Structural basis of binding of homodimers of the nuclear receptor NR4A2 to selective Nur-responsive DNA elements

Longying Jiang¹, Shuyan Dai¹, Jun Li¹, Xujun Liang¹, Lingzhi Qu¹, Xiaojuan Chen¹, Ming Guo¹, Zhuchu Chen¹, Lin Chen¹, Lin Chen¹,2, Hudie Wei¹,*¹, Yongheng Chen¹,*¹

¹ Department of Oncology, Laboratory of Structural Biology, NHC Key Laboratory of Cancer Proteomics, Xiangya Hospital, Central South University, Changsha, Hunan 410008, China.
² Molecular and Computational Biology Program, Department of Biological Sciences and Department of Chemistry, University of Southern California, Los Angeles, CA 90089, USA

* To whom correspondence should be addressed. Tel: +86 731 84327542; Fax: +86 731 84327542; Email: yonghenc@163.com (Y.C.); hudiewei18@163.com (H.W.)

Running Title: Structural basis of NR4A2 homodimer

Keywords: Nuclear receptor subfamily 4 group A member 2, nuclear receptor, transcription factor, Nur-responsive elements, crystal structure, protein-DNA binding, protein-protein interaction, gene regulation
Abstract

Proteins of nuclear receptor subfamily 4 group A (NR4A), including NR4A1/NGFI-B, NR4A2/Nurr1, and NR4A3/NOR-1, are nuclear transcription factors that play important roles in metabolism, apoptosis, and proliferation. NR4A proteins recognize DNA response elements as monomers or dimers to regulate the transcription of a variety of genes involved in multiple biological processes. In this study, we determined two crystal structures of the NR4A2 DNA-binding domain (NR4A2-DBD) bound to two Nur-responsive elements (NurREs), namely, an inverted repeat and an everted repeat at 2.6–2.8 Å resolutions. The structures revealed that two NR4A2-DBD molecules bind independently to the everted repeat, whereas two other NR4A2-DBD molecules form a novel dimer interface on the inverted repeat. Moreover, substitution of the interfacial residue Valine 298 to lysine, as well as mutation of DNA bases involved in the interactions, abolished the dimerization. Overall, our structural, biochemical, and bioinformatics analyses provide a molecular basis for the binding of the NR4A2 protein dimers to NurREs and advance our understanding of the dimerization specificity of nuclear receptors.

Introduction

The NR4A subfamily, including NR4A1 (NGFI-B), NR4A2 (Nurr1) and NR4A3 (NOR-1), belongs to the nuclear receptor superfamily (1,2). Most nuclear receptor members are regulated by small lipophilic ligands such as retinoids and steroids, while NR4A proteins are orphan members for which no ligand has been identified (3). The NR4A proteins function as transcription factors (TFs) that regulate the expression of many key genes involved in proliferation (Wnt/β-catenin) (4), metabolism (peroxisome proliferator-activated receptor γ coactivator 1a (PGC1α)), apoptosis (Fas-ligand, TRAIL), inflammation (interleukin-8), DNA repair (arginase-1) (5) and angiogenesis (VEGF) (6). Altered expression of NR4A receptors has been implicated in a wide variety of cancers, such as melanoma, colon cancer and oral squamous cell carcinoma (7-9). Increasing evidence has shown roles of the NR4A receptors in cancer immunity. NR4A receptors are expressed at high levels in CD8+ T cells from humans with cancer or chronic viral infections and are linked to the induction of T cell dysfunction (10–12). NR4A receptors repress effector gene expression by inhibiting the function of the TF AP-1 and activate tolerance-related genes by promoting acetylation of histone 3 at lysine 27 (12). The treatment of tumor-bearing mice with T cells expressing chimeric antigen receptors (CAR-T cells) lacking all three NR4A receptors in vivo resulted in tumor regression and prolonged survival (13).

Nuclear receptors engage the ‘hormone response elements’ (HREs) of target genes to regulate transcription. Most nuclear receptors dimerize on target DNA, which contains two repeats of the hexamer AGAACA or AGGTCA (14). The two half-sites can be organized as direct, inverted or everted repeats with spacers of varying lengths (15) (Figure S1). For example, steroid receptors bind as homodimers to inverted palindromic repeats separated by a three-nucleotide spacer (IR3). Nonsteroid nuclear receptors, such as retinoic acid receptors (RAR), retinoid X receptors (RXR), and vitamin D3 receptors (VDR), form homodimers or heterodimers with RXR on response elements consisting of direct repeats separated by a one-to-five nucleotide spacer (DR1-DR5).

NR4A receptors were initially found to bind as monomers to a NGFI-B-responsive element (NBRE, AAAGGTCA), an
octanucleotide consisting of the canonical nuclear receptor binding motif AGGTCA preceded by two adenines (16). In addition, NR4A1 and NR4A2 (but not NR4A3) can bind as homodimers or heterodimers to the Nur-responsive element (NurRE), which consists of two repeats of the NBRE-related octanucleotides (17). NR4A1 and NR4A2 can also form heterodimers with RXR (18). A large sample analysis of most human TFs by high-throughput SELEX (HT-SELEX) and ChIP sequencing revealed that NR4A2 can bind to two different NurRE motifs. One motif, named ER0, consists of two everted palindromic eight-nucleotide half-sites with no spacer. The other motif is IR5, which consists of two inverted eight-nucleotide repeats spaced by five nucleotides (19). NR4A proteins also bind to the promoter regions of the pituitary proopiomelanocortin (POMC) gene to mediate the physiological response of the POMC gene to corticotropin-releasing hormone (CRH) (20). CRH signals rapidly activate nuclear DNA binding activity of NR4A dimers but not monomers (21). In turn, NR4A dimers synergistically enhance the transcription of NurRE reporters (22).

