increasing concentrations of NE, and the fraction of NR proteins increased from 0.5% to 5% of the total proteins pulled down by the DR1(12X) (Figs. 3B, 3C). This demonstrated that the DR1(12X) sequence was highly specific for the NR proteins, and that our protocol could be used to estimate changes in the DNA binding activity of NR family members.

To find the optimal copy number of HREs for NR enrichment, we used a series of DNA constructs containing 3, 6, 9, 12, and 15 copies of DR1. Surprisingly, we did not find a dramatic difference in the number or quantity of NR proteins pulled down by these DNA sequences (Figs. 4A, 4B), suggesting that under our experimental conditions three copies of HREs are sufficient, and there is no further gain in NR binding with additional HRE repeats.

Comparison of NR Proteins Enriched by Different DNA Binding Elements—The principles of NR binding specificity have been revealed through structural studies of DNA–NR complexes (32–37). These in vitro studies suggest that the binding specificity of NRs is determined not only by the sequence of DNA, but also by the geometric (direct or inverted) and spatial arrangement between the half-sites (23). We first tested the influence of spacing ($n = 1$ to 5 bp) between DRs on the efficiency of NR recovery. Twenty-one NR proteins were detected in this set of experiments (Fig. 5A). Most of the NR proteins were insensitive to the number of spacing base pairs between indirect repeats, with the exception of Ppard and...
Esrrg, which prefer IR3, and Rorc, which prefers the IR4 sequence.

We also observed that among the 25 NR proteins detected by DR and IR sequences, the majority of them bound HREs in both orientations. The exceptions are Nr1i2, which binds only DR sequences, and Ppard, Rora, Esrrg, and Rorc, which bind only IR sequences. We conclude that these unique NRs have more strict preferences regarding the orientation of the conserved binding half-sites.

The JASPA database curates specific consensus HRE motifs for several NRs, including those of Nurr1 (Nr4a2), Rora, Pparg, Esr1, and Esr2. We tested how these HREs alone behaved in their cognate NR protein binding with mouse liver extracts. Eleven members of NR proteins were identified by Esr1E and Esr2E, 14 by RoraE, and 18 and 19 by Nurr1E and PpargE, respectively (Fig. 5C). However, none of these enriched their specific cognate NRs to a level that could be detected by mass spectrometry.

Finally, we tested the NR binding capability of the DRm/IRm(s), constructs that contain tandem arrays of the DRn or IRn (n from 1 to 5) sequence. We hypothesized that DRm and IRm might bind more NR proteins than the single DRn and IRn because they provide more NR binding structures. However, only 18 and 17 NR proteins were detected using DRm and IRm, respectively (Fig. 5E), suggesting that DRm and IRm are not superior to single DRn and IRn for enriching NR proteins.

Nucleotide Composition in the Spacer Is Important for NR Protein Binding—The spacer between the half-sites also contributes to DNA binding by NR proteins, and it has been proposed that the “true” NR half-site is therefore eight bases long (38–40). As the number of base pairs in the spacer does not seem to impact the binding of NRs significantly, we investigated the impact of the spacer nucleotide composition on NR protein binding. We analyzed the difference between NR proteins isolated by a group of IR3 DNAs with different spacer sequences (Fig. 6A). This group of HREs included

![Fig. 3. NR proteins were enriched by means of HRE-containing DNA affinity pull-down.](image-url)
Enrichment and Detection of NR Proteins

**Fig. 4.** The influence of the copy number of an HRE on its NR binding capability. A series of DNA sequences containing 3, 6, 9, 12, or 15 copies of DR1-HRE were used to pull down NR proteins from mouse liver NE. The isolated protein mixtures were digested with trypsin on beads and identified via mass spectrometry. A, the total amount of NR superfamily proteins is not changed based on the number of DR1 copies in HRE-DNAs. B, a heatmap of relative protein amounts for individual NRs that were recovered with DR1 DNAs with increasing HRE repeat numbers.