The three NR4A receptors all share a common structural organization with other nuclear receptors. These receptors share a high degree of sequence homology in their ligand-binding domains (LBDs) and DNA-binding domains (DBDs), but exhibit divergent N-terminal regions containing the activation function 1 (AF1) domain (16). A previously reported three-dimensional structure of a rat NR4A1-DBD protein showed the DNA binding mode of the NR4A1 monomer (16), while the DNA binding mode of NR4A receptors as dimers remains unclear.

In this study, we report two crystal structures of NR4A2-DBD bound to ER0 and IR5 at 2.5 Å and 2.7 Å, respectively. These structures reveal that NR4A2-DBD can bind the two DNAs in different manners. We further analyzed the roles of the protein-protein interactions between two NR4A2 molecules and protein-DNA interactions in the promotion of the formation of the NR4A2-DBD/IR5 complex. Then, a bioinformatic analysis of the endogenous NR4A-binding motif identified the existence and ratio of the ER0- and IR5-binding motifs in vivo. Overall, our structural, biochemical and bioinformatics analyses will help elucidate the molecular basis of the DNA binding specificity of NR4A dimers.

**Results**

**Ability of NR4A2 to bind to different DNAs**

As mentioned above, NR4A2 can bind to different response elements as a monomer or dimer (16,19). Here, we carried out an electrophoretic mobility shift assay (EMSA) to analyze the ability of NR4A2-DBD to bind to three different DNAs in vitro (Figure 1A). Purified recombinant NR4A2-DBD protein was incubated with NBRE, ER0 and IR5. Then, samples were detected on a native polyacrylamide gel. As shown in Figure 1B, NR4A2-DBD formed a complex band with NBRE (Figure 1B, lanes 2 and 3). In contrast, NR4A2-DBD formed two mobility complex bands when incubated with ER0 or IR5 (Figure 1B, lanes 6 and 10). The lower band migrated to the same position as the NR4A2-DBD monomer bound to NBRE. The upper, slow-migrating band presumably represented the dimeric complex. With the decrease in free DNA levels upon protein binding, NR4A2-DBD formed an enhanced dimeric complex with ER0 and IR5 when the amount of NR4A2 protein increased (Figure 1B, lanes 7 and 11). The dimer/monomer ratio was quantified using ImageJ (Figure 1C). For site ER0, the dimer/monomer ratio was 0.17, 0.28 and 0.70 with a 1-, 2- and 3-fold
increase in protein concentration, respectively. For site IR5, the ratio increased to 0.74, 1.62, and 3.09 for the same protein concentrations. These results indicated that NR4A-DBD could bind as a dimer to the ER0 and IR5 sequences in a cooperative manner, and the IR5 site was more cooperative and stronger than site ER0.

**Overall structures of NR4A2-DBD bound to two NurREs**

To better characterize the mechanism by which NR4A2 proteins recognize NurRE dimers, we determined the crystal structures of human NR4A2-DBD bound to ER0 and IR5, respectively. The methods of crystallization and structure determination are described in the Experimental procedures section. The final refinement statistics are summarized in Table 1. The NR4A2-DBD/IR5 structure was solved at 2.8 Å and crystallized in the P 21 space group with one complex per asymmetric unit. The NR4A2-DBD/ER0 structure was solved at 2.6 Å and crystallized in the P 43 space group with one complex per asymmetric unit. The initially screened NR4A2-DBD/ER0 crystals were twinned. After multiple rounds of optimization, the twinning problem was not solved. The best set of diffraction data for the NR4A2-DBD/ER0 crystals was collected, which was estimated to be 47.8% twinned by phenix.xtriage (23). Although we applied detwinning during refinement, the R-factors of the final model remained higher than that expected for non-twinned crystals of similar resolution range (R-work/R-free: 0.25/0.29, resolution: 2.60 Å).

The global conformation of NR4A2-DBD had few structural variations in these two complex structures, which were composed of two highly conserved α-helical zinc modules (H1 and H2), an N-terminal loop, a loop linking H1 and H2, and a C-terminal extension (CTE) (Figure 2A). The core DBD comprises two highly conserved four-cysteine/zinc nucleated modules. The CTE includes a T-box and an A-box (Figure 2C). The structure of NR4A2-DBD also showed an analogous tertiary structure with the previously determined rat NR4A1-DBD structure (16), and the root mean square deviation (RMSD) for the superimposition of Ca atoms was approximately 0.33 Å.

In the structure of the NR4A2-DBD/ER0 complex, two NR4A2-DBD molecules were bound to the same face of the double-stranded DNA in a tail-to-tail orientation (Figure 2B). The two molecules were independently bound to the half-sites. In the structure of the NR4A2-DBD/IR5 complex, two NR4A2-DBD molecules were arranged in head-to-head orientation. Moreover, the two NR4A2-DBD molecules formed an interface through the reversely arrayed loop from residues 294-300. The protein-protein interactions might help NR4A2-DBD homodimers cooperatively bind IR5, and we will analyze this aspect later in the results. Overall, NR4A2 bound in different manners to the two NurRE sequences. The difference might result from the relative orientation (inverted or everted repeats) and number of spacer nucleotides of the half-sites.

**DNA recognition by NR4A2**

We analyzed the detailed information of protein-DNA interactions by NUCPLOT (24) as shown in Figure 3. The interactions were basically conserved for half-site recognition, and we used the NR4A2-DBD bound downstream of IR5 as a representative example. The N-terminal helix H1 of NR4A2-DBD docked into and predominantly interacted with the major groove. Three conserved residues, namely, Glu281, Lys284 and Arg289, formed hydrogen bonds with -5Cyd, 4′Gua and -7Gua, respectively (Figure 3A). The C terminus of NR4A2-DBD formed a second independent DNA-binding surface to interact with the minor groove. The conserved Arg-Gly-Arg motif (Arg342,
Gly343 and Arg344) made hydrogen bonds contacts with the flanking extension bases -3Thy, -3'Ade and -1Thy, respectively (Figure 3B). In addition to specific base contacts, the NR4A2-DBD protein also formed numerous hydrogen bonds and van der Waals interactions with the phosphate backbone to further stabilize DNA binding (Figure 3C).

**Protein-protein interactions in the NR4A2-DBD/IR5 complex**

In the structure of the NR4A2-DBD/IR5 complex, a dimerization interface between two NR4A2-DBD molecules was identified. This interface involves the same region of the two NR4A2 molecules (amino acids 294-300) in an antiparallel manner (Figure 4A). From an analysis of PBDePISA (25), the dimer interface buried a solvent-accessible area of approximately 419 Å². The dimerization contacts involved residues Asn294, Val298 and Leu300 of the upstream subunit and residues Asn294, Lys296, Val298 and Leu300 of the downstream subunit (Figure 4B).

To identify the role of the dimer interface, we mutated the key hydrophobic residue Val298 to a charged lysine. Then, the effect of mutant V298K on DNA binding ability was measured using EMSA. Compared with wild-type NR4A2-DBD, the dimeric complex was hardly detected when the mutant V298K was incubated with IR5 (Figure 4C). The results indicated that Val298 was crucial for NR4A2 dimerization on IR5, and the cooperative assembly was inhibited without the dimer interaction.

We also investigated the effect of the IR5 sequence on the DNA binding properties of NR4A2. The three bases (-7Gua, -5Cyd and -3Thy) in the downstream half-site involved in the specific protein-base interactions were mutated to other bases. As shown in Figure 4D, the dimeric complexes were abolished when wild-type NR4A2-DBD was incubated with these IR5 variants, apart from a -7Gua to Ade mutant. The results indicated the crucial roles of these bases for NR4A2 dimerization. Taken together, these results suggested that the protein-protein interactions between two NR4A2 molecules, as well as the specific protein-base interactions, play similar important roles in favoring the cooperative formation of the NR4A2-DBD/IR5 complex.

**Analysis of the NR4A2 and RXR heterodimer**

To investigate whether NR4A2 can heterodimerize with RXR on the IR5 response element, we modeled the NR4A2/RXR heterodimer bound to IR5 by superposition of one NR4A2 molecule with a RXR structure (PDB: 4CN2) (Figure 5A). Given the sequence similarities and high homology, the RXR exhibited a good fit and showed no obvious clashes. Then, we carried out EMSA to analyze the heterodimeric abilities of NR4A2 and RXR bound to IR5 response elements (Figure 5B). NR4A2 or RXR alone could form a monomer band with IR5 (Figure 5B, lanes 2 and 3). When IR5 was incubated with NR4A2 and RXR at the same time, a higher complex band was observed (Figure 5B, lane 4). The results suggested that NR4A2 could form a heterodimer with RXR on the IR5 response element in vitro.

**Identification of the endogenous NR4A binding motifs ER0 and IR5**

To identify whether the NR4A protein binds ER0 and IR5 in vivo, we analyzed the occurrence of these motifs in endogenous NR4A-binding sites. We first obtained the ER0 and IR5 matrixes from the footprintDB database (http://floresta.eead.csic.es/footprintdb/index.php) (Figure 6A and 6B). Then, we searched for the motif IR5 or ER0 in the NR4A ChIP-seq database (GSE123629) using these two matrixes. For motif ER0, 9178 peaks were found, accounting for 9.2% of the total
NR4A-binding sites. For motif IR5, 33780 peaks were found, accounting for 33.9% of the total NR4A-binding sites (Figure 6C). The ER0 motif was identified in the promoters of genes such as Mmp9, Foxo1, Wnt2b, Hdac7, and Fgf1 (Figure 6D). The IR5 motif was identified in the promoters of genes such as Cyp17a1, Bcl6, Nfkb1, Bach2, and Gata3 (Figure 6D). These data suggested that NR4A proteins could indeed bind ER0 or IR5 motifs in vivo.

**Discussion**

In this paper, we first report the cocrystal structures of NR4A2-DBD bound to two different NurREs. Both structures demonstrate that NR4A2 makes similar base interactions with half-sites as those seen in the structure of rat NR4A1 bound to NBRE (16). The highly conserved helix H1 forms specific contacts with the identity elements of the major groove (Figure 3). The C-terminal residues form a unique substructure to interact extensively with the minor groove, in particular with the characteristic 5’-flanking extended A-T base pairs of the canonical nuclear receptor binding motif AGGTCA (Figure 3).

In addition to exhibiting similar base recognition, NR4A2 subunits were bound to these two NurREs in distinct modes: (1) the everted repeats mode, in which two molecules independently bind to each half-site of DNA in a tail-to-tail orientation (Figure 2B), and (2) the inverted repeats mode, in which the two molecules form a dimerization interface and synergistically bind to each half-site of DNA in a head-to-head orientation (Figure 2A). The primary association across the interface is via van der Waals contacts. The residue Val298 contributes to the hydrophobic character of the interface, and substitution of the hydrophobic valine with a charged lysine is likely to strongly inhibit homodimer formation (Figure 4). In addition to the dimer interface, the presence of two inverted DNA binding sites also plays important roles in driving formation of the dimer on DNA (Figure 4). Moreover, bioinformatic analysis of ChIP-seq data revealed that 33.9% and 9.2% of the total NR4A binding sites contain IR5 and ER0 sites, respectively (Figure 6). These results suggest that NR4A proteins could bind these sites in vivo.