**Fig. 5.** Comparison of the NR proteins enriched by different DNA elements. Twenty DNA sequences belonging to five kinds of HREs were used to enrich NR proteins from mouse liver NE. The bound proteins were identified via mass spectrometry. A, the influence of spacing between DR-HREs on the efficiency of NR protein binding. B, the influence of spacing between IR-HREs on the efficiency of NR protein binding. C, NR proteins pulled down by specific HREs containing a CHS motif. D, NR proteins pulled down by specific HREs containing a CHS-like motif. E, NR proteins pulled down by DRm and IRm whose sequences contain a tandem array of mixed DRn(s) or IRn(s).

Esr1E with a 5′-CCC-3′ spacer, Esr2E with a 5′-CGG-3′ spacer, PpargE with a 5′-CAG-3′ spacer, and IR3 with a 5′-AGT-3′ spacer sequence. Both Esr1E and Esr2E, whose spacers are composed of GC, allowed the detection of 11 NR proteins, whereas PpargE and IR3, whose spacers contain A/T, allowed the detection of 19 and 21 NR proteins, respectively. Furthermore, the quantity of NR proteins enriched by PpargE and IR3 was much greater than that enriched by Esr1E and Esr2E (Fig. 6A). Given that the only difference between Esr2E (CAG) and PpargE (CAG) is the A residue in the middle, it appears that the residue A/T in the middle of the 3-bp spacer is preferred for NR protein binding to the residue.
G/C, thus the nucleotide composition of the spacer is an important parameter for NR proteins binding in HREs.

**Pre-separation with IEF or SDS-PAGE Improves the Detection of NR Proteins**—We evaluated whether off-gel IEF or SDS-PAGE separation could help further enrich the low abundant NR proteins. The DR1 sequence was used to isolate NR proteins from mouse liver NE for this purpose. The DNA pull-down sample was directly digested on beads (SOB), digested first and separated into 12 peptide fractions by IEF, or run on SDS-PAGE and cut into six bands before digestion with trypsin. As anticipated, more proteins were identified when the sample was separated with IEF or SDS-PAGE (Fig. 6B).

Nineteen NR proteins were detected when the sample was digested on beads, whereas 27 and 30 NR proteins were detected when samples were separated with IEF and SDS-PAGE, respectively. In addition, the quantity of recovered NR proteins was dramatically increased with separation of DNA pull-down samples by SDS-PAGE. Considering the low cost and minimal additional effort required, SDS-PAGE is the preferred approach for the detection of NR proteins by DNA pull-down.

**In-depth Profiling of NR Protein DNA Binding Activity in Mouse Liver**—Although the global NR protein expression patterns in mouse liver are currently unknown, their “active,” or DNA-bound, pool can be measured with the method described above. Yang and colleagues showed that 20 NR genes exhibit circadian-like patterns of expression in mouse liver at the mRNA level (18), and we sought to investigate whether NR DNA binding activity is also rhythmic.

We first prepared male mouse liver NE at 6:00 a.m. and used DR1 and IR3 as affinity reagents to enrich NR proteins. Pull-down samples were fractioned on SDS-PAGE, and six bands were analyzed with mass spectrometry to identify a total of 30 NR proteins (Fig. 6C). We then made liver NE from male mice sacrificed at 10:00 a.m. and 2:00 p.m. These times were chosen because many NRs exhibited diurnal rhythmicity at the mRNA level (18), and most of these cycling NR transcripts peaked at 10:00 a.m. or 2:00 p.m. (supplemental Table S2; data obtained from NURSA).
In order to investigate gender differences, liver NE was also generated from female mice at 6:00 a.m. DR1 and IR3 were used to pull down NR proteins, followed by SDS-PAGE separation and sequencing of six protein bands. A total of 35 NRs were identified (Fig. 6C); most of these were identified in at least two samples, and only a few members were specifically identified in one sample. Orphan nuclear receptor Nr4a1 (Nurr77) and Esr1 (Esr) were specifically identified in female mouse liver. Nr0b2, also known as Shp, was detected only in male mouse liver collected at 2:00 p.m., coinciding with its peak expression at the mRNA level. Collectively, as many as 35 non-redundant members of NR proteins were identified in mouse liver. They constitute 85% of the 41 NR family members that are expressed in mouse liver at the mRNA level (Fig. 7A).