Previous study reported that NR4A2 could heterodimerize with RXRα or RXRγ in midbrain dopaminergic (DAergic) neurons (26). The important role of the NR4A2/RXR heterodimer in vivo has been well studied, and several synthetic ligands that bind to the RXR-binding pocket are thought to be potential therapeutic agents that act by activating NR4A2/RXR heterodimers (27). Based on our structure of the NR4A2 homodimer, we modeled the structures of the NR4A2/RXR heterodimer bound to the IR5 motif (Figure 5). It will help explain the molecular mechanism by which the NR4A2/RXR heterodimer binds to DNA. Moreover, our EMSA results verified that NR4A2 could heterodimerize with RXR on the IR5 response element (Figure 5).

Nuclear receptors form dimers on their target DNAs via highly cooperative assembly of their DBDs. With highly conserved core DBDs and similar response elements, nuclear receptors can recognize and differentiate between specific DNAs. This specificity can be partly attributed to the ability of nuclear receptors to make distinct protein-protein contacts to reinforce their protein-DNA contacts, as observed for glucocorticoid receptor (GR), RXR, and VDR (Figure S2). GR forms the largest dimerization interface on IR3 (PDB code: 3G6P) (28). Salt bridges, numerous hydrogen bonds and van der Waals forces contribute to the dimer interactions (Figure S2A). In the structure of the RXR/DR1 complex (PDB code: 4CN2) (29),
Glu207 of the downstream subunit makes hydrogen bond interactions with Arg182 of the upstream subunit, and Gln210 of the downstream subunit makes contacts with residues Arg172 and Arg186 of the upstream subunit (Figure S2B). VDR binds to DR3 with a dimerization interface involving the side chains of Pro61, Phe62 and His75 of the upstream subunit and residues Asn37, Glu92 and Phe93 of the downstream subunit (Figure S2C) (30). In the NR4A2/IR5 structure, the hydrophobic residue Val298 plays a key role in the dimerization interaction, as well as numerous surrounding van der Waals contacts (Figure S2D). Compared to these previously published nuclear receptor dimer structures, the NR4A2-DBD homodimer utilizes a different region to form a compact and nonpolar dimer interface.

The polarity and strength of dimers are modulated by the spacing and relative orientation of the half-sites. DNAs with direct repeats, inverted repeats and everted repeats exhibit nuclear receptor binding as head-to-tail, head-to-head and tail-to-tail dimers, respectively. For example, RXR (29), Rev-Erb (31) and VDR (30) engage with DR1, DR2 and DR3, respectively, as head-to-tail homodimers (Figure S3A-C). The structure was similar in the case of DR4 sequences recognized by a RXR/TR heterodimer (32) (Figure S3D). A dimer interface was formed between the CTE sequence of the downstream subunit and the second zinc finger of the upstream subunit in all these structures. The 1- to 4-bp spacers between two half-sites diversified the relative displacement of the two protein subunits on the DNAs. For inverted repeats, steroid receptors such as GR (33) and NR4A2 bind to IR3 and IR5 sites as head-to-head homodimers, respectively. The different spacers drove the participation of varying residues in the dimer interaction. Two GR subunits form an extensive dimer interface through residues along with the zinc ions in the second zinc-binding motif (Figure S3E), while two NR4A2 subunits engaged with the IR5 sequence form a narrow interface involving N-terminal residues close to the second zinc-binding motif (Figure 2A). For everted repeats, the structure of NR4A2/ER0 determined here demonstrated that two NR4A2 molecules formed a tail-to-tail orientation (Figure 2B). Overall, the relative orientation and diverse spacers of the two half-sites dictated the different displacements and dimer interfaces of nuclear receptor homodimers.

In addition to DNA recognition by DBDs of nuclear receptors, the LBDs can also affect the DNA binding properties of these receptors (34). Classic LBD-mediated dimerization interactions have been observed in the structure of the full-length PPARγ-RXRα heterodimer bound to DR1 (PDB ID: 3DZY) (34). The LBD of PPARγ not only forms contacts with the LBD of RXRα through helices 7, 9 and 10 of each receptor but also interacts with the DBD CTE region of RXRα. These dimerization interfaces contribute to stabilizing DR1 binding. Although the structure of full-length NR4A receptors has not been determined, the sequences that share homology with other nuclear receptors indicate that the LBDs might also help in defining the preferred dimerization to bring the full-length proteins in close proximity to DNA.

In summary, we determined two crystal structures of NR4A2-DBD/DNA complexes and provided the molecular basis for DNA recognition by NR4A dimers. We revealed a new mode of nuclear receptor binding as a dimer to IR5. The two DBDs formed a novel dimer interface, primarily via van der Waals contacts between residues formed by the loop preceding the second zinc finger motif of each subunit. The dimer interface and protein-DNA
interactions both play important roles in favoring cooperative dimer formation. Our structural, biochemical and bioinformatic analyses may provide a better understanding of the dimerization specificity of nuclear receptors.

**Experimental procedures**

**Expression and purification**

Human NR4A2-DBD (residues 259-348) was cloned into a modified pMAL-C5X vector (35). The NR4A2 plasmid was transformed into *Escherichia coli* Rosetta (DE3) cells. After purification by an amylose resin column (BioLabs), the N-terminal maltose-binding protein (MBP) tag was removed by PreScission protease at 4 °C overnight. The cleaved protein was further purified by a Mono S cation exchange column and a Superdex 75 column (36). The final protein was concentrated to approximately 20 mg/ml in 20 mM HEPES (pH 7.5), 200 mM NaCl, and 0.5 mM TCEP (37). DNA was synthesized by Genewiz (Nanjing, China) and purified as previously described (38). Human RXR-DBD was cloned into the pGEX-6P1 vector and purified as previously described (29).

**Crystallization and data collection**

Protein and DNA complexes were prepared by mixing protein and DNA at a 5:3 molar ratio. Crystals of NR4A2-DBD/ER0 were grown at 18 °C by the hanging drop method with a reservoir buffer containing 150 mM NaCl, 50 mM MES (pH 5.93), 10 mM MgCl$_2$, 5 mM CaCl$_2$, and 10-12% PEG4K (w/v). Crystals of NR4A2-DBD/IR5 were grown at 18 °C with a reservoir buffer containing 50 mM Na acetate (pH 4.7), 200 mM NaCl, 10 mM MgCl$_2$, 5 mM Li$_2$SO$_4$, 5 mM DTT, and 15-19% PEG4K (w/v). Then, crystals were transferred into a well solution containing an additional 20% glycerol (w/v) and flash-frozen in liquid nitrogen. Data were collected at the BL17U1 beamline and BL19U1 beamline of Shanghai Synchrotron Radiation Facility (SSRF).

**Structure determination**

Data were reduced using HKL2000 (39). The structures were solved by molecular replacement (MR) using phaser from the PHENIX package (40). A previously solved rat NR4A1 structure (PDB Code: 1CIT) (16) was used as an initial search model. Further refinements were performed using Phenix and CCP4 (41,42). Graphical representations of structures were generated using PyMol (43). The statistics of the crystallographic analysis are presented in Table 1. Schematic diagrams of protein-DNA interactions were generated by NUCPLOT (24). Figures were generated using PyMol (44).

**Site-directed mutagenesis**

Site-directed mutagenesis of NR4A2-DBD was performed according to the manufacturer’s instruction for the ClonExpress II One Step Cloning Kit (Vazyme) using the pMAL-C5X-NR4A2 plasmid as the template (45). These mutants were verified by DNA sequencing (Tsingke, Changsha, China). Mutant NR4A2 proteins were expressed and purified as wild-type protein.

**Electrophoretic mobility shift assay (EMSA)**

Protein and DNA samples were prepared at a concentration of 45 μM. DNA was incubated with protein in a total volume of 6 μl in a buffer containing 20 mM HEPES (pH 7.5), 10 mM MgCl$_2$, 200 mM NaCl, and 0.01% TritonX 100 for 20 minutes at room temperature. A native 8% (w/v) polyacrylamide gel (PAGE) in 0.5×TBE buffer was used to separate the free DNA from the protein/DNA complex (46). The gel was visualized after staining with GoldView.

**Bioinformatics analysis**

The bed files of the ChIP-seq data
The core sequences ER0 and IR5 were searched against the ChIP-seq peak sequences using the R package TFBSTools (48).

**Accession numbers**

The coordinates and structure factors are deposited in the Protein Data Bank under the accession codes 6L6Q (NR4A2-DBD/ER0), and 6L6L (NR4A2-DBD/IR5).

**Acknowledgments**

This work was supported by the National Natural Science Foundation of China (grants 81372904 and 81570537 to Y.C.; grant 81272971 to Z.C.), the Science and Technology Department of Hunan Province (2015SK2037). We thank the staff from BL17U1 beamline and BL19U1 beamline of the Shanghai Synchrotron Radiation facility (SSRF) for help with data collection. We thank Dr. Michael R. Stallcup for proofreading.

**Conflict of interest**

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

**REFERENCES**

1. Wang, J., Yang, J., Zou, Y., Huang, G. L., and He, Z. W. (2013) Orphan nuclear receptor nurr1 as a potential novel marker for progression in human prostate cancer. *Asian Pac J Cancer Prev* **14**, 2023-2028

2. Mohan, H. M., Aherne, C. M., Rogers, A. C., Baird, A. W., Winter, D. C., and Murphy, E. P. (2012) Molecular Pathways: The Role of NR4A Orphan Nuclear Receptors in Cancer. *Clinical Cancer Research An Official Journal of the American Association for Cancer Research* **18**, 3223

3. Zhan, Y. Y., Chen, Y., Zhang, Q., Zhuang, J. J., Tian, M., Chen, H. Z., Zhang, L. R., Zhang, H. K., He, J. P., Wang, W. J., Wu, R., Wang, Y., Shi, C., Yang, K., Li, A. Z., Xin, Y. Z., Li, T. Y., Yang, J. Y., Zheng, Z. H., Yu, C. D., Lin, S. C., Chang, C., Huang, P. Q., Lin, T., and Wu, Q. (2012) The orphan nuclear receptor Nur77 regulates LKB1 localization and activates AMPK. *Nature chemical biology* **8**, 897-904

4. Berwick, D. C., and Harvey, K. (2014) The regulation and deregulation of Wnt signaling by PARK genes in health and disease. *Journal of molecular cell biology* **6**, 3-12

5. Yin, K., Chhabra, Y., Tropée, R., Lim, Y. C., Fane, M., Dray, E., Sturm, R., and Smith, A. G. (2017) NR4A2 Promotes DNA Double-strand Break Repair Upon Exposure to UVR. *Molecular Cancer Research* **15**, molcanres.0002.2017

6. Mohan, H. M., Aherne, C. M., Rogers, A. C., Baird, A. W., Winter, D. C., and Murphy, E. P. (2012) Molecular pathways: the role of NR4A orphan nuclear receptors in cancer. *Clinical cancer research : an official journal of the American Association for Cancer Research* **18**, 3223-3228

7. Zhang, L., Wang, Q., Liu, W., Liu, F., Ji, A., and Li, Y. (2018) The Orphan Nuclear Receptor 4A1: A Potential New Therapeutic Target for Metabolic Diseases. *Journal of Diabetes Research*

8. Holla, V. R., Mann, J. R., Qiong, S., and Dubois, R. N. (2006) Prostaglandin E2 regulates the nuclear receptor NR4A2 in colorectal cancer. *Journal of Biological Chemistry* **281**, 2676-2682