To investigate the relationships between the DNA binding properties of NRs and the HRE sequence, we analyzed our data via unsupervised hierarchical clustering using Gene Cluster 3.0 (Fig. 7B). The resulting dendrogram on the DNA element axis shows that these DNA elements branch into two major clusters (labeled I and II). Cluster II includes ArE and GrE that contain CHS-like recognition motifs and are slightly different from the conserved CHSs; cluster I includes all other HREs with the conserved CHSs. Cluster II DNA response elements tend to enrich specific NRs, whereas cluster I DNA response elements demonstrate a pan-binding property and can enrich different

---

**Fig. 7.** MS-based profiling reveals NR proteome in mouse liver and DNA-binding preferences of NR proteins on different types of HREs. A, proteomic and transcriptomic (described by Yang et al. (18)) profiling in liver cover >90% of NR superfamily members. B, a heatmap of the Gene Cluster 3.0–generated unsupervised hierarchical clustering for the NRs enriched on 20 DNA sequences. DNA sequences separate into two main clusters (I and II), and NR proteins are divided into Group I and Group II.
NRs. This indicates that the sequence of half-sites plays a major role in determining NR binding activities.

Cluster I further branches into two subordinate clusters (labeled IA and IB). Cluster IA includes seven DNA elements, among which DR4, DR5, and Nurr1E belong to DR HREs and IR1, Esr1E, Esr2E, and IR4 belong to IR HREs. Eleven DNA elements were grouped into cluster IB, also including DRs and IRs of half-sites. These clusters suggest that DRs or IRs, as well as the spacing, are not the sole determinants of NR binding specificity. It is the combination of DR/IR and spacing that perhaps determines the binding specificity.

The dendrogram on the NR axis also reveals a higher-order hierarchy among NRs on the basis of their DNA binding patterns. At the top of this hierarchy, the NRs are divided into two major groups (labeled I and II). Group I contains only Nr3c1 (Gr), which preferentially binds to the DNA sequences of ArE and GrE; group II contains the rest of the detectable NRs, and they tend to bind various kinds of DNA elements. A few members of NR proteins bind to specific DNA structures, such as Nr112, which specifically binds to DR1, and Rorc, which binds to IR4.

**DISCUSSION**

We reported an affinity-based method to enrich and detect endogenous NR proteins using nuclear hormone receptor response element DNA and mass spectrometry. We systematically evaluated the efficiency of 20 DNA sequences belonging to five types of HREs for the enrichment of NR proteins from mouse liver and found that 12X DR1 and 9X IR3 are the most efficient HRE constructs that can be used to enrich the majority of NR proteins from mouse liver. Furthermore, we optimized this protocol for in-depth coverage of NR proteins, including the preparation of NE from mouse liver and the separation of DNA-bound NR proteins with SDS-PAGE followed by mass spectrometry.

With these affinity reagents and our protocol, we were able to detect 35 NRs in mouse liver. To our knowledge, this is the first report of such in-depth coverage for endogenous NR proteins in a tissue. The analysis of the NR transcriptome in mouse liver via QPCR revealed that 41 members of the NR superfamily are expressed at the mRNA level (18). Although proteins are more directly correlated with their functions, the exact numbers of NR proteins expressed in liver were previously unknown. We have now shown that 35 NR proteins are also detectable in mouse liver at the protein level. Regarding the six transcripts for which no protein product was measured by mass spectrometry, it is possible that the proteins might be of extremely low abundance or unstable, such that they are below the detection limit of the current method.

Besides establishing an analytical method to measure the DNA binding activities of the NR family in mouse liver, the current study also sheds light on the DNA binding preferences of NR proteins. Hierarchical clustering analysis of enriched NR proteins indicates that the 20 different DNA sequences fall into five HRE classes, suggesting that the major factor in determining NR protein DNA binding is the consensus sequence of the CHS or the CHS-like binding motif. Most NR proteins bind to CHS DNAs with limited selectivity, and only a few NRs bind to DNAs with a CHS-like motif with high specificity.

HREs are arranged as DRs or IRs, and a 1–5 rule has been described for the target selectivity of DNA response elements with one to five base pairs of spacing (DR1 to DR5 and IR1 to IR5) (40, 41). Clustering analysis of our data indicates that all HREs containing CHS form cluster I, which branches into two subordinate clusters, IA and IB. Although differences in NR protein binding patterns are noticeable between clusters IA and IB, both IA and IB contain DNA response elements with DR and IR sequences and with different spacing. Thus, our in vitro pull-downs suggest that DRs and IRs, as well as the spacer length between the half-sites, do not seem to be the sole determinants of the selectivity of NR protein binding. Instead, a combination of these features might determine the efficiency of NR protein binding to DNA.