9. Shigeishi, H., Higashikawa, K., Hatano, H., Okui, G., Tanaka, F., Tran, T. T., Rizqiawan, A., Ono, S., Tobiume, K., and Kamata, N. (2011) PGE 2 targets squamous cell carcinoma cell with the
activated epidermal growth factor receptor family for survival against 5-fluorouracil through NR4A2 induction. *Cancer Letters* **307**, 227-236

10. Philip, M., Fairchild, L., Sun, L., Horste, E. L., Camara, S., Shakiba, M., Scott, A. C., Viale, A., Lauer, P., and Merghoub, T. (2017) Chromatin states define tumour-specific T cell dysfunction and reprogramming. *Nature* **454**, 452

11. Scott-Browne, J. P., Lopez-Moyado, I. F., Trifari, S., Wong, V., Chavez, L., Rao, A., and Pereira, R. M. (2016) Dynamic Changes in Chromatin Accessibility Occur in CD8(+) T Cells Responding to Viral Infection. *Immunity* **45**, 1327-1340

12. Liu, X., Wang, Y., Lu, H., Li, J., Yan, X., Xiao, M., Hao, J., Alekseev, A., Khong, H., Chen, T., Huang, R., Wu, J., Zhao, Q., Wu, Q., Xu, S., Wang, X., Jin, W., Yu, S., Wang, Y., Wei, L., Wang, A., Zhong, B., Ni, L., Liu, X., Nurieva, R., Ye, L., Tian, Q., Bian, X. W., and Dong, C. (2019) Genome-wide analysis identifies NR4A1 as a key mediator of T cell dysfunction. *Nature* **567**, 525-529

13. Chen, J., Lopez-Moyado, I. F., Seo, H., Lio, C. J., Hempleman, L. J., Sekiya, T., Yoshimura, A., Scott-Browne, J. P., and Rao, A. (2019) NR4A transcription factors limit CAR T cell function in solid tumours. *Nature* **567**, 530-534

14. Khorasanizadeh, S., and Rastinejad, F. (2001) Nuclear-receptor interactions on DNA-response elements. *Trends in biochemical sciences* **26**, 384-390

15. Helsen, C., Kerkhofs, S., Clinkemalie, L., Spans, L., Laurent, M., Boonen, S., Vanderschueren, D., and Claessens, F. (2012) Structural basis for nuclear hormone receptor DNA binding. *Molecular and cellular endocrinology* **348**, 411-417

16. Meinke, G., and Sigler, P. B. (1999) DNA-binding mechanism of the monomeric orphan nuclear receptor NGFI-B. *Nature structural biology* **6**, 471-477

17. Maira, M., Martens, C., and Drouin, J. (1999) Heterodimerization between members of the Nur subfamily of orphan nuclear receptors as a novel mechanism for gene activation. *Molecular & Cellular Biology* **19**, 7549-7557

18. Aarnisalo, P., Kim, C. H., Lee, J. W., and Perlmann, T. (2002) Defining requirements for heterodimerization between the retinoid X receptor and the orphan nuclear receptor Nurr1. *The Journal of biological chemistry* **277**, 35118-35123

19. Jolma, A., Yan, J., Whittington, T., Toivonen, J., Nitta, K. R., Rastas, P., Morgunova, E., Enge, M., Taipale, M., Wei, G., Palin, K., Vaquerizas, J. M., Vincentelli, R., Luscombe, N. M., Hughes, T. R., Lemaire, P., Ukkonen, E., Kivioja, T., and Taipale, J. (2013) DNA-binding specificities of human transcription factors. *Cell* **152**, 327-339

20. Murphy, E. P., and Conneely, O. M. (1997) Neuroendocrine regulation of the hypothalamic pituitary adrenal axis by the nur77 subfamily of nuclear receptors. *Molecular Endocrinology* **11**, 39

21. Maira, M., Martens, C., Batsche, E., Gauthier, Y., and Drouin, J. (2003) Dimer-specific potentiation of NGFI-B (Nur77) transcriptional activity by the protein kinase A pathway and AF-1-dependent coactivator recruitment. *Mol Cell Biol* **23**, 763-776

22. Maira, M., Martens, C., Philips, A., and Drouin, J. (1999) Heterodimerization between members of the Nur subfamily of orphan nuclear receptors as a novel mechanism for gene activation. *Mol Cell Biol* **19**, 7549-7557

23. Zwart, P. H., Grosse-Kunstleve, R. W., Lebedev, A. A., Murshudov, G. N., and Adams, P. D. (2010) Surprises and pitfalls arising from (pseudo)symmetry. *Acta Crystallographica* **64**, seudo-seudo
24. Luscombe, N. M., Laskowski, R. A., and Thornton, J. M. (1997) NUCPLOT: a program to generate schematic diagrams of protein-nucleic acid interactions. *Nucleic Acids Research* **25**, 4940-4945
25. Krissinel, E., and Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state. *Journal of molecular biology* **372**, 774-797
26. Wallen-Mackenzie, A., Urquiza, A. M. D., Petersson, S., Rodriguez, F. J., and Perlmann, T. (2003) Nurr1-RXR heterodimers mediate RXR ligand-induced signaling in neuronal cells. *Genes Dev* **17**, 3036-3047
27. Scheepstra, M., Andrei, S. A., de Vries, R. M. J. M., Meijer, F. A., Ma, J.-N., Burstain, E. S., Olsson, R., Ottnann, C., Milroy, L.-G., and Brunsved, L. (2017) Ligand Dependent Switch from RXR Homo- to RXR-NURR1 Heterodimerization. *Acs Chemical Neuroscience* **8**, 2065-2077
28. 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
29. Osz, J., McEwen, A. G., Poussin-Courmontagne, P., Moutier, E., Birck, C., Davidson, J., Moras, D., and Rochel, N. (2015) Structural basis of natural promoter recognition by the retinoid X nuclear receptor. *Scientific reports* **5**, 8216
30. Shaffer, P. L., and Gewirth, D. T. (2002) Structural basis of VDR–DNA interactions on direct repeat response elements. *Embo Journal* **21**, 2242-2252
31. Zhao, Q., Khorasanizadeh, S., Miyoshi, Y., Lazar, M. A., and Rastinejad, F. (1998) Structural elements of an orphan nuclear receptor-DNA complex. *Molecular Cell* **1**, 849-861
32. 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
33. Meijsing, S. H., Pufall, M. A., So, A. Y., Bates, D. L., Lin, C., and Yamamoto, K. R. (2009) DNA binding site sequence directs glucocorticoid receptor structure and activity. *Science* **324**, 407-410
34. Chandra, V., Huang, P., Hamuro, Y., Raghuram, S., Wang, Y., Burris, T. P., and Rastinejad, F. (2008) Structure of the intact PPAR-gamma-RXR- nuclear receptor complex on DNA. *Nature* **456**, 350-356
35. Li, J., Jiang, L., Liang, X., Qu, L., Wu, D., Chen, X., Guo, M., Chen, Z., Chen, L., and Chen, Y. (2017) DNA-binding properties of FOXP3 transcription factor. *Acta biochimica et biophysica Sinica* **49**, 792-799
36. Yongheng, C., Chunxia, C., Zhe, Z., Chun-Chi, L., Johnson, M. E., Espinoza, C. A., Edsall, L. E., Bing, R., Xianghong Jasmine, Z., and Grant, S. F. A. (2015) DNA binding by FOXP3 domain-swapped dimer suggests mechanisms of long-range chromosomal interactions. *Nucleic Acids Research* **43**, 1268-1282
37. Yan, N., Li, J., Chen, X., Chen, Y., Chen, L., and Chen, Z. (2015) [Expression and purification of DNA binding domain of NR4A1]. *Zhong nan da xue xue bao. Yi xue ban = Journal of Central South University. Medical sciences* **40**, 345-350
38. Li, J., Dantas Machado, A. C., Guo, M., Sagendorf, J. M., Zhou, Z., Jiang, L., Chen, X., Wu, D., Qu, L., Chen, Z., Chen, L., Rohs, R., and Chen, Y. (2017) Structure of the Forkhead Domain of FOXA2 Bound to a Complete DNA Consensus Site. *Biochemistry* **56**, 3745-3753
39. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* **276**, 307-326
40. Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Echols, N., Headd, J. J., Hung, L. W., Jain, S., Kapral, G. J., Grosse Kusnetzov, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R. D., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2011) The Phenix software for automated determination of macromolecular structures. *Methods* **55**, 94-106

41. project, n. C. C. (2010) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* **50**, 760-763

42. Adams, P. D. (2010) PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallographica* **58**, 1948-1954

43. Warren, L. D. (2004) PyMOL User’s Guide. *Delano Scientific, San Carlos*

44. Bramucci, E., Paiardini, A., Bossa, F., and Pascarella, S. (2012) PyMod: sequence similarity searches, multiple sequence-structure alignments, and homology modeling within PyMOL. *BMC bioinformatics* **13 Suppl 4**, S2

45. Wu, D., Guo, M., Min, X., Dai, S., Li, M., Tan, S., Li, G. Q., Chen, X., Ma, Y., and Li, J. (2018) LY2874455 potently inhibits FGFR gatekeeper mutant and overcomes mutation-based resistance. *Chemical Communications*

46. Chen, X., Wei, H., Li, J., Liang, X., Dai, S., Jiang, L., Guo, M., Qu, L., Chen, Z., Chen, L., and Chen, Y. (2019) Structural basis for DNA recognition by FOXC2. *Nucleic Acids Res* **47**, 3752-3764

47. Zhu, L. J., Gazin, C., Lawson, N. D., Pagès, H., Lin, S. M., Lapointe, D. S., and Green, M. R. (2010) ChiPpeakAnno: a Bioconductor package to annotate ChiP-seq and ChiP-chip data. *Bmc Bioinformatics* **11**, 237

48. Tan, G., and Lenhard, B. (2016) TFBSTools: an R/bioconductor package for transcription factor binding site analysis. *Bioinformatics* **32**, 1555-1556

**Abbreviations and nomenclature:**

NR4A, nuclear receptor subfamily 4, group A; NR4A2-DBD, NR4A2 DNA-binding domain; NurRE, Nur-responsive elements; ER0, two everted palindromic half-sites; IR5, two inverted repeats spaced by five nucleotides; DR1, direct repeats of half-site separated by one nucleotides spacer; EMSA, Electrophoretic mobility shift assay.
Table 1. Data collection and refinement statistics.