Structural studies suggest that DNA could act as an allosteric effector to modulate NR activities (42–46). For example, Tr4 orphan receptor functions as a repressor of its target genes when interacting with DR1-HRE and DR2-HRE, but as an activator when interacting with DR4-HRE (45). Incidentally, DR1/2 and DR4 separate into different sub-clusters, indicating that several NR proteins bind to these DNA sequences in patterns similar to that of Tr4, and they might contribute to the differentiation of Tr4 functions as a repressor or an activator.

Our analysis also revealed several unexpected findings. One surprise is the importance of the nucleotide composition of the spacer in determining the selectivity of NR proteins. Considerable differences were found in NR proteins enriched by Esr2E and PpargE, both of which belong to the IR3 class. Given that the only difference between Esr2E(CGG) and PpargE(CAG) is the A residue in the middle of the 3-bp spacer, this suggests that the nucleotide composition in the spacer between the half-sites plays an important role in NR protein binding. The other surprise is that copy numbers of HREs greater than three are not beneficial for the purpose of enriching NR proteins. Neither the number of NR proteins nor the quantity of NR proteins pulled down showed any obvious difference among constructs containing 3 to 15 HRE copies.

NR proteins in liver direct a broad range of key hepatic functions by sensing lipid-soluble hormones and dietary lipids and by governing the expression of key liver metabolic genes. The method presented here allows the analysis of NR proteins on the sub-proteome scale. The workflow could also be applied to cytoplasmic and even membrane extracts of cells, as many NRs have been shown to be expressed in other regions of the cell in addition to the nucleus. The workflow could also be applied to measure NR post-translational modifications. We anticipate that this method will find widespread use in...
future investigations of NR actions in liver biology and pathology.

* This work was supported by the National High-Tech Research and Development Program of China (863 Program) (2012AA020202), the State Key Lab of Proteomics (SKLP-K201001), the Nuclear Receptor Signaling Atlas (NURSA), and the Cancer Prevention Research Institute of Texas (CPRIT) (RP110784 to J.Q.).

S This article contains supplemental material.

To whom correspondence should be addressed: Jun Qin, E-mail: jqin@bcm.edu; and Bei Zhen, E-mail: zp1963@sina.com.

REFERENCES

1. Margolis, R. N., and Christakos, S. (2010) The nuclear receptor superfamily of steroid hormones and vitamin D gene regulation. An update. Ann. N.Y. Acad. Sci. 1192, 208–214

2. Mangelsdorf, D. J., Thummel, C., Beato, M., Herrlich, P., Schutz, G., Umesono, K., Blumberg, B., Kastner, P., Mark, M., Chambon, P., and Evans, R. M. (1995) The nuclear receptor superfamily: the second decade. Cell 83, 835–839

3. Zhang, Z., Burch, P. E., Cooney, A. J., Lanz, R. B., Pereira, F. A., Wu, J., Gibbs, R. A., Weinstock, G., and Wheeler, D. A. (2004) Genomic analysis of the nuclear receptor family: new insights into structure, regulation, and evolution from the rat genome. Genome Res. 14, 580–590

4. Jin, L., and Li Y. (2010) Structural and functional insights into nuclear receptor signaling. Adv. Drug Deliv. Rev. 62, 1218–1226

5. Sladek, F. M. (2011) What are nuclear receptor ligands? Mol. Cell. Endocrinol. 334, 3–13

6. Chawla, A., Repa, J. J., Evans, R. M., and Mangelsdorf, D. J. (2001) Nuclear receptors and lipid physiology: opening the X-files. Mol. Endocrinol. 580–590

7. Arrese, M., and Karpen, S. J. (2010) Nuclear receptors, inflammation, and liver disease: insights for cholestatic and fatty liver diseases. Cln. Pharmacol. Ther. 87, 473–478

8. Trauner, M., and Halilbasic, E. (2011). Nuclear receptors as new perspective for the management of liver diseases. Gastroenterology 140, 1120–1125, e1–e12

9. Vacca, M., Degirolamo, C., Mariani-Costantini, R., Palasciano, G., and Moschetta, A. (2011) Lipid-sensing nuclear receptors in the pathophysiology and treatment of the metabolic syndrome. Wiley Interdiscip. Rev. Syst. Biol. Med. 3, 562–587

10. Hansen, M. K., and Connolly, T. M. (2008) Nuclear receptors as drug targets in obesity, dyslipidemia and atherosclerosis. Curr. Opin. Investig. Drugs 9, 247–265

11. Overwjin, J. P., Al-Lazikani, B., and Hopkins, A. L. (2006) How many drug targets are there? Nat. Rev. Drug Discov. 5, 993–996

12. Krid, H., Dorison, A., Salti, A., Cheval, L., and Crambert, G. (2012) Expression profile of nuclear receptors along male mouse nephron segments reveals a link between ERbeta and thick ascending limb function. PLoS One 7, e34223

13. Shahbazi, M., Jeddhi-Tehrani, M., Zareie, M., Salek-Moghaddam, A., Akhondi, M. M., Bahmani, M., Sadeghi, M. R., and Zarnani, A. H. (2011) Expression profiling of vitamin D receptor in placenta, decidua and ovary of pregnant mice. Placenta 32, 657–664

14. Alexiadis, M., Eriksson, N., Jamieson, S., Davis, M., Drummond, A. E., Chu, S., Clyne, C. D., Muscat, G. E., and Fuller, P. J. (2011) Nuclear receptor profiling of ovarian granulosa cell tumors. Horm. Cancer 2, 157–169

15. Geong, Y., Xia, Y., Xiao, G., Behrens, C., Girard, L., Wistuba, I., Minna, J. D., and Mangelsdorf, D. J. (2010). Nuclear receptor expression defines a set of prognostic biomarkers for lung cancer. PLoS Med 7, e1000378

16. Yenamandra, S. P., Lundin, A., Arulampalam, V., Yurchenko, M., and Petry, V. (2007) Structural and functional insights into nuclear receptor signaling. Adv. Drug Deliv. Rev. 62, 1218–1226

17. Weinberg, A. L., Carter, D., Ahonen, M., Alarid, E. T., Murdoch, F. E., and Fritsch, M. K. (2007) The DNA binding domain of estrogen receptor alpha is required for high-affinity nuclear interaction induced by estradiol. Biochemistry 46, 8933–8942

18. Shaffer, P. L., and Gewirth, D. T. (2004) Vitamin D receptor-DNA interactions. Vitam. Horm. 68, 257–273

19. Khorasanizadeh, S., and Rastinejad, F. (2001) Nuclear-receptor interactions on DNA-response elements. Trends Biochem. Sci. 26, 384–390

20. Kobayashi, T., Kodani, Y., Nozawa, A., Endo, Y., and Sawasaki, T. (2008) DNA-binding profiling of human hormone nuclear receptors via fluorescence correlation spectroscopy in a cell-free system. FEBS Lett. 582, 2737–2744

21. Rochel, N., Ciesielksi, F., Godet, J., Moman, E., Roessle, I., Chovelon, M., Haerter, J. O., Svergun, D. I., and Moras, D. (2001) Common architecture of nuclear receptor heterodimers on DNA direct repeat elements with different spacings. Nat. Struct. Mol. Biol. 18, 564–570

22. Umesono, K., Murakami, K. K., Thompson, C. C., and Evans, R. M. (1991) Direct repeats as selective response elements for the thyroid hormone, retinoic acid, and vitamin D3 receptors. Cell 65, 1255–1266

23. Mangelsdorf, D. J., and Evans, R. M. (1993) The RXR heterodimers and orphan receptors. Cell 73, 841–850

24. Orlov, I., Rochel, N., Moras, D., and Klalhoz, B. P. (2012) Structure of the full human RXR/VDR nuclear receptor heterodimer complex with its DR3 target DNA. EMBO J. 31, 291–300

25. Navalite, I., Schwend, T., and Gustafsson, J. A. (2010) Proteomics analysis of the estrogen receptor alpha receptosome. Mol. Cell. Proteomics 9, 1411–1422

26. Li, W. C., Raiphs, K. L., and Tosh, D. (2010) Isolation and culture of adult mouse hepatocytes. Methods Mol. Biol. 633, 185–196

27. Schwanhausser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., and Selbach, M. (2011) Global quantification of mammalian gene expression control. Nature 473, 337–342

28. Luisi, B. F., Xu, W. X., Otwinowski, Z., Freedman, L. P., Yamamoto, K. R., and Sigler, P. B. (1991) Crystallographic analysis of the interaction of the glucocorticoid receptor with DNA. Nature 354, 567–578

29. Schwabe, J. W., Chapman, L., Finch, J. T., and Rhodes, D. (1993) The crystal structure of the estrogen receptor DNA-binding domain bound to DNA: how receptors discriminate between their response elements. Cell 75, 567–578

30. Zhao, Q., Chasse, S. A., Devarakonda, S., Sierk, M. L., Ahvazi, B., and Rastinejad, F. (2000) Structural basis of RXR-VDR DNA interactions. J. Mol. Biol. 299, 509–520

31. Rastinejad, F., Wagner, T., Zhao, Q., and Khorasanizadeh, S. (2000) Structure of the RXR-RAR DNA-binding complex on the retinoic acid response element DR1. EMBO J. 19, 1045–1054

32. Zhao, Q., Khorasanizadeh, S., Miyoshi, Y., Lazar, M. A., and Rastinejad, F. (1998) Structural elements of an orphan nuclear receptor-DNA complex. Mol. Cell 1, 849–861

33. Meineke, C., and Sigler, P. B. (1999) DNA-binding mechanism of the mono- neric orphan nuclear receptor NGFI-B. Nat. Struct. Biol. 6, 471–477

34. Harbers, M., Wahlstrom, G. M., and Vennstrom, B. (1996) Transactivation of the thyroid hormone and retinoic acid receptors: 3,4,5 rule modified. Trends Endocrinol. Metab. 7, 88–98

35. Meineke, C., and Sigler, P. B. (1999) DNA-binding mechanism of the mono- neric orphan nuclear receptor NGFI-B. Nat. Struct. Biol. 6, 471–477

36. Harbers, M., Wahlstrom, G. M., and Vennstrom, B. (1996) Transactivation by the thyroid hormone receptor is dependent on the spacer sequence in hormone response elements containing directly repeated half-sites. Nucleic Acids Res. 24, 2252–2259

37. Olson, D. P., and Koenig, R. J. (1997) S’-flanking sequences in thyroid hormone response element half-sites determine the requirement of retinoid X receptor for receptor-mediated gene expression. J. Biol. Chem. 272, 9907–9914

38. Phan, T. Q., Jow, M. M., and Privalsky, M. L. (2010) DNA recognition by thyroid hormone and retinoic acid receptors: 3,4,5 rule modified. Mol. Cell. Endocrinol. 319, 88–98

39. Xie, W., Barwick, J. L., Simon, C. M., Pierce, A. M., Safe, S., Blumberg, B.,
Guzelian, P. S., and Evans, R. M. (2000) Reciprocal activation of xenobiotic response genes by nuclear receptors SXR/PXR and CAR. *Genes Dev.* **14**, 3014–3023
42. Rastinejad, F., Perlmann, T., Evans, R. M., and Sigler, P. B. (1995) Structural determinants of nuclear receptor assembly on DNA direct repeats. *Nature* **375**, 203–211
43. Gronemeyer, H., and Bourguet, W. (2009). Allosteric effects govern nuclear receptor action: DNA appears as a player. *Sci. Signal* **2**, pe34
44. Landel, C. C., Kushner, P. J., and Greene, G. L. (1995) Estrogen receptor accessory proteins: effects on receptor-DNA interactions. *Environ. Health Perspect.* **103 Suppl 7**, 23–28
45. Lee, Y. F., Pan, H. J., Burbach, J. P., Morkin, E., and Chang, C. (1997) Identification of direct repeat 4 as a positive regulatory element for the human TR4 orphan receptor. A modulator for the thyroid hormone target genes. *J. Biol. Chem.* **272**, 12215–12220
46. Meijsing, S. H., Pufall, M. A., So, A. Y., Bates, D. L., Chen, L., and Yamamoto, K. R. (2009) DNA binding site sequence directs glucocorticoid receptor structure and activity. *Science* **324**, 407–410