|                         | NR4A2-DBD/ER0 complex | NR4A2-DBD/IR5 complex |
|-------------------------|------------------------|------------------------|
| **Data collection**     |                        |                        |
| Wavelength (Å)          | 0.97915                | 0.97915                |
| Space group             | P 4₃                   | P 2₁                   |
| a, b, c (Å)             | 35.696, 35.696, 259.273| 42.405, 64.315, 60.776 |
| α, β, γ (°)             | 90, 90, 90             | 90, 96.84, 90          |
| Resolution (Å)          | 50-2.601 (2.693-2.601) | 50-2.781(2.88-2.781)   |
| Unique reflections      | 9915 (990)             | 8163 (737)             |
| Redundancy              | 13.8 (13.3)            | 3.3(3.0)               |
| Completeness (%)        | 99.95 (100.00)         | 98.67 (89.99)          |
| Mean I/sigma(I)         | 16.81 (4.42)           | 6.69 (2.01)            |
| **Refinement**          |                        |                        |
| Resolution (Å)          | 32.41-2.601            | 35.23-2.781            |
| Wilson B-factor         | 47.51                  | 37.19                  |
| R-work/R-free           | 0.25/0.29              | 0.22/0.25              |
| Number of non-hydrogen atoms | 2172                 | 2205                   |
| macromolecules          | 2147                   | 2177                   |
| Protein residues        | 210                    | 211                    |
| RMS (bonds)             | 0.008                  | 0.008                  |
| RMS (angles)            | 1.17                   | 1.12                   |
| Ramachandran favored (%)| 95                     | 96                     |
| Ramachandran outliers (%)| 0                    | 0                      |
| Clashcore               | 15.47                  | 11.25                  |
| Average B-factor        | 53.10                  | 41.70                  |
| Macromolecules          | 53.20                  | 41.90                  |
| Ligands                 | 42.70                  | 31.10                  |
| Solvent                 | 39.30                  | 20.50                  |
FIGURES LEGENDS

Figure 1. Binding ability of NR4A2-DBD with different DNAs. (A) DNA sequences of the response elements for NR4A2. (B) Binding features of NR4A2-DBD with different DNAs were determined by EMSA. Lanes 1, 4, and 8 show free DNA without protein. The primary concentrations of DNA and protein used were both 45 µM. + indicates that the molar ratio of protein to DNA is 1:1, and ++ indicates that the molar ratio is 2:1. (C) Quantification of the intensity of dimer/monomer bands. The graph shows the relative density of the dimer/monomer bands detected by EMSA.

Figure 2. Overall structures of NR4A2-DBD/DNA complexes. (A) Overall structure of the NR4A2-DBD/IR5 complex. NR4A2-DBD is colored cyan, and DNA targets are colored light pink. (B) Overall structure of the NR4A2-DBD/ER0 complex. NR4A2-DBD is colored magenta, and ER0 is colored lemon. The sequences of the DNA duplex used in the structures are listed below, and the core octanucleotides are colored red. The arrow shows the orientation of each core half-site. Secondary structure elements (H1 and H2) are labeled, and the region in the box is involved in the protein-DNA interactions. (C) Amino acid sequences of NR4A2-DBD. Arrangement of residues 261–348 of NR4A2 shows the classic nuclear receptor-type zinc finger motif and a C-terminal extension (CTE). The red and orange letters correspond to the P-box and D-box, respectively. The blue and purple letters correspond to the T-box and A-box, respectively. Solid-line boxes indicate α-helical segments. The residues in bold and italic are involved in dimerization.

Figure 3. Detailed protein-base contacts in the NR4A2-DBD/IR5 structure. (A) Stereo diagram of major groove recognition by helix H1 of NR4A2-DBD. (B) Stereo diagram of minor groove recognition by the C-terminal extension of NR4A2-DBD. The color is shown in Figure 2. All hydrogen bonds are shown as black dotted lines. (C) Summary of protein-DNA interactions between the NR4A2-DBD and the upstream half-site DNA generated by NUCPLOT. Hydrogen bonds are shown as black dotted lines. Van der Waals interactions are shown as red dotted lines. Asterisks indicate residues that are on the plot more than once. W means water molecule.

Figure 4. Importance of protein-protein interactions and protein-DNA interactions in the NR4A2-DBD/IR5 structure. (A) Surface representation of the dimerization interface in two NR4A2-DBD molecules. (B) Detailed stereodiagram of the residues (shown as sticks) involved in the dimer interface. (C) The DNA binding ability of the wild type (WT) and V298K variant of NR4A2-DBD was measured by EMSA. The concentrations of the DNA and proteins used here were both 45 µM. (D) Binding features of NR4A2-DBD with different IR5 DNA variants. The core sequences of the IR5 DNA duplex used are listed. The arrow shows the orientation of each core half-site. The bases in red and bold are involved in the base-protein interaction and are mutated in the EMSA. The concentrations of DNA used here were 45 µM, and the proteins were used at 90 µM.

Figure 5. Analysis of the NR4A2/RXR heterodimer. (A) Model of the NR4A2/RXR heterodimer
on IR5. The RXR structure was obtained from a previously reported RXR/DR1 structure (PDB identification: 4CN2) and superimposed onto our NR4A2/IR5 complex structure. The RXR and NR4A2 structures were aligned to create the NR4A2/RXR/DNA structural model. (C) DNA binding ability of the NR4A2/RXR heterodimer bound to DNAs.

**Figure 6.** Occurrence of the IR5 and ER0 binding motifs in NR4A-binding sites in human cells. (A) Analysis of the ratio of motifs IR5 and ER0 in the NR4A ChIP-seq data. (B) Representative gene promoters near NR4A-binding sites containing motif IR5.
Figure 1. Binding ability of NR4A2-DBD with different DNAs.
Figure 2. Overall structure of NR4A2-DBD/DNA complexes
Figure 3. Detailed protein-base contacts in the NR4A2-DBD/IR5 structure.
Figure 4. Importance of protein-protein interactions and protein-DNA interaction in the NR4A2-DBD/IR5 structure
Figure 5. Analysis of the NR4A2 and RXR heterodimer
Figure 6. Occurrence of the IR5 and ER0 binding motif in NR4A binding sites in human cells.
Structural basis of binding of homodimers of the nuclear receptor NR4A2 to selective Nur-responsive DNA elements
Longying Jiang, Shuyan Dai, Jun Li, Xujun Liang, Lingzhi Qu, Xiaojuan Chen, Ming Guo, Zhuchu Chen, Lin Chen, Hudie Wei and Yongheng Chen

J. Biol. Chem. published online November 13, 2019

Access the most updated version of this article at doi: 10.1074/jbc.RA119.010730

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts