Published online 18 January 2023

Nucleic Acids Research, 2023, Vol. 51, No. 3 ***1443–1457***
https://doi.org/10.1093/nar/gkac1259

Structures of human $TR4_{LBD}$–JAZF1 and $TR4_{DBD}$–DNA complexes reveal the molecular basis of transcriptional regulation

Yunlong Liu[1], Lulu Ma[1], Min Li[2], Zizi Tian[1], Meiting Yang[1], Xi Wu[1], Xue Wang[1], Guohui Shang[1], Mengjia Xie[1], Yiyun Chen[3], Xin Liu[1], Lun Jiang[1], Wei Wu[1], Chaoqun Xu[4], Liqun Xia[5], Gonghui Li[5], Shaodong Dai[6] and Zhongzhou Chen[1,*]

[1]State Key Laboratory of Agrobiotechnology and Beijing Advanced Innovation Center for Food Nutrition and Human Health, College of Biological Sciences, China Agricultural University, Beijing 100193, China, [2]National Protein Science Facility, Tsinghua University, Beijing 100084, China, [3]Department of Biochemistry, University of Colorado, Boulder, CO 80303, USA, [4]Key Laboratory of Bioactive Substances and Resources Utilization of Chinese Herbal Medicines, Ministry of Education, Institute of Medicinal Plant Development, Peking Union Medical College and Chinese Academy of Medical Sciences, Beijing 100193, China, [5]Department of Urology, Sir Run Run Shaw Hospital, Zhejiang University School of Medicine, Hangzhou 310016, China and [6]Department of Pharmaceutical Sciences, Skaggs School of Pharmacy and Pharmaceutical Sciences, University of Colorado Anschutz Medical Campus, Aurora, CO 80045, USA

Received September 14, 2022; Revised November 30, 2022; Editorial Decision December 15, 2022; Accepted January 11, 2023

ABSTRACT

Testicular nuclear receptor 4 (TR4) modulates the transcriptional activation of genes and plays important roles in many diseases. The regulation of TR4 on target genes involves direct interactions with DNA molecules via the DNA-binding domain (DBD) and recruitment of coregulators by the ligand-binding domain (LBD). However, their regulatory mechanisms are unclear. Here, we report high-resolution crystal structures of $TR4_{DBD}$, $TR4_{DBD}$–DNA complexes and the $TR4_{LBD}$–JAZF1 complex. For DNA recognition, multiple factors come into play, and a specific mutual selectivity between TR4 and target genes is found. The coactivators SRC-1 and CREBBP can bind at the interface of TR4 originally occupied by the TR4 activation function region 2 (AF-2); however, JAZF1 suppresses the binding through a novel mechanism. JAZF1 binds to an unidentified surface of TR4 and stabilizes an α13 helix never reported in the nuclear receptor family. Moreover, the cancer-associated mutations affect the interactions and the transcriptional activation of TR4 *in vitro* and *in vivo*, respectively. Overall, our results highlight the crucial role of DNA recognition and a novel mechanism of how JAZF1 reinforces the autorepressed conformation and influences the transcriptional activation of TR4, laying out important structural bases for drug design for a variety of diseases, including diabetes and cancers.

INTRODUCTION

The human testicular receptor 4 (TR4), also known as TAK1, encoded by *NR2C2*, is a member of the orphan nuclear receptors (NRs). TR4 was first cloned in 1994 from humans and rats (1). Two isoforms were identified in humans, and one containing 596 residues was chosen as the canonical sequence. TR4 is widely expressed throughout the body and is particularly concentrated in the testis, cerebellum, prostate and hippocampus (2).

TR4 shares the typical structural features of NRs, containing an N-terminal domain (NTD), a DNA-binding domain (DBD), a hinge region and a ligand-binding domain (LBD). The conserved DBD determines the specific DNA-binding property of AGGTCA repeats spaced by 0–6 nucleotides (direct repeats: DR0–DR6) (3). Understanding the structural mechanism of how TR4 recognizes and binds to different DNAs is essential to uncover the selective activation or inhibitory mechanisms. There is no structural information on the TR4–DNA complex. Interestingly, the LBD is involved in ligand binding, receptor dimerization, nuclear translocation and transcriptional activation (4–8). TR4 plays important roles in many normal physiological processes, such as fertility (9), glucose and lipid metabolism (10,11), bone metabolism (12), erythroid differentiation (13) and oxidative stress (14). Moreover, TR4 affects the progression and therapy of cancers, including prostate

*To whom correspondence should be addressed. Tel: +86 10 62734078; Fax: +86 10 62734078; Email: chenzhongzhou@cau.edu.cn

cancer (15), renal cell carcinoma (16), seminoma (17), hepatic cell carcinoma (18), non-small cell lung cancer (19) and cervical cancer (20). Since the important roles of TR4 in tumors have been revealed, it is critical to find its activators and repressors in order to develop cancer therapy. However, the autorepressed conformation (21) limits information on its small molecule modulators. Several polyunsaturated fatty acids (PUFAs) and their metabolites (22), and the synthetic peroxisome proliferator-activated receptor gamma (PPARγ) agonist rosiglitazone (10), might work as regulators of TR4. Retinol and its derivatives were identified as novel groups of TR4 activators (21). Moreover, several FDA-approved drugs including genistein, nilotinib and bexarotene (23) were found to regulate TR4 transcriptional activity dose dependently. In addition, several phosphorylation regulators (24–26) were also reported to inhibit TR4-induced transcriptional activation. Despite some progress in developing TR4 regulators, no specific drug has been reported yet.

NRs regulate the expression of target genes through interacting with intermediary proteins (27). Only one coactivator, p300/CREB-binding protein-associated factor (PCAF), was identified to promote the transcriptional activation of TR4 independent of its LBD (28). Although TR4-associated protein (TRA16) is another interactor of TR4 in the absence of an LBD, its interaction suppressed the transcriptional activation via inhibiting DNA binding (29). As reported previously, many coregulators, such as SRC-1, PGC1α and RIP140, could bind NR LBDs with their LXXLL motifs (30–32). Although the interface of these coregulators on LBD is occupied by the activation function region 2 (AF-2) of TR4, some coregulators still interact with the TR4 LBD. According to previous studies (33), AF-2 is critical for the interaction with JAZF1 (TIP27), as the deletion of AF-2 causes an 85% reduction in binding. As a repressor, JAZF1 may affect the coactivators' binding to TR4 rather than inhibiting its homodimerization or DNA-binding ability. JAZF1–TR4 was reported to be involved in the regulations of obesity and diabetes (34), di(2-ethylhexyl) phthalate (DEHP)-induced reproductive toxicity (35), cardiac malformation (36) and atherosclerosis (37). The development of specific peptide drugs targeting NRs is still preliminary. For example, the original corepressor NCoR/SMRT can only inhibit non-specifically. As a TR4-specific repressor, JAZF1 is of great significance in the therapy of TR4-related diseases; the molecular details of the interaction will help to promote the development of TR4-specific peptide drugs.

In this study, we report four crystal structures with three complex structures solved. A novel tail helix of TR4 and its complex with corepressor JAZF1 in the NR family were determined for the first time, and an unidentified interface was found. We also uncovered the specific mechanism of how JAZF1 suppresses TR4-induced transcriptional activation, and found that multiple factors contributed to the recognition of DNA. In addition, a novel mode of DBD dimerization on DNA was observed. We also found the mutual selectivity of the interaction between TR4 and target DNA and determined the characteristics of the target DNA as well. By oligomeric analysis, we demonstrated that the DBD and LBD share the contribution to TR4 homodimerization. Furthermore, cancer-associated mutations affect the interactions and the transcriptional activation activity of TR4 *in vitro* and *in vivo*. Overall, our results provided extensive information on the structural basis to promote drug design for many diseases, such as diabetes and cancers.

MATERIALS AND METHODS

Expression and purification

DNA encoding the wild-type TR4 and JAZF1 was amplified from human cDNA. Fragments encoding various amino acid fragments were amplified by polymerase chain reaction (PCR) and ligated into pET-28a (Novagen) as well as a modified pET28a vector encoding a 6× His-maltose-binding protein (MBP) fusion tag and a tobacco etch virus (TEV) protease cleavage site. Mutants of TR4 and JAZF1 were generated by site-directed mutagenesis and verified by DNA sequencing. The recombinant plasmid containing the target gene was transformed into *Escherichia coli* strain BL21 (DE3) cells. Cells were then grown in Luria–Bertani (LB) medium at 37°C until the OD_{600} reached 0.8–1.0 and were induced overnight at 18°C by adding 0.2 mM isopropyl-β-D-1-thiogalactopyranoside and 0.2 mM $ZnSO_4$. The cultures were harvested by centrifuging at 4000 *g* for 15 min, and the cell pellets were resuspended in lysis buffer (20 mM Tris–HCl pH 7.3, 1 M NaCl, 2 mM β-mercaptoethanol) replenished with 0.1% (v/v) Triton X-100 and 1 mM phenylmethylsulfonyl fluoride (PMSF; Invitrogen). Cells were lysed by sonication and clarified by centrifugation at 18 300 *g* for 35 min. The lysate was filtered through a 0.45 μm filter membrane to remove cell debris and other impurities before being loaded onto an Ni^{2+}-chelating column. The TEV protease was added to cleave the 6× His-MBP fusion tag at a 1:10 (w/w, protease/protein) ratio overnight at 4°C. Further purification was performed by size-exclusion chromatography (SEC) using a Superdex 200 column (GE Healthcare) with SEC buffers: 20 mM Tris–HCl pH 7.3, 100 mM NaCl, 2 mM β-mercaptoethanol for $TR4_{DBD}$ (residues 113–189) and $TR4_{DBD\text{-}CTE}$ (residues 113–200, CTE: C-terminal extension of the DBD), and 20 mM Bis-Tris propane pH 8.5, 1 M NaCl, 2 mM β-mercaptoethanol for the $TR4_{LBD}$ (residues 341–596)–$JAZF1_{TID}$ (residues 51–75) chimera. Peak fractions were collected and examined using Coomassie Blue-stained sodium dodecylsulfate (SDS)–polyacrylamide gels. The protein was concentrated to 5 mg/ml for crystallization.

Crystallization and data collection

Crystallizations of $TR4_{DBD}$, and $TR4_{LBD}$–JAZF1, $TR4_{DBD}$–DNA and $TR4_{DBD\text{-}CTE}$–DNA complexes were all performed at 16°C with multiple commercial crystallization kits in 24-well plates using the hanging-drop vapor diffusion method. The conditions were optimized, and high-quality crystals were finally obtained in 2.4 M malonate pH 6.5 for $TR4_{DBD}$ and 1.6 M $(NH_4)_2SO_4$, 0.1 M HEPES pH 7.5, 0.2 M NaAc and 5% MPD for the $TR4_{LBD}$–$JAZF1_{TID}$ chimera. Two partial complementary oligonucleotides (5′-GGCAGAGGTCAAAGGTCA-3′ and 5′-CTGACCTTTGACCTCTGC-3′) were used to

obtain the 18 bp duplex DNA with single-nucleotide overhangs. The $TR4_{DBD}$–DNA and $TR4_{DBD\text{-}CTE}$–DNA complexes were obtained by a second SEC purification with a 3:1 (protein:DNA) ratio incubated overnight. The incubated complex was further purified by SEC (Superdex 200, GE Healthcare) and concentrated to 8 mg/ml. The $TR4_{DBD}$–DNA complex was obtained in 30% polyethylene glycol (PEG) 4000, 0.1 M sodium citrate pH 6.0, 0.2 M NH_4Ac. The $TR4_{DBD\text{-}CTE}$–DNA complex appeared in many solutions containing PEG and finally grew to high-quality crystals in 20% PEG8000, 0.1 M Tris pH 8.0, 0.2 M $(NH_4)_2SO_4$. All crystals were cryoprotected in the mother liquor supplemented with 25% (v/v) glycerol and flash-frozen in liquid nitrogen. The data were respectively collected on beamlines BL02U1, BL10U2 and BL18U1 at Shanghai Synchrotron Radiation Facility (SSRF). Data collection and processing statistics are shown in Supplementary Table S1.

Structure determination and refinement

All diffraction data were indexed, integrated and scaled with the HKL2000 suite of programs (38). For the Holo-$TR4_{DBD}$ structure, using the structure of the PPARγ DBD (PDB 3DZU) (39) as the search model, molecular replacement solutions were found using the BALBES server (40). The crystal contains two proteins per asymmetric unit and belongs to space group $P4_1$. After 10 more cycles of manual rebuilding by COOT (41) and refinement with REFMAC5 (42), the structure was refined to 1.6 Å with an R_{work} of 17.6% and R_{free} of 18.1%. To solve the $TR4_{DBD}$–DNA and $TR4_{DBD\text{-}CTE}$–DNA complexes, we built a search model combined with the TR4 DBD and core motifs of 13 bp DNA from PDB 3DZU (39), and molecular replacement solutions were found using PHASER (43) and MOLREP (44). The details of the $TR4_{DBD}$–DNA and $TR4_{DBD\text{-}CTE}$–DNA complexes are shown in Supplementary Table S1. For the $TR4_{LBD}$–JAZF1 complex structure, we used the apo-TR4 LBD (PDB 3P0U) (21) as the search model. An initial molecular replacement solution was obtained from the BALBES server. The $TR4_{LBD}$–$JAZF1_{TID}$ chimera belongs to space group $P4_322$ and there is one TR4 LBD and one JAZF1 TID protein in the asymmetric unit. After the rebuilding by COOT (41) and refinement with REFMAC5 (42) for several cycles, the final structure was refined to 1.86 Å resolution with a R_{work}/R_{free} value of 19.6%/21.9%.

All structural figures in this article were prepared using PyMOL (45) (version 2.3.4, Schrödinger LLC).

MBP pull-down assay

All the constructs fused with 6× His-MBP or 6× His-tag were first purified using the appropriate affinity columns. For the coactivator experiments, 3 μM 6× His-MBP–$TR4_{LBD}$ or 6× His-MBP–$TR4_{LBD}$–$JAZF1_{TID}$ were incubated with 12 μM 6× His-$TR4_{DBD}$–$SRC\text{-}1_{1434–1441\,(LLQQLLTE)}$, 6× His-$TR4_{DBD}$–$CBP_{68–77\,(KQLSELLRGG)}$ and 6× His-$TR4_{DBD}$–$CBP_{356–365\,(QQLVLLLHAH)}$, respectively. For the cancer-associated mutation experiments, 4 μM 6× His-MBP–$JAZF1_{39–79}$ was incubated with 5 μM TR4 LBD. The incubation was at 4°C in 500 μl of a buffer of 20 mM HEPES pH 7.0, 1 M NaCl, 2 mM β-mercaptoethanol and 0.1% (v/v) Triton X-100 in the presence of Amylose Resin beads (NEB) for 6 h. The resin was extensively rinsed with the same buffer to remove unbound or non-specifically bound proteins. Proteins left on the beads were separated by SDS–polyacrylamide gel electrophoresis (PAGE) and analyzed by western blotting.

Electrophoretic mobility shift assays (EMSAs)

The protein used was $TR4_{113–200}$. DNA (5 μM) was incubated with protein (10 μM) in 20 mM HEPES buffer (pH 7.9), 50 mM KCl, 2.5 mM $MgCl_2$, 1 mM dithiothreitol (DTT), 10% (v/v) glycerol at 4°C. After 1 h incubation, 4 μl of the mixture was separated on a 6% native polyacrylamide gel in 1× TB buffer (45 mM Tris pH 8.3, 45 mM boric acid) at 180 V for ∼30 min. The gel was visualized by staining with GelRed and scanned with the ChemiDoc MP Imaging System (BIO-RAD) at 312 nm.

Microscale thermophoresis

Microscale thermophoresis (MST) measured the binding affinities between double-stranded DNA (dsDNA) oligonucleotides and $TR4_{113–200}$ (wild-type and mutations). The sequence of the wild-type DR1 dsDNA used is GGCA GAGGTCAAAGGTCAAACGT. The TR4 proteins fused with the 6× His-tag were bound to the RED-tris-NTA 2nd Generation dye. For binding assay, 10 μl of 100 nM labeled proteins in a buffer containing 20 mM HEPES pH 7.5, 100 mM NaCl, 1 mM DTT, 0.15% (v/v) Nonidet P-40 (NP-40), 0.05% (w/v) PEG8000 was mixed with DNA solutions with different concentrations prepared by consecutive 2-fold dilution in the same buffer. After incubation for 30 min at room temperature, the samples were loaded into silica capillaries and temperature-induced fluorescence changes were measured on a Monolith™ (NanoTemper) at 25°C by using 60% LED power and 40% infrared laser power. Data analyses were performed by using the NTAnalysis software (NanoTemper Technologies). The data contained three independent measurements using different protein preparations. The fitting function is derived from the law of mass action:

$$\text{fraction bound} = \frac{[AB]}{[B]} = \frac{[A]+[B]+K_D-\sqrt{([A]+[B]+K_D)^2-4[A][B]}}{2[B]}$$

where [AB] = concentration of bound complex; [A] = concentration of total ligand molecules; [B] = concentration of total target molecules; and K_D = dissociation constant.

Small-angle X-ray scattering

Small-angle X-ray scattering (SAXS) measurements were performed on beamline BL19U2 at the SSRF following previously published methods (46,47). Briefly, all proteins were subjected to SEC with a buffer containing 20 mM HEPES pH 7.5, 80 or 200 mM NaCl and 1 mM DTT. Various concentrations of protein or complex were used, and the data

were collected at 1.03 Å with a distance of 2.68 m from the detector. Individual data were processed by RAW (48). The scattering data from the buffer alone were measured before and after each sample measurement, and the average scattering data were used for background subtraction. Comparison of the scattering of $TR4_{DBD}$, $TR4_{LBD}$–$JAZF1_{TID}$ chimera and $TR4_{DBD\text{-}CTE}$–DNA complex structures with SAXS experimental data were computed with FoXS (49).

DNA preparation

All DNA oligonucleotides were synthesized by Shanghai Generay. DNA duplexes used for crystallization, EMSA and MST experiments were first annealed in a buffer of 20 mM HEPES pH 7.5 and 100 mM NaCl by heating the mixture at 95°C for 5 min and slowly cooling to 12°C for 46 min.

DNA shape analysis and electrostatic potential calculation

The groove width and electrostatic potential of the minor groove for dsDNA in the crystal structure were calculated using the CURVES + ver.1.31 (50) and DNAshape server (51), respectively. The minor groove width of the free DNA predicted was calculated using the DNAshape server (51).

Cell culture and reagents

The COS-7 cell line was purchased from the Cell Bank of Type Culture Collection of the Chinese Academy of Sciences. COS-7 cells were cultured in Dulbecco's modified Eagle's medium (DMEM) (Gibco) supplemented with 10% fetal bovine serum (FBS; Cellmax, Beijing, China) and penicillin–streptomycin (Gibco). The cells were incubated at 37°C and in a 5% CO_2 atmosphere.

Plasmid transfection and reporter assays

Upon reaching 70–90% confluence in 24-well plates, the COS-7 cells were transfected with 250 ng of pGL3-POMC-Luc plasmid containing the ∼1800 bp rat POMC (proopiomelanocortin) promoter–luciferase reporter described previously (52). At the same time, 25 ng of PRL-TK renilla control plasmid, 250 ng of pcDNA3.1/$TR4_{1–596}$ or pcDNA3.1 vector and an additional 250 ng of pcDNA3.1/$JAZF1_{1–243}$ in some treatments were transfected by using Lipofectamine 3000 (Invitrogen). After 5 h of transfection, the medium was replaced with a fresh 4% FBS medium, and the cells were incubated for 24 h in total before the dual-luciferase assay. Luciferase activity was measured with a Dual-Glo® Luciferase Assay System (Promega) according to the manufacturer's instructions.

RESULTS

JAZF1 binds TR4 at a novel interface

A previous study indicated that JAZF1 interacts with the C-terminal region of TR4 via its TR4-interacting domain (TID; residues 39–79) (33). In this study, $TR4_{341–596}$ and full-length JAZF1 (residues 1–243) were cloned and expressed in *E. coli*. TR4 bound stably with full-length JAZF1 or $JAZF1_{TID}$ (Supplementary Figure S1A, B). Despite strong interactions (Supplementary Figure S1C), full-length JAZF1 was subjected to degradation and failed to co-crystallize. Therefore, we truncated JAZF1 into fragments of various lengths to incubate with $TR4_{341–596}$. We crystallized the $TR4_{341–596}$/$JAZF1_{39–79}$ complex (Supplementary Figure S1D) and solved the complex structure at 4.08 Å resolution (Supplementary Figure S1E, F). In the complex structure, the electron density of the fragment of $JAZF1_{39–50}$ was absent due to the lack of interactions with TR4 or with $JAZF1_{51–79}$. Thus, the flexible $JAZF1_{39–50}$ fragment might limit the complex diffraction resolution. To further clearly understand the interaction details between TR4 and JAZF1, we constructed a $TR4_{341–596}$–$JAZF1_{51–75}$ chimera and obtained a 1.86 Å high-resolution structure. The crystal structure was in space group $P4_322$, with $R_{work} = 19.6\%$ and $R_{free} = 21.9\%$ (Supplementary Table S1). In an asymmetric unit, there was one TR4 LBD with well-defined residues 358–595 (the electron density of residues 455–461 was missing) and one JAZF1 with well-defined residues 51–73 as well as 119 water molecules. As expected, the overlay of the $TR4_{341–596}$–$JAZF1_{51–75}$ chimera and $TR4_{341–596}$/$JAZF1_{39–79}$ complex structures showed little difference between the two structures (Supplementary Figure S1G).

The TR4 LBD in the complex structure was composed of multiple α-helices to form a multilayered sandwich structure. JAZF1 (51–75) formed an α-helix and was attached to the upper layer of the sandwich (Figure 1A). Unlike the traditional AF-2 (α12) positions for binding the corepressor (NCoR/SMRT) or the coactivators, JAZF1 stably bound to a novel surface formed by the tail helices α12 and α13 and the crossed helix α3 of TR4 (Figure 1A). The composite simulated-annealing F_o–F_c 'omit' map confirms the presence of a bound JAZF1 (Figure 1B). Upon binding to JAZF1, a specific tail helix α13 from the F domain (53) was stabilized, which had not been observed in the apo-TR4 LBD (21) (Figure 1C) or other NR family members previously, heralding new functions. In addition, the conformation of TR4 changed noticeably; the loops before α3 and α12 shifted at the largest distances of 10.8 Å and 4.5 Å, respectively. Moreover, a new helix α6 was formed, and the electron density of missing residues 446–454 between helices α5 and α7 in apo-$TR4_{LBD}$ became visible after binding JAZF1 (Figure 1C). Despite these obvious changes, the conformations of α1, α3, α4, α5, α7, α8 and α9 were quite similar to the apo-$TR4_{LBD}$, and the autorepressed conformation of the LBD caused by the bend of α10/α11 was also maintained (Supplementary Figures S1H and S2). Taken together, compared with the apo-$TR4_{LBD}$ (21) structure containing only 10 identified α-helices, the resolution has been greatly improved, and the overall conformation of TR4 was more compact and complete after JAZF1 bound at the C-terminus of the LBD (Figure 1A, C).

JAZF1 blocks the coactivators' binding by reinforcing the autorepressed conformation of the TR4 LBD

Traditional coactivators and corepressors utilize their LXXLL motifs to bind to the coregulator-binding surface of NRs. According to our results, JAZF1 bound to a novel coregulator-binding surface formed by TR4 AF-2

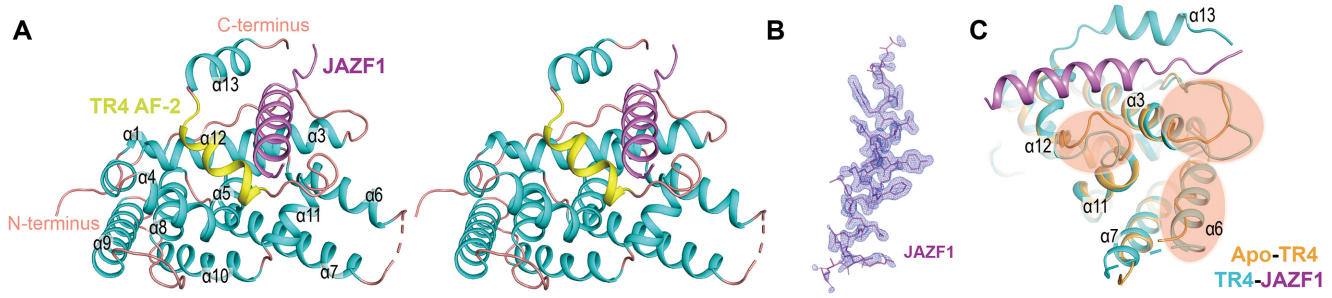

Figure 1. Overall structure of the $TR4_{LBD}$–JAZF1 complex. (**A**) Stereoview of the $TR4_{LBD}$–JAZF1 complex. TR4, TR4 AF-2 (helix α12) and JAZF1 are colored cyan, yellow and magenta, respectively. (**B**) The composite simulated-annealing F_o–F_c 'omit' electron density map of JAZF1 in the complex, contoured at 2.5 σ. (**C**) Structural overlay of apo-$TR4_{LBD}$ (PDB 3P0U) (21) and the $TR4_{LBD}$–JAZF1 complex shows the conformational changes of the TR4 LBD caused by JAZF1 binding. Obvious changes are indicated by magenta ovals.

(helix α12), helix α3 and tail helix α13 (Figure 1A). Hydrogen bonds and multiple hydrophobic interactions jointly make JAZF1 a strong interactor of TR4 (Figure 2A). More specifically, the binding involves the hydrogen bonds between JAZF1 Asp67 and TR4 Tyr578. Moreover, JAZF1 Leu58, Ala57 and Val56 formed hydrogen bonds with TR4 Glu388. Gln52, Thr54, Tyr55 and Tyr60 of JAZF1 and Pro374, Gln382, Arg392, Gln590 and Ser595 of TR4 were also involved in hydrogen bonding. Interestingly, a water molecule stabilized the hydrogen bond interaction between TR4 Pro374, Arg392 and JAZF1 Thr54, forming a stable interaction network. In addition, the crystal packing showed more interaction details (Supplementary Figure S3). Both termini of JAZF1 formed hydrogen bonds with the symmetric TR4 LBDs, and the helix α13 of TR4 also formed hydrogen bonds with another symmetric helix α13, jointly contributing to the stability of the crystal. In addition, the formation of five hydrophobic interfaces of TR4 also plays a crucial role in the interaction with JAZF1. Pro53, Tyr55, Leu58, Ile61 and Met65 of JAZF1 bound at two hydrophobic interfaces formed by helix α3 and the loops before α3, as well as before α12, respectively. Likewise, Val56, Tyr60 and Phe64 of JAZF1 bound at another two hydrophobic interfaces on the other side formed by the C-terminus of helix α12, helix α13 and Phe395 of helix α3. Meanwhile, Ala68 and Ala69 of JAZF1 bound at the interface formed at the N-terminus of helix α12. Overall, JAZF1 stably binds at a novel interface of TR4 through the formation of hydrogen bonds and hydrophobic interactions. All the residues of JAZF1 involved in binding are conserved among the species (Figure 2B). Interestingly, the interacting residues of TR4 are variable in the NR superfamily (Figure 2C), while most of them are similar in TR2, indicating that JAZF1 is a TR4-specific suppressor and may have effects on TR2 as well.

Based on structural superposition, the space of the coactivators' binding sites of the TR4 LBD is occupied by its AF-2 helix in apo-$TR4_{LBD}$ and the $TR4_{LBD}$–JAZF1 complex (Figure 2D). The sequence of the TR4 AF-2 helix showed that the core sequence of AF-2 is similar to the LXXLL motifs of the coactivators, in which the leucines might be functionally replaced with isoleucines, binding at the coactivators' interface and forming a natural autorepressed conformation. Both AF-2 and AF-2-binding sites of the LBD are conserved in the NR superfamily (Figure 2C). However, the results of the pull-down assay showed that the LXXLL motifs of the coactivators SRC-1 and CREBBP (CBP) could hijack the AF-2-binding sites (Figure 2E), while the interactions were mostly blocked when TR4 was bound by JAZF1 (Figure 2E). Combined with our solved structure, JAZF1 reinforced the AF-2 helix at the coactivators' interface by binding to the helices α3, α12 and α13, and thus suppressing the regulation of TR4-induced transcriptional activation by coactivators SRC-1 and CBP.

The complex structures of DBD–DNA and C-terminal-extended DBD–DNA

The highly conserved DBD plays a crucial role in transcriptional regulation via DNA binding. Here, the TR4 DBD (residues 113–189) was constructed and expressed in *E. coli*. We obtained a 1.6 Å high-resolution Holo-$TR4_{DBD}$ crystal structure (Figure 3A) containing two DBD molecules per asymmetric unit. Each DBD contains two conserved C4-type zinc fingers: ZnF-I (Cys117, Cys120, Cys134 and Cys137) at the N-terminus of the DBD, and ZnF-II (Cys153, Cys159, Cys169 and Cys172) located inside the DBD (Figure 3B, C). The outward extensions of both termini of DBD confer the ability to bind dsDNA independently and stably.

TR4 binds to AGGTCA DNA sequence motifs in direct repeat (DR) orientation with a variable number of spacer nucleotides to regulate target genes such as CD36, stearoyl-CoA desaturase 1 (SCD1), proopiomelanocortin (POMC), phosphoenolpyruvate carboxykinase (PEPCK) and apolipoprotein E (ApoE) (10,11,25,55–56). We screened the binding of a series of dsDNAs ranging from 13 to 35 bp with two AGGTCA half-sites separated by one nucleotide (DR1) using MST experiments. Although the MST results (Supplementary Figure S4A) indicated that the dsDNA of 23 bp had the highest binding affinity for TR4, the quality of the crystal was poor. Luckily, we finally obtained high-quality complex crystals with an 18 bp DR1 fragment (K_D = 38.3 μM) and solved a 3.2 Å resolution $TR4_{113–189}$–DNA complex structure (Figure 3D). In the $TR4_{DBD}$–DNA structure, every half-site of dsDNA bound a DBD molecule. In each DBD molecule, the conserved ZnF-I played a crucial role in DNA recognition. Helix I (HI) in ZnF-I inserted into the major groove of DNA to recognize and bind target genes (Figure 3D). We also obtained a 2.3 Å high-resolution structure of the $TR4_{DBD-CTE}$–

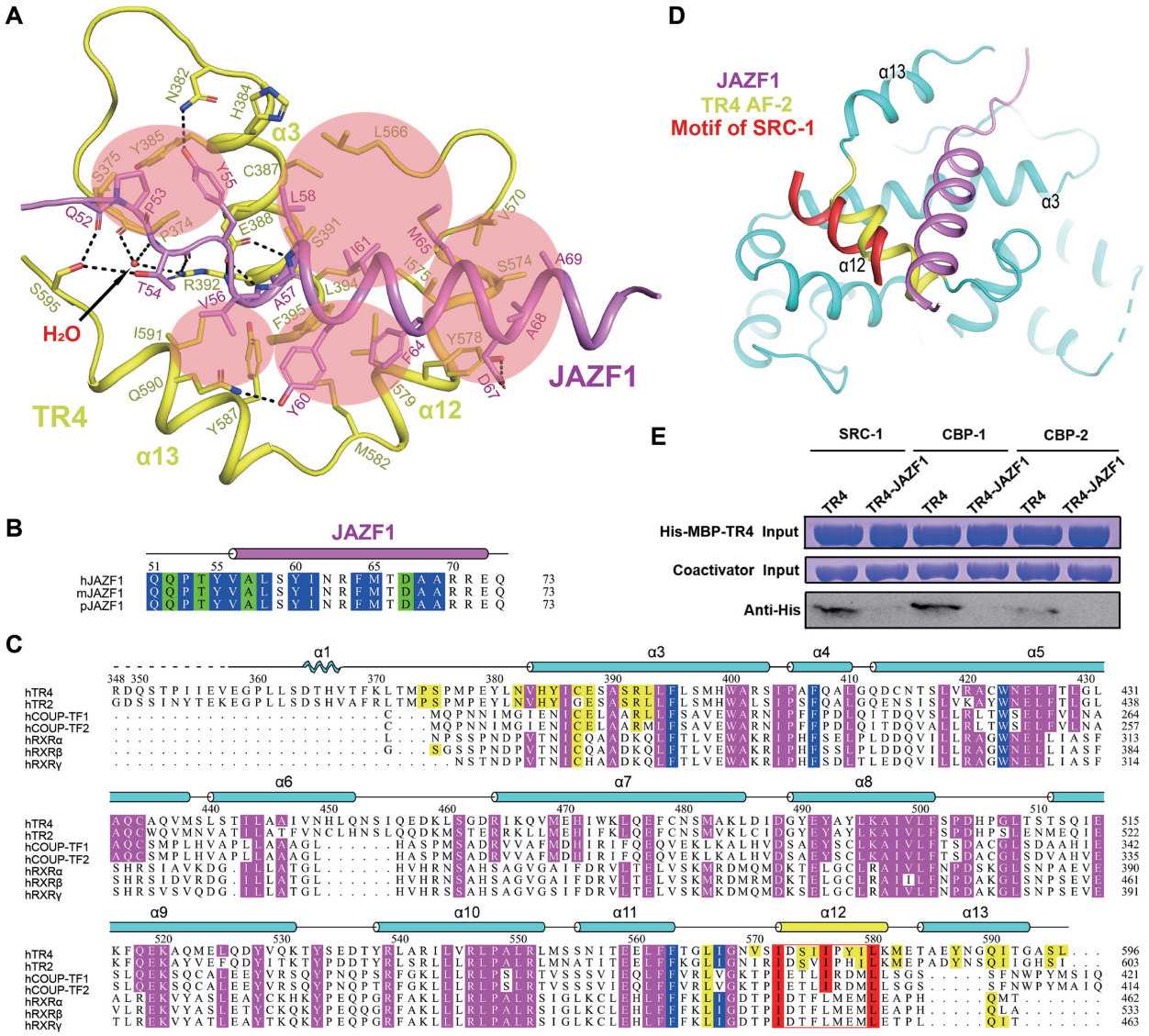

Figure 2. JAZF1 suppresses the binding of coactivators to TR4. (**A**) Detailed representation of interactions between the TR4 LBD (yellow) and JAZF1 (magenta). Interacting residues are shown as sticks; hydrophobic bindings, pink ovals; hydrogen bonds, dashed lines; and water molecules, red sphere. (**B**) Structure-based sequence alignments of JAZF1 from different species reveal a conserved TR4-interacting motif. hJAZF1, JAZF1 from *Homo sapiens* (Q86VZ6); mJAZF1, JAZF1 from *Mus musculus* (Q80ZQ5); pJAZF1, JAZF1 from *Pongo abelii* (Q5RDF5). (**C**) Structure-based sequence alignments of the LBDs of TR4 and other NR family members. The NRs used are all from *Homo sapiens*. Secondary structural elements of the $TR4_{LBD}$–JAZF1 structure are calculated using DSSP and colored in cyan and yellow (AF-2, α12). The conserved residues are colored purple, red (AF-2) and blue (AF-2 binding sites). The residues involved in JAZF1 binding are colored yellow. (**D**) Overlay of the $TR4_{LBD}$–JAZF1 complex structure and its AF-2 helix (yellow) with the SRC-1 LXXLL motif (red) from the RXRα structure (PDB 1xv9) (54). (**E**) JAZF1 suppresses the binding of coactivators to TR4. The TR4 DBD is fused with the LXXLL motifs of coactivators (32), $SRC\text{-}1_{1434–1441}$, $CBP_{68–77}$ (CBP-1) and $CBP_{356–365}$ (CBP-2), respectively. Meanwhile, TR4–JAZF1 input is the fusion of 6× His-MBP, $TR4_{341–596}$ and $JAZF1_{51–75}$.

DNA complex using the same 18 bp DR1 fragment with a longer length DBD (residues 113–200) (Figure 3E). The composite simulated-annealing F_o–F_c 'omit' map showed the presence of a bound dsDNA and a C-terminal extension of the DBD (Figure 3E). Interestingly, the electron density of the CTE (residues 190–196) could only be seen in the DBD bound at the 3′ half-site. The interface analysis between each DBD molecule and dsDNA suggested that the existence of CTE induced a 114 Å^2 increase in the interface area (Supplementary Table S2), which might lead to an improvement of DNA binding ability (K_D = 16.4 μM) (Supplementary Figure S4B). Furthermore, the resolution of the complex was also improved (from 3.2 Å to 2.3 Å) (Supplementary Table S1).

Multiple factors contribute to the recognition of DNA

The total buried interface between $TR4_{DBD\text{-}CTE}$ and dsDNA reaches 1617.5 Å^2 (Supplementary Table S2), indicating strong interactions. According to the PISA analysis, >41 hydrogen bonds were formed between $TR4_{DBD\text{-}CTE}$ and dsDNA (Figure 4A). Nearly half of the hydrogen-bonding

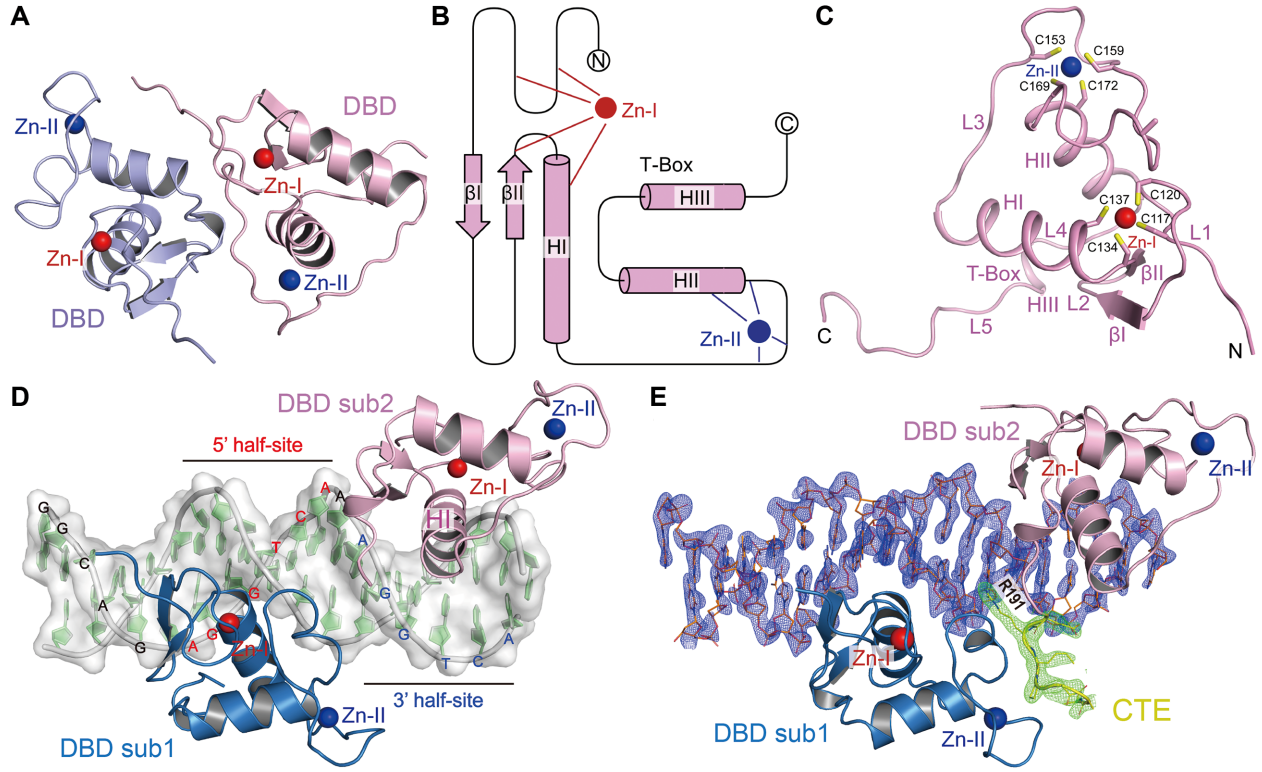

Figure 3. Structural characterization of $TR4_{DBD}$ and $TR4_{DBD}$–dsDNA complexes. (**A**) Crystal structure of $TR4_{DBD}$. Zn^{2+} ions in the structure are shown as spheres and colored red (Zn-I) and blue (Zn-II). (**B**) Topology of the TR4 DBD (pink) with its zinc fingers (red and blue). (**C**) Overall structure of $TR4_{DBD}$. The residues coordinating the Zn^{2+} ions are shown as sticks. (**D**) Structure of the $TR4_{113–189}$–dsDNA complex. DBDs binding at the 5′ and 3′ half-site are colored blue and pink, respectively. The DNA sequence containing direct repeats of the canonical AGGTCA half-site separated by one nucleotide (DR1) is shown. (**E**) Structure of the $TR4_{113–200}$–dsDNA complex. The CTE of TR4 DBD is colored yellow. The composite simulated-annealing F_o–F_c 'omit' electron density maps of CTE and dsDNA are shown at 2.5 σ.

interactions occurred on the phosphodiester backbone, involving the side chains of Arg127, Tyr129, Lys142, Arg146, Arg166, Asn167, Gln170, Arg173 and Gln188. In addition, there was also a specific recognition network between $TR4_{DBD\text{-}CTE}$ and DNA bases. Firstly, there were direct hydrogen-bonding interactions between Glu135, Lys138, Lys142 and Arg143, and bases G_2, G_3 and C_5 of the $A_1G_2G_3T_4C_5A_6$ half-site. Secondly, many water molecules mediated the formation of hydrogen bonds between the TR4 side chains and DNA bases (Figure 4A), forming multiple and stable recognition networks between the residues Tyr129, Glu135, Lys138, Lys142, Arg143, Arg191 and all the bases except A_6 in the repeated $A_1G_2G_3T_4C_5A_6$ half-sites (Figure 4B). Glu135, Lys138, Lys142 and Arg143 can form hydrogen bonds with DNA both independently and by water molecules. Of these, the residues Glu135, Lys138 and Arg143 are conserved in RAR (57), EcR (58), LXRβ (59) and VDR (60) and could form direct hydrogen-bonding interactions with DNA bases, while Lys142 was replaced by arginine and could only form hydrogen bonds mediated by water. Interestingly, though K138 and K142 are conserved in RXR, whether they can directly recognize DNA bases varies among the DNA targets (57,61).

Hydrogen-bonding interactions play an important role in DNA recognition. By further analyzing the complex structure, we found that the side chains of Gln188 and Arg191 were anchored in the minor groove of DNA (Figures 3E and 4C), which was also observed in the RXR–DNA complex (39). The minor groove regions bound by the Gln188 and Arg191 side chains were narrower than their adjacent regions, with a minimum width of 5.1 Å (versus >6.6 Å for other parts) (Figure 4C). The narrower minor groove might be an intrinsic structural feature of the DNA sequence itself [the A-rich sequence provides a narrower minor groove width (62)] (Figure 4C, dashed line). Theoretically, the negative electrostatic potential in the minor groove is enhanced as the groove width decreases. The electrostatic potential of the narrow minor groove bound by the side chain of Arg191 is significantly more negative than those of the adjacent minor grooves (Figure 4C, red line). Thus, the positively charged Arg191 favorably binds to the intrinsically narrow minor groove with enhanced negative electrostatic potential. This observation suggests that the binding of Gln188 and especially Arg191 side chains to the target DNA is a DNA shape readout mechanism. Arginine, which is positively charged and abundant on the protein surface, was confirmed to have an extremely prominent contribution to the shape readout mechanism in previous studies (63). As further confirmed by the deletion and mutagenesis results (Figure 4D, E; Supplementary Figure S5), this type of recognition by Arg191 plays a crucial role in enhancing the binding between TR4 and DNA.

In the $TR4_{DBD\text{-}CTE}$–DNA complex structure, the electron density of CTE (residues 190–196) in the DBD subunit 2

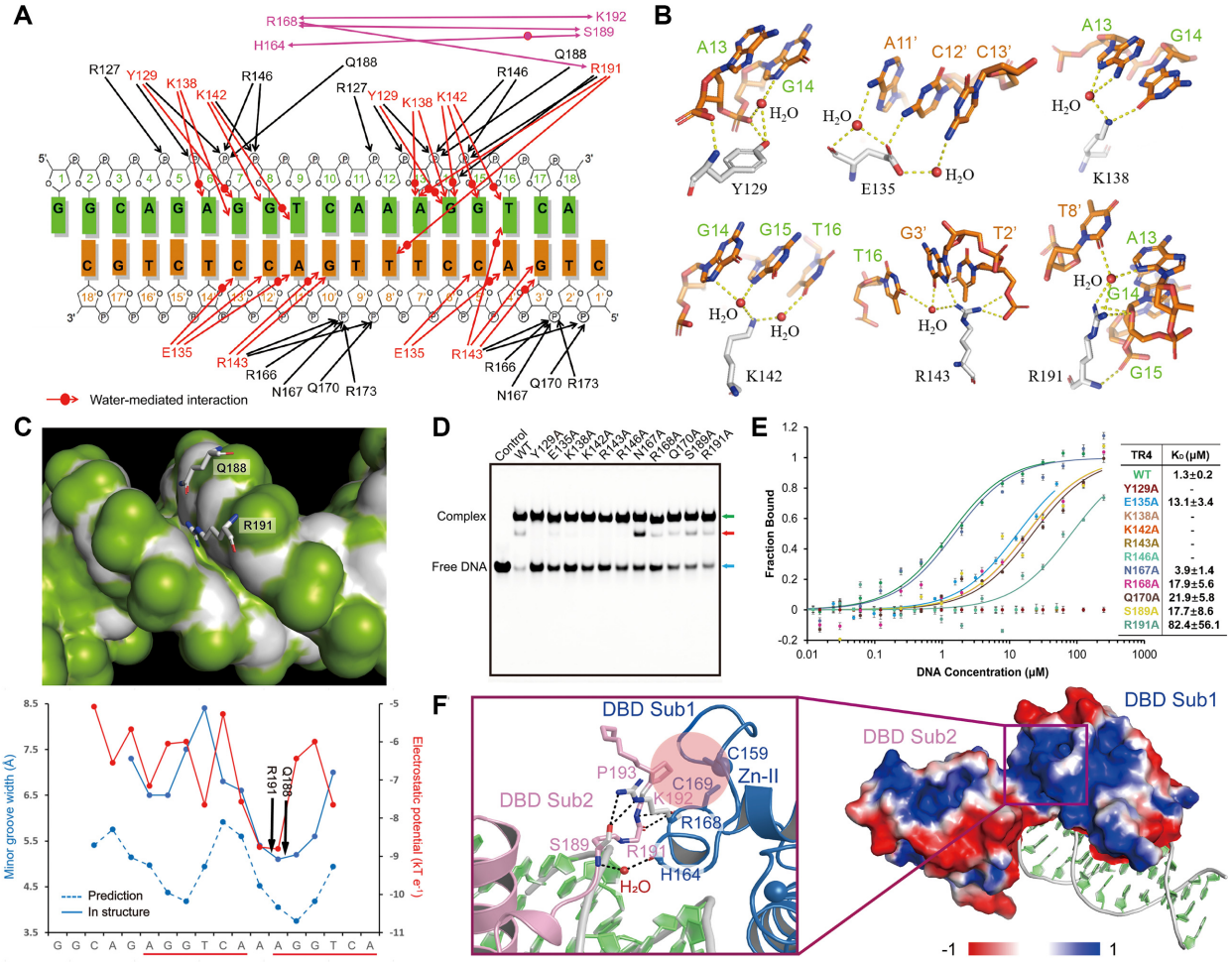

Figure 4. The interaction between $TR4_{DBD\text{-}CTE}$ and dsDNA. (**A**) Schematic representation summarizing the TR4–dsDNA interaction. The sequence of the dsDNA is shown with two complementary strands. An additional base (G/C) appeared at the 5′ end of each strand. The residues involved in the interaction with DNA bases are colored red, with the DNA phosphodiester skeleton colored black and the residues of another $TR4_{DBD\text{-}CTE}$ colored pink, respectively. (**B**) Close-up view of the water-mediated interaction between key residues and DNA bases. Water molecules, red spheres; hydrogen bonds, dashed lines. (**C**) Recognition of the narrow minor groove geometry of DNA. The minor groove width and electrostatic potential are calculated by CURVES + ver.1.31 (50) and DNAshape server (51). The enhanced negative electrostatic potential in the narrower groove region attracts Gln188 and Arg191 side chains through favorable electrostatic interactions, a mechanism known as DNA shape readout. (**D**) EMSA of different $TR4_{DBD\text{-}CTE}$ mutants with the dsDNA. Positions of free dsDNA, blue arrow; $TR4_{(1)}$–dsDNA complex, red arrow; $TR4_{(2)}$–dsDNA complex, green arrow. (**E**) Measurement of the binding affinity of the wild-type and mutated $TR4_{DBD\text{-}CTE}$ with 23 bp DR1 dsDNA by MST. The experiments were repeated three times. (**F**) Cooperativity of dimerization enhances DNA binding. The dimeric interfaces that involve residues of the second Zn module of DBD Sub1 and the CTE of DBD Sub2 are shown as sticks. Hydrogen bonds, black dashed lines; hydrophobic interaction area, pink oval. Right: electrostatic surface representation of the complex (red, negative; blue, positive; light gray, neutral).

(Sub2) bound at the 3′ half-site of DNA was rather clear while that in Sub1 bound at the 5′ half-site of DNA was completely missing (Figure 3E). To determine the reason for this, we performed structural analysis. Hydrogen bonds were formed between Arg168 of DBD Sub1 and Ser189, Arg191 and Lys192 of DBD Sub2 (Figure 4F). In addition, water molecules mediated the formation of the hydrogen bonds between His164 of Sub1 and Ser189 of Sub2 (Figure 4F, A). In addition, Pro193 of the Sub2 CTE bound at the hydrophobic interface formed by Sub1 ZnF-II. Thus, the CTE of TR4 DBD Sub2 was stabilized by another DBD. Although part of the CTE was already described as interacting with dimeric protein in other NR–DR1 structures (39,57), it had not been reported that CTE formed hydrophobic interactions with the ZnF-II of another DBD (Supplementary Figure S6A). Moreover, the key residue Pro193 is not conserved among NRs (Supplementary Figure S6B). With the help of the CTE, the two DBDs formed an additional 327.6 $Å^2$ dimeric interface and a larger positive charge area, cooperatively enhancing the binding to DNA (Supplementary Figure S7). Consistently, both the R168A and S189A mutations caused a 12-fold decrease in the ability to bind to dsDNA (Figure 4E), indicating that the cooperativity of the $TR4_{DBD\text{-}CTE}$ dimer prominently enhances its binding to target DNA.

To test the roles of the above key residues, site mutations of $TR4_{DBD\text{-}CTE}$ were generated and EMSA was performed to understand the qualitative characteristics of DNA binding among the $TR4_{DBD\text{-}CTE}$ mutants, while MST experiments were used to determine the equilibrium dissocia-

tion constants. The mutations of Y129A, K138A, K142A, R143A and R146A caused the most obvious damage to the binding. The mutation R191A involved in shape readout also decreased the binding ability by ∼60 times (Figure 4E). In the complex structure, the DNA contains two repeated AGGTCA half-sites, forming a heterotrimeric complex with two DBD molecules. However, a small number of $TR4_{DBD\text{-}CTE}$–DNA heterodimers were also found in the EMSA experiments. The proportion of heterodimers varied with different $TR4_{DBD\text{-}CTE}$ mutants. N167A significantly increased the proportion of heterodimers, while Y129A, K138A, K142A, R143A and R146A decreased the proportion of heterodimers (Figure 4D). Consistently, all mutants had a lower binding affinity for DNA than the wild-type. Taken together, these findings demonstrate that these key residues jointly determine the recognition and multiple factors contribute to the binding of DNA.

Mutual selectivity of the interaction between TR4 and dsDNA

To verify the transcriptionally regulatory function of TR4, we analyzed the potential promoter sequences of some downstream target genes (Supplementary Table S3), and performed MST experiments to determine the equilibrium dissociation constants of TR4 for each target sequence. As a reference, we constructed a 23 bp ideal DR1 response element (idRE) with two canonical AGGTCA half-sites separated by an adenine. According to the results (Figure 5A), TR4 bound each promoter of these target genes, and the binding affinity varied with sequences. The promoters of gene ciliary neurotrophic factor receptor (CNTFRα) and cellular retinol-binding protein type II (CRBPII) had the same half-sites or only one nucleotide replaced compared with idRE, showing better abilities to bind to TR4 (CNTFRα, K_D = 6.8 μM; CRBPII, K_D = 4.1 μM). In contrast, the promoter of the gene stearoyl-CoA desaturase 1 (SCD1) had the largest sequence differences and bound TR4 the most weakly (K_D = 142 μM). In contrast, although the genes α-MHC, S14, LHR, RARβ, SV40, ApoE, CD36 and PEPCK all had two or three nucleotides replaced at half-sites, their abilities to bind to TR4 were very different (ranging from 9.5 μM to 94.2 μM), indicating that the difference in binding affinity might be related to the sequence characteristics of the half-sites. After analyzing the sequence differences between natural genes and the canonical AGGTCA half-site, we constructed several dsDNAs with the same site mutations in both half-sites and performed MST experiments (Figure 5B). The substitution of G_2 and T_4 in the standard $A_1G_2G_3T_4C_5A_6$ half-site had the most obvious impact on binding, with a hundreds-fold loss of binding affinity, while the substitution of G_3 and C_5 of the $A_1G_2G_3T_4C_5A_6$ half-site caused a 30- to 60-fold loss of binding affinity, indicating that $G_2G_3T_4C_5$ was vital for the target genes bound by TR4. Meanwhile, A_6 in the 5′ $A_1G_2G_3T_4C_5A_6$ half-site together with the spaced adenine (A_7) and A_1 in the 3′ $A_1G_2G_3T_4C_5A_6$ half-site form an AAA sequence, providing a narrower minor groove (Figure 4C), and are crucial to the shape readout by Arg191. Thus, we deduce that the sequence PuGGTCA is an optimal target motif of TR4.

In fact, the promoter sequences of the target genes regulated by TR4 do not always have two perfect repeated AGGTCA half-sites. One or two half-sites may have some deviations from the ideal AGGTCA sequence, causing the differences in regulation between target genes. To determine the respective contributions of the 5′ half-site and the 3′ half-site to binding, we performed MST experiments. The results showed that although both substitutions affected the binding, the 5′ half-site (K_D = 2.7 μM) played a more important role than the 3′ half-site (K_D = 15.9 μM) (Figure 5C). However, excluding the CTE (only solved in DBD Sub2), the total buried interfaces between each DBD and its target half-site were equivalent (754.6 $Å^2$ for the 5′ half-site, 748.9 $Å^2$ for the 3′ half-site) (Supplementary Table S2). Moreover, a more intricate hydrogen bond network was formed between $TR4_{DBD\text{-}CTE}$ and the 3′ half-site (Figure 4A), which was puzzling to us for a long time until we calculated the free energy of the interface dissociation (ΔG^{diss}) of each half-site. The ΔG^{diss} of the 5′ half-site is 7.6 kcal/mol, while that for the 3′ half-site is only 5.3 kcal/mol (Supplementary Table S2). The higher free energy demonstrates a stronger interaction, indicating that the 5′ half-site plays a much more important role, and TR4 tends to bind the target with a complete 5′ half-site of the AGGTCA sequence.

TR4 Arg191 recognizes a narrow minor groove, which plays a crucial role in DNA binding. The minor groove is located at the interval of the half-sites, indicating that the characteristics of the spacing sequence may affect the recognition of target genes by TR4. Previous studies had demonstrated that AT-rich sequences tend to form narrow minor grooves (63). Therefore, we designed sequences with different numbers of spaced adenines. The MST results showed that the sequences spaced by 0–4 adenines had no significant difference when binding to TR4, while the DR5 (K_D = 0.2 μM) sequence with an AAAAA interval had a significantly high binding affinity with TR4 (Figure 5D), indicating that some genes/regions with this sequence feature in humans might be preferentially regulated by TR4 (Supplementary Table S5). Through the study of sequences containing different numbers of spaced bases (DR0–DR5) from the EMSA results, we also found that the DR1 sequence had prominent stability to form heterotrimers with TR4, while other types of sequences formed an observable proportion of TR4–DNA heterodimers (Figure 5E). These results revealed that the binding of TR4 and DNA was dynamic; the TR4–DNA heteromers might recruit other NRs to coregulate some target genes.

In the crystal structure, both termini of the TR4 DBD extend independently from the core DNA-binding region (Figure 3E) to form linkers to connect the NTD (residues 1–112) and the hinge region (residues 190–341). It is difficult to obtain a fixed conformation for the flexible NTD and the hinge, so we can only determine their effects on the DNA-binding ability of the DBD by measuring the equilibrium dissociation constants. The results showed that the presence of the NTD did not significantly affect the binding of DNA, while the presence of a part of the hinge region increased the binding affinity by ∼15 times (Figure 5F), suggesting that the hinge region played an important role in DNA binding, and this might be the effect of the cooperativity of the $TR4_{DBD\text{-}CTE}$ dimer.

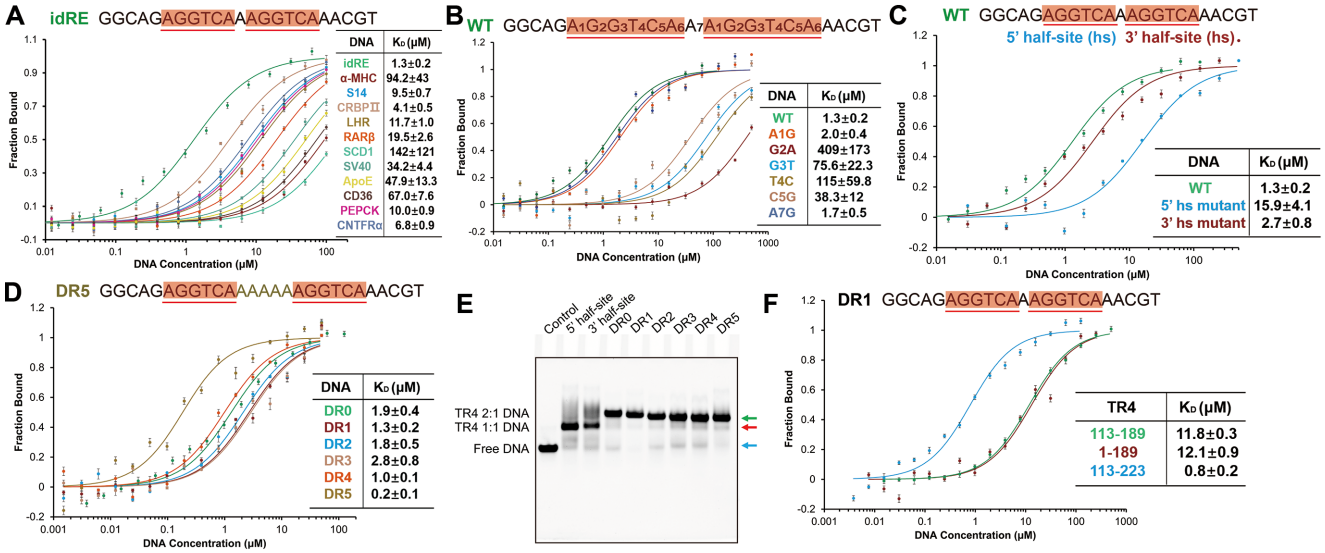

Figure 5. Recognition of target genes by TR4. The protein used is $TR4_{113–200}$. The DNAs used are all dsDNAs. (**A**) MST measurement of the binding affinity of target gene promoters. idRE: ideal DR1 response element, with two canonical AGGTCA half-sites separated by an adenine in its sequence. (**B**) MST measurement of the TR4 binding affinity of different mutants of idRE. (**C**) Comparison of the binding affinity between the sequences containing each single or double AGGTCA half-site. (**D**) MST measurement of the TR4 binding affinity of DR0–DR5 (0–5 base A) sequences indicates a priority for the sequence containing 5-spaced base A. (**E**) EMSA of TR4 with different dsDNAs. (**F**) The effects of the NTD and hinge region on DNA binding. The fragments of TR4 used are shown.

Both the DBD and LBD contribute to the formation of the TR4 homodimer

As we know, TR4 uses the homodimer to bind target genes (64). Previous studies have demonstrated that the TR4 LBD exists in the form of a homodimer both in solution and in the crystal form (21). To clarify the contribution of the DBD and LBD to TR4 homodimerization, we first obtained the crystal structure of the TR4 DBD (Figure 3A), which contained two DBD molecules per asymmetric unit with a total buried interface of 451.8 $Å^2$. At the same time, we analyzed the oligomeric state of the DBD in solution by SAXS experiments (Supplementary Table S4). The results implied that DBD had two types of oligomeric states in solution. Although monomers accounted for a larger portion, the DBD could also form homodimers itself (Figure 6A).

We also analyzed the oligomeric state of the TR4 DBD in the presence of DNA. The SAXS results showed that the TR4 DBD formed a $DBD_{(2)}$–DNA heterotrimer (Figure 6B), consistent with our $TR4_{DBD\text{-}CTE}$–DNA complex structure (Figure 3E). The dimerization of the DBD played a crucial role in the recognition of DNA (Figure 4F). The mutations of key residues involved in dimerization caused obvious loss of binding to target DNA (Figure 4D, F), indicating that the DBD contributed to TR4 homodimerization and played an important role in its functions. Moreover, the oligomeric state analysis of the $TR4_{LBD}$–JAZF1 complex in solution implied that the TR4 LBD maintained its dimerization state (Figure 6C), consistent with our solved structure. Taken together, these findings demonstrate that both DBD and LBD contribute to form TR4 homodimers.

Cancer-associated mutations of TR4 affect their binding to dsDNA *in vitro*

TR4 is involved in the progress of many diseases by regulating the transcriptional activation of many disease-associated genes, and many mutations have been linked to cancers such as adenocarcinoma, bladder carcinoma, endometrioid carcinoma and squamous cell carcinoma. We searched the COSMIC and Cancer3Dv2 cancer databases, and summarized cancer-associated mutations in Supplementary Tables S6 and S7. Among the mutations reported, many are located at the $TR4_{DBD\text{-}CTE}$–DNA interface. These residues involved in DNA recognition are all highly conserved in most species (Supplementary Figure S8). The R168L mutation occurs at the dimeric interface of the DBD, thereby it may affect DNA binding. On the other hand, the R191W mutation may weaken the shape readout at the narrow minor groove of DNA. Other mutations may directly affect the interaction between their side chains and DNA. To test these hypotheses, some cancer-associated mutations were generated (Figure 7A). As expected, R168L and R191W mutations influenced the binding to DNA, and the R173Q mutation destroyed the binding severely. However, the results of the R127C and N167K mutations were quite surprising. The R127C mutant formed a strong interaction with DNA and the binding affinity was increased by ~280 times. Meanwhile, the mutation of N167K shortened the distance from the binding site of DNA, forming a shorter and stronger hydrophilic interaction, which increased the binding affinity for DNA by >60 times. Overall, mutations of R127C and N167K improved the binding affinity between TR4 and target DNA significantly.

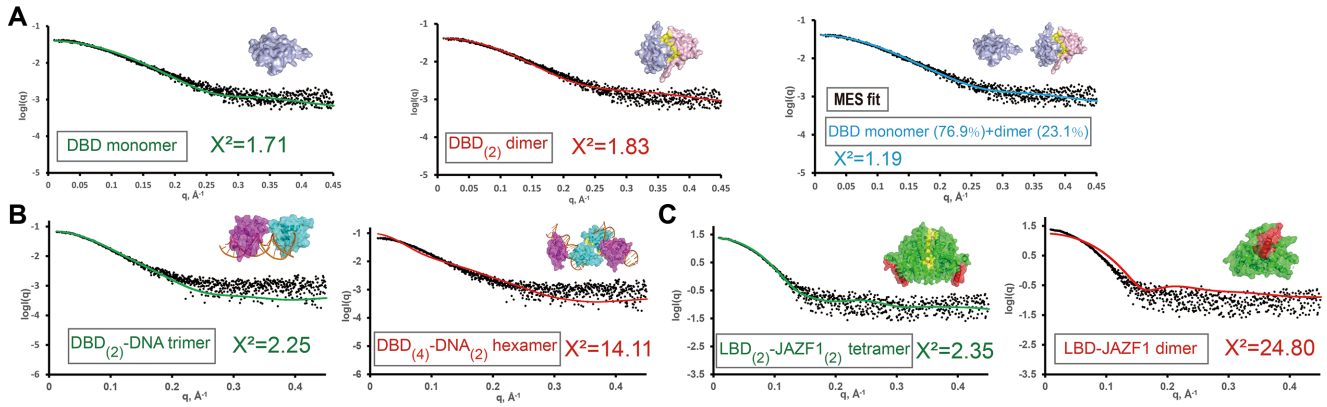

Figure 6. The oligomeric analysis of TR4. Oligomeric states of (**A**) $TR4_{DBD}$, (**B**) the $TR4_{DBD\text{-}CTE}$–DNA complex and (**C**) the $TR4_{LBD}$–JAZF1 complex. Experimental data are represented by black dots. The theoretical scattering curves of oligomers are colored differently. The dimeric interfaces are colored yellow.

Cancer-associated mutations disrupt the interaction between TR4 and JAZF1 *in vitro*

JAZF1 suppresses the transcriptional activation of downstream target genes by TR4. This suppression changes the function of TR4 in many cancers. Many cancer-associated mutations were found at the TR4–JAZF1 interface (Supplementary Tables S6 and S7), indicating that the diseases may be caused by the weaker interaction between JAZF1 and TR4. Thus, some cancer-associated mutations of TR4 and JAZF1 were generated. The pull-down results indicated that the TR4 R392C mutant almost lost the ability to bind JAZF1, which might be due to the disruption of the hydrogen bonds formed by Arg392. The mutation of JAZF1 F64I also caused a great loss of the ability to bind to TR4, as isoleucine was weaker than phenylalanine in the hydrophobic interaction network (Figure 7B).

Cancer-associated mutations affect the transcriptional activation of target genes by TR4 *in vivo*

To test the contribution of the key cancer-associated mutations to TR4-induced transcriptional activation *in vivo*, we overexpressed wild-type TR4 and JAZF1 and some cancer-associated mutants in COS-7 cells, and a dual-luciferase reporter gene system was constructed to determine the transcriptional activation of target genes by TR4 (52). The results (Figure 7C) showed that the overexpression of TR4 significantly induced the expression of target gene. The R173Q mutation disrupted the binding to target DNA, and thus destroyed most of the transcriptional activation activity of TR4 (P <0.0001). Surprisingly, the N167K mutation increased the binding affinity for DNA *in vitro*, while inhibiting the transcriptional activation of the target gene (P <0.001) *in vivo*.

When JAZF1 was also overexpressed in the TR4-overexpressing cells, the results (Figure 7D) indicated that JAZF1 significantly inhibited most of the TR4-induced transcriptional activation (P <0.0001). The F64I mutation of JAZF1 reduced the inhibition of TR4 by weakening its interaction (P <0.0001) (Figure 7D). After we mutated hydrophobic Phe64 to a hydrophilic serine, the interaction between JAZF1 and TR4 was further compromised, leading to a weaker inhibition by more than half (P <0.0001). Collectively, these findings suggest that cancer-associated mutations affect the transcriptional activation of target genes by TR4, providing critical target information for the drug design and therapy of cancers.

DISCUSSION

TR4 binds to the promoter regions of many disease-associated genes and regulates their transcriptional activation, participating in the regulation of many important physiological processes (65). The function of TR4 in most of these diseases could be suppressed by its repressor JAZF1. Due to the lack of a ligand pocket and coregulator complex structure, no specific drugs for TR4 have been developed.

In this report, we solved high-resolution $TR4_{DBD\text{-}CTE}$–DNA and $TR4_{LBD}$–JAZF1 complex structures. In the $TR4_{LBD}$–JAZF1 complex structure, a unique tail helix, α13, was observed. This α-helix played an important role in the recognition of JAZF1 and had not been reported in other NR structures. In addition, we observed substantial conformational changes of the TR4 LBD after binding JAZF1 (Figure 1C). The apo-TR4 LBD structure obtained previously was a 3 Å low-resolution structure containing 10 α-helices (21). Our $TR4_{LBD}$–JAZF1 complex structure not only increased the resolution to 1.86 Å, but also showed the missed secondary structures, redefining the LBD as a more complete structure consisting of 12 α-helices. Moreover, the TR4-interacting domain (TID) of JAZF1 was minimally redefined to residues 51–75.

For NRs, the participation of coactivators is required for regulating the transcriptional activation of target genes. The coactivators bind at the hydrophobic interface of NRs with their LXXLL motifs. Similarly, the classical NR corepressor NCoR/SMRT interacts with specific residues on the same surface required for coactivator binding via their consensus sequence of LXXI/HIXXXI/L motifs (66). However, it is observed that the binding interface of JAZF1 on TR4 was very different from those of the classic coactivators and corepressors. JAZF1 bound to a new surface formed by the AF-2, α3 and the unique helix α13 (Figure 2A). Sequence alignments showed that the residues of JAZF1

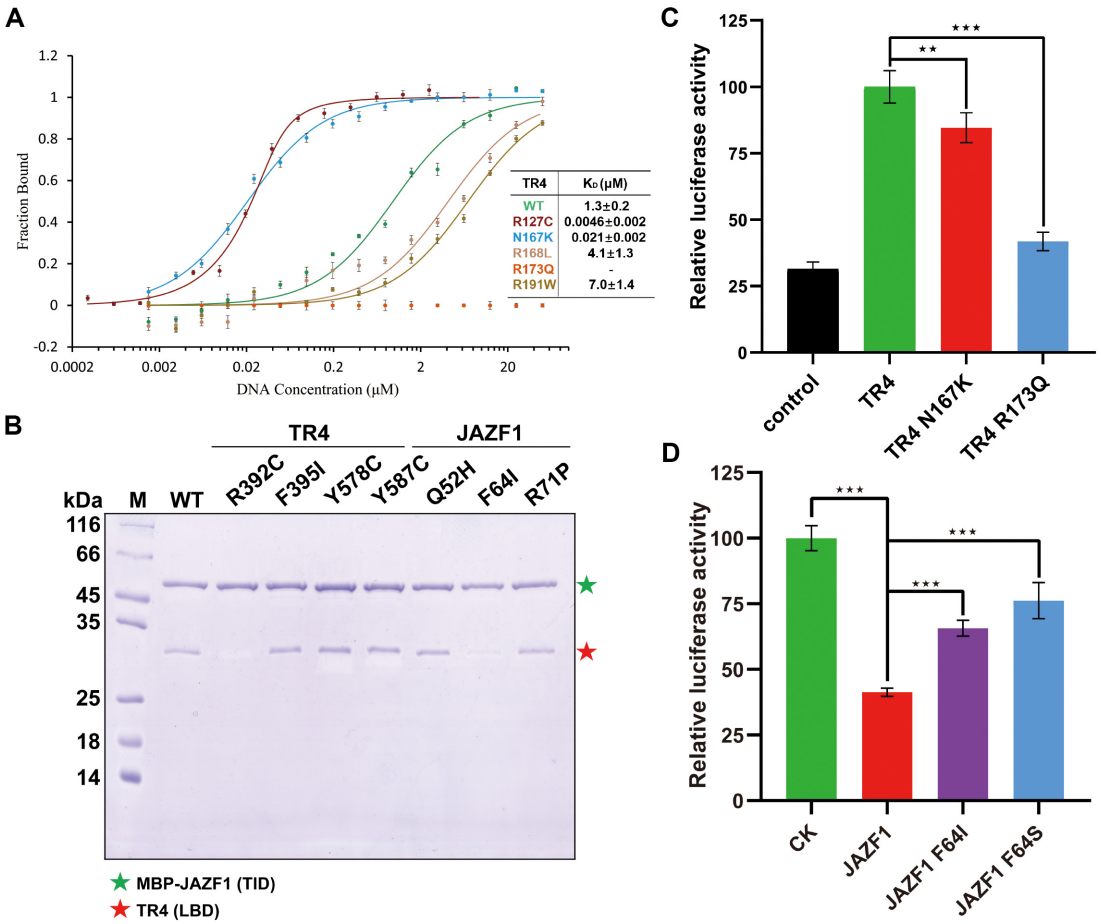

Figure 7. Mutations of TR4 and JAZF1 in cancers. (**A**) MST measurement of the binding affinity between different cancer-associated mutations of the $TR4_{DBD\text{-}CTE}$ with idRE. (**B**) Some cancer-associated mutations in the TR4 LBD and JAZF1 destroyed their interaction by MBP pull-down assay. (**C**) Effects of cancer-associated mutations of the $TR4_{DBD}$ on TR4-induced transcriptional activation. (**D**) Effects of JAZF1 and JAZF1 cancer-associated mutants on TR4-induced transcriptional activation. Luciferase activities were measured and presented as fold change compared with the control (CK). Data are the means ± standard deviation of three independent experiments. One-way ANOVA was used to test differences for statistical significance. $^{**}P < 0.001$ and $^{***}P < 0.0001$.

involved at the interface were well conserved (Figure 2B), while the residues of TR4 were variable among NRs (Figure 2C), indicating that the interaction of JAZF1 is rather specific for TR4 in the NR family. This hypothesis is further supported by an *in vivo* study (33). This particular suppression is critical for the development of TR4-specific peptide drugs. Most importantly, our structure and pull-down results demonstrated that JAZF1 reinforced the AF-2 helix on its binding sites and thus blocked the binding of the coactivators (Figure 2D, E), revealing the mechanism of the suppression from JAZF1 to TR4.

Targeting downstream genes is the most characteristic feature of NRs. Research on NRs and DNA has been conducted over a long period (57–61), while for TR4 it is still in its infancy. To clarify the sequence characteristics of TR4 target genes and determine the mechanism of how TR4 regulates target genes, we solved the complex structures of TR4 and dsDNA. In the solved structures, except in the case of direct recognition, water molecules mediated the binding of DNA bases and TR4 side chains (Figure 4A, B). Meanwhile, similarly to RXR (39), Gln188 and Arg191 recognized the narrower minor groove of DNA (Figure 4C), revealing a conserved shape readout mechanism. Furthermore, again like RXR (39), the existence of the CTE induced the dimerization of the DBD on DNA (Figure 4F). In contrast, serine barely existed at the dimeric interface in most previous DBD synergy studies (59,67–68), while it made a great contribution to TR4 dimerization (Figure 4F). Further, the CTE formed hydrophobic interactions with ZnF-II of another DBD (Supplementary Figure S6A), which had not been observed in other NR members. Moreover, the key residue Pro193 is not conserved among the NR members except TR2 (Supplementary Figure S6B), indicating a specific mode of TR4 and TR2.

By MST experiments, we defined the half-site recognized by TR4 as PuGGTCA, which was slightly different from the targets of other NRs (Supplementary Table S8). Furthermore, due to the difference in free energy on dissociating from TR4, the 5′ half-site played a much more important role than the 3′ half-site (Figure 5C). This finding was demonstrated for the first time in the NR family, explaining the target selectivity preference of TR4. The sequence characteristics of the spacer between the repeated half-sites

are also important for binding. Our results showed that TR4 had a binding preference for the DR5 sequence with the AAAAA spacer, the binding affinity was 5- to 10-fold higher than for DR0–DR4 (Figure 5D). As we searched the human genome, we indeed found the same sequence in intergenic and non-coding regions of genes associated with metabolism and cancers (Supplementary Table S5), indicating that TR4 might have the potential to regulate some of them and play important roles in related physiological processes in humans. By EMSA experiments, we also found that TR4 formed an observable proportion of TR4–DNA heterodimers, indicating that TR4 might form heteromers with other NRs to recognize different targets. For example, RXR can form not only homodimers but also heterodimers with a variety of other NRs to recognize multiple target genes (51–59,69). There has also been some research reporting that TR4 could form heteromers with AR (70), TR (71), TR2 (72) or ER (73) to recognize different target genes. Since the LBD dimerizes easily, the TR4 DBD has not been the focus of research. Our structure results (Figure 3E) and SAXS experiments (Figure 6) revealed that the TR4 DBD exists as homodimers to bind DNA, indicating that both the DBD and LBD make a contribution to TR4 homodimerization.

Finally, we found that many mutations in cancer tissues were located at the interfaces of the TR4 DBD, LBD and JAZF1 TID. Diseases may arise from the disruption of the interactions caused by mutations. Mutations located at PPARγ DBD and LBD played an important role in partial lipodystrophy, dyslipidemia, insulin resistance and colon cancer by disrupting the interaction with DNAs, ligands and coactivators (74–77). As expected, our results also showed that cancer-associated mutations significantly affected the interactions between TR4 and DNA (Figure 7A) as well as that between TR4 and JAZF1 (Figure 7B). Specifically, the N167K mutation found in endometrioid carcinoma and the R127C mutation reported in adenocarcinoma increased the binding affinity by ∼60 and ∼280 times, respectively. These were new findings and might offer guidance for other NR members. Moreover, the R173Q mutant, also found in adenocarcinoma, greatly disrupted DNA binding and severely destroyed the activation of target genes (Figure 7A, C). The JAZF1 F64I mutant in squamous cell carcinoma also severely disrupted the interaction with TR4 and significantly weakened the suppression of TR4-induced transcriptional activation (Figure 7B, D). Taken together, these cancer-associated mutations have remarkable effects on the interactions involving TR4 both *in vitro* and *in vivo*.

Briefly, we successfully solved three complex structures of TR4 for the first time, showing a novel helix α13 never reported in the NR family. We proposed a working model (Supplementary Figure S9) and explained how JAZF1 suppressed TR4-mediated transcriptional activation. The details of the interaction are helpful for the development of TR4-specific peptide drugs. Furthermore, multiple factors were found to contribute to DNA recognition, the characteristics of the target DNAs were summarized and the mutual selectivity between TR4 and target DNA was also analyzed. Finally, the cancer-associated mutations were found and verified both *in vitro* and *in vivo*. Overall, we provided extensive structural information and key targets to develop the diagnostic tools and drug design strategies for a variety of diseases, including diabetes and cancers.

DATA AVAILABILITY

Coordinates and structure factors have been deposited in the Protein Data Bank under accession codes 7XV6, 7XV8, 7XV9 and 7XVA.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We are grateful to Dr X. Edward Zhou (University of Michigan Medical School) for kindly providing the pCR3.1-TR4 LBD plasmid. We would also like to thank the staff of beamlines BL02U1, BL10U2, BL18U1 and BL19U2 at the Shanghai Synchrotron Radiation Facility for their excellent technical assistance.

FUNDING

This work was supported by National Key Research and Development Program of China [2018YFE0113100]; the National Natural Science Foundation of China [31872713 and 32071210]; and the Extramural Scientists of State Key Laboratory of Agrobiotechnology [2022SKLAB6-2].
Conflict of interest statement. None declared.

REFERENCES

1. Chang,C., Da Silva,S.L., Ideta,R., Lee,Y., Yeh,S. and Burbach,J.P. (1994) Human and rat TR4 orphan receptors specify a subclass of the steroid receptor superfamily. *Proc. Natl Acad. Sci. USA*, **91**, 6040–6044.
2. Bookout,A.L., Jeong,Y., Downes,M., Yu,R.T., Evans,R.M. and Mangelsdorf,D.J. (2006) Anatomical profiling of nuclear receptor expression reveals a hierarchical transcriptional network. *Cell*, **126**, 789–799.
3. 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.
4. de Vera,I.M.S., Munoz-Tello,P., Zheng,J., Dharmarajan,V., Marciano,D.P., Matta-Camacho,E., Giri,P.K., Shang,J., Hughes,T.S., Rance,M. *et al.* (2019) Defining a canonical ligand-binding pocket in the orphan nuclear receptor Nurr1. *Structure*, **27**, 66–77.
5. Arao,Y. and Korach,K.S. (2021) The physiological role of estrogen receptor functional domains. *Essays Biochem.*, **65**, 867–875.
6. de Vries,R., Meijer,F.A., Doveston,R.G., Leijten-van de Gevel,I.A. and Brunsveld,L. (2021) Cooperativity between the orthosteric and allosteric ligand binding sites of RORγt. *Proc. Natl Acad. Sci. USA*, **118**, e2021287118
7. Shang,J. and Kojetin,D.J. (2021) Structural mechanism underlying ligand binding and activation of PPARγ. *Structure*, **29**, 940–950.
8. Reitzel,A.M., Macrander,J., Mane-Padros,D., Fang,B., Sladek,F.M. and Tarrant,A.M. (2018) Conservation of DNA and ligand binding properties of retinoid X receptor from the placozoan *Trichoplax adhaerens* to human. *J. Steroid Biochem. Mol. Biol.*, **184**, 3–10.
9. Shyr,C.R., Collins,L.L., Mu,X.M., Platt,K.A. and Chang,C. (2002) Spermatogenesis and testis development are normal in mice lacking testicular orphan nuclear receptor 2. *Mol. Cell. Biol.*, **22**, 4661–4666.
10. Kim,E., Liu,N.C., Yu,I.C., Lin,H.Y., Lee,Y.F., Sparks,J.D., Chen,L.M. and Chang,C. (2011) Metformin inhibits nuclear receptor TR4-mediated hepatic stearoyl-CoA desaturase 1 gene expression with altered insulin sensitivity. *Diabetes*, **60**, 1493–1503.

11. Xie,S., Lee,Y.F., Kim,E., Chen,L.M., Ni,J., Fang,L.Y., Liu,S., Lin,S.J., Abe,J., Berk,B. *et al.* (2009) TR4 nuclear receptor functions as a fatty acid sensor to modulate CD36 expression and foam cell formation. *Proc. Natl Acad. Sci. USA*, **106**, 13353–13358.
12. Lin,S.J., Ho,H.C., Lee,Y.F., Liu,N.C., Liu,S., Li,G., Shyr,C.R. and Chang,C. (2012) Reduced osteoblast activity in the mice lacking TR4 nuclear receptor leads to osteoporosis. *Reprod. Biol. Endocrinol.*, **10**, 43.
13. Lee,M.P., Tanabe,O., Shi,L., Jearawiriyapaisarn,N., Lucas,D. and Engel,J.D. (2017) The orphan nuclear receptor TR4 regulates erythroid cell proliferation and maturation. *Blood*, **130**, 2537–2547.
14. Lee,Y.F., Liu,S., Liu,N.C., Wang,R.S., Chen,L.M., Lin,W.J., Ting,H.J., Ho,H.C., Li,G., Puzas,E.J. *et al.* (2011) Premature aging with impaired oxidative stress defense in mice lacking TR4. *Am. J. Physiol. Endocrinol. Metab.*, **301**, E91–E98.
15. Zhu,J., Yang,D.R., Sun,Y., Qiu,X., Chang,H.C., Li,G., Shan,Y. and Chang,C. (2015) TR4 nuclear receptor alters the prostate cancer CD133+ stem/progenitor cell invasion via modulating the EZH2-related metastasis gene expression. *Mol. Cancer Ther.*, **14**, 1445–1453.
16. Wang,M., Sun,Y., Xu,J., Lu,J., Wang,K., Yang,D.R., Yang,G., Li,G. and Chang,C. (2018) Preclinical studies using miR-32-5p to suppress clear cell renal cell carcinoma metastasis via altering the miR-32-5p/TR4/HGF/Met signaling. *Int. J. Cancer.*, **143**, 100–112.
17. Chen,Y., Lu,J., Xia,L., Xue,D., Yu,X., Shen,D., Xu,L. and Li,G. (2018) Testicular orphan receptor 4 promotes tumor progression and implies poor survival through AKT3 regulation in seminoma. *Cancer Sci.*, **109**, 384–394.
18. Jin,R., Lin,H., Li,G., Xu,J., Shi,L., Chang,C. and Cai,X. (2018) TR4 nuclear receptor suppresses HCC cell invasion via downregulating the EphA2 expression. *Cell Death Dis.*, **9**, 283.
19. Zhang,L., Zhang,J., Ma,Y., Chen,J., Dong,B., Zhao,W., Wang,X., Zheng,Q., Fang,F. and Yang,Y. (2015) Testicular orphan receptor 4 (TR4) is a marker for metastasis and poor prognosis in non-small cell lung cancer that drives the EMT phenotype. *Lung Cancer*, **89**, 320–328.
20. Zhang,Y., Li,P., Hu,J., Zhao,L.N., Li,J.P., Ma,R., Li,W.W., Shi,M. and Wei,L.C. (2019) Role and mechanism of miR-4778-3p and its targets NR2C2 and Med19 in cervical cancer radioresistance. *Biochem. Biophys. Res. Commun.*, **508**, 210–216.
21. Zhou,X.E., Suino-Powell,K.M., Xu,Y., Chan,C.W., Tanabe,O., Kruse,S.W., Reynolds,R., Engel,J.D. and Xu,H.E. (2011) The orphan nuclear receptor TR4 is a vitamin A-activated nuclear receptor. *J. Biol. Chem.*, **286**, 2877–2885.
22. Tsai,N.P., Huq,M., Gupta,P., Yamamoto,K., Kagechika,H. and Wei,L.N. (2009) Activation of testicular orphan receptor 4 by fatty acids. *Biochim. Biophys. Acta*, **1789**, 734–740.
23. Xia,L., Shen,D., Wang,H., Ren,L., Chen,Y. and Li,G. (2020) Identification of small-molecule regulators of testicular receptor 4 via a drug repurposing screening. *ACS Omega*, **5**, 30625–30632.
24. Bendell,J.C., Javle,M., Bekaii-Saab,T.S., Finn,R.S., Wainberg,Z.A., Laheru,D.A., Weekes,C.D., Tan,B.R., Khan,G.N., Zalupski,M.M. *et al.* (2017) A phase 1 dose-escalation and expansion study of binimetinib (MEK162), a potent and selective oral MEK1/2 inhibitor. *Br. J. Cancer*, **116**, 575–583.
25. Du,L., Bergsneider,M., Mirsadraei,L., Young,S.H., Jonker,J.W., Downes,M., Yong,W.H., Evans,R.M. and Heaney,A.P. (2013) Evidence for orphan nuclear receptor TR4 in the etiology of Cushing disease. *Proc. Natl Acad. Sci. USA*, **110**, 8555–8560.
26. Zhang,D., Bergsneider,M., Wang,M.B. and Heaney,A.P. (2016) Targeting the ERK pathway for the treatment of Cushing's disease. *Oncotarget*, **7**, 69149–69158.
27. Cheong,M.C., Wang,Z., Jaleta,T.G., Li,X., Lok,J.B., Kliewer,S.A. and Mangelsdorf,D.J. (2021) Identification of a nuclear receptor/coactivator developmental signaling pathway in the nematode parasite *Strongyloides stercoralis*. *Proc. Natl Acad. Sci. USA*, **118**, e2021864118.
28. Huq,M.D., Gupta,P., Tsai,N.P. and Wei,L.N. (2006) Modulation of testicular receptor 4 activity by mitogen-activated protein kinase-mediated phosphorylation. *Mol. Cell. Proteomics*, **5**, 2072–2082.
29. Yang,Y., Wang,X., Dong,T., Kim,E., Lin,W.J. and Chang,C. (2003) Identification of a novel testicular orphan receptor-4 (TR4)-associated protein as repressor for the selective suppression of TR4-mediated transactivation. *J. Biol. Chem.*, **278**, 7709–7717.
30. Clarisse,D., Thommis,J., Van Wesemael,K., Houtman,R., Ratman,D., Tavernier,J., Offner,F., Beck,I. and De Bosscher,K. (2017) Coregulator profiling of the glucocorticoid receptor in lymphoid malignancies. *Oncotarget*, **8**, 109675–109691.
31. Liu,X., Wang,Y. and Ortlund,E.A. (2019) First high-resolution crystal structures of the glucocorticoid receptor ligand-binding domain-peroxisome proliferator-activated γ coactivator 1-α complex with endogenous and synthetic glucocorticoids. *Mol. Pharmacol.*, **96**, 408–417.
32. Carcache,D.A., Vulpetti,A., Kallen,J., Mattes,H., Orain,D., Stringer,R., Vangrevelinghe,E., Wolf,R.M., Kaupmann,K., Ottl,J. *et al.* (2018) Optimizing a weakly binding fragment into a potent RORγt inverse agonist with efficacy in an in vivo inflammation model. *J. Med. Chem.*, **61**, 6724–6735.
33. Nakajima,T., Fujino,S., Nakanishi,G., Kim,Y.S. and Jetten,A.M. (2004) TIP27: a novel repressor of the nuclear orphan receptor TAK1/TR4. *Nucleic Acids Res.*, **32**, 4194–4204.
34. Jang,W.Y., Bae,K.B., Kim,S.H., Yu,D.H., Kim,H.J., Ji,Y.R., Park,S.J., Park,S.J., Kang,M.C., Jeong,J.I. *et al.* (2014) Overexpression of Jazf1 reduces body weight gain and regulates lipid metabolism in high fat diet. *Biochem. Biophys. Res. Commun.*, **444**, 296–301.
35. Shi,Y.Q., Fu,G.Q., Zhao,J., Cheng,S.Z., Li,Y., Yi,L.N., Li,Z., Zhang,L., Zhang,Z.B., Dai,J. *et al.* (2019) Di(2-ethylhexyl)phthalate induces reproductive toxicity via JAZF1/TR4 pathway and oxidative stress in pubertal male rats. *Toxicol. Ind. Health*, **35**, 228–238.
36. Bae,K.B., Kim,M.O., Yu,D.H., Shin,M.J., Kim,H.J., Yuh,H.S., Ji,Y.R., Kim,J.Y., Kim,J.M., Hyun,B.H. *et al.* (2011) Overexpression of Jazf1 induces cardiac malformation through the upregulation of pro-apoptotic genes in mice. *Transgenic Res.*, **20**, 1019–1031.
37. Li,X., Yang,M., Wang,H., Jia,Y., Yan,P., Boden,G., Yang,G. and Li,L. (2014) Overexpression of JAZF1 protected ApoE-deficient mice from atherosclerosis by inhibiting hepatic cholesterol synthesis via CREB-dependent mechanisms. *Int. J. Cardiol.*, **177**, 100–110.
38. Otwinowski,Z. and Minor,W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.*, **276**, 307–326.
39. 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.
40. Long,F., Vagin,A.A., Young,P. and Murshudov,G.N. (2008) BALBES: a molecular-replacement pipeline.*Acta Crystallogr. D Biol. Crystallogr.*, **64**, 125–132.
41. Emsley,P., Lohkamp,B., Scott,W.G. and Cowtan,K. (2010) Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 486–501.
42. Murshudov,G.N., Skubák,P., Lebedev,A.A., Pannu,N.S., Steiner,R.A., Nicholls,R.A., Winn,M.D., Long,F. and Vagin,A.A. (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D Biol. Crystallogr.*, **67**, 355–367.
43. McCoy,A.J., Grosse-Kunstleve,R.W., Adams,P.D., Winn,M.D., Storoni,L.C. and Read,R.J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr.*, **40**, 658–674.
44. Vagin,A. and Teplyakov,A. (2010) Molecular replacement with MOLREP. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 22–25.
45. Alexander,N., Woetzel,N. and Meiler,J. (2011) bcl::Cluster: a method for clustering biological molecules coupled with visualization in the Pymol Molecular Graphics System. *IEEE Int. Conf. Comput. Adv. Biol. Med. Sci.*, **2011**, 13–18.
46. Hu,Y., Chen,Z., Fu,Y., He,Q., Jiang,L., Zheng,J., Gao,Y., Mei,P., Chen,Z. and Ren,X. (2015) The amino-terminal structure of human fragile X mental retardation protein obtained using precipitant-immobilized imprinted polymers. *Nat. Commun.*, **6**, 6634.
47. Zhang,X., Zhang,Q., Xin,Q., Yu,L., Wang,Z., Wu,W., Jiang,L., Wang,G., Tian,W., Deng,Z. *et al.* (2012) Complex structures of the abscisic acid receptor PYL3/RCAR13 reveal a unique regulatory mechanism. *Structure*, **20**, 780–790.
48. Nielsen,S.S., Toft,K.N., Snakenborg,D., Jeppesen,M.G., Jacobsen,J.K., Vestergaard,B., Kutter,J.P. and Arleth,L. (2009) BioXTAS RAW, a software program for high-throughput automated small-angle X-ray scattering data reduction and preliminary analysis. *J. Appl. Crystallogr.*, **42**, 959–964.

49. Schneidman-Duhovny,D., Hammel,M., Tainer,J.A. and Sali,A. (2013) Accurate SAXS profile computation and its assessment by contrast variation experiments. *Biophys. J.*, **105**, 962–974.
50. Lavery,R., Moakher,M., Maddocks,J.H., Petkeviciute,D. and Zakrzewska,K. (2009) Conformational analysis of nucleic acids revisited: curves+. *Nucleic Acids Res.*, **37**, 5917–5929.
51. Zhou,T., Yang,L., Lu,Y., Dror,I., Dantas Machado,A.C., Ghane,T., Di Felice,R. and Rohs,R. (2013) DNAshape: a method for the high-throughput prediction of DNA structural features on a genomic scale. *Nucleic Acids Res.*, **41**, W56–W62.
52. Xia,L., Shen,D., Zhang,Y., Lu,J., Wang,M., Wang,H., Chen,Y., Xue,D., Xie,D. and Li,G. (2021) Targeting the TR4 nuclear receptor with antagonist bexarotene can suppress the proopiomelanocortin signalling in AtT-20 cells. *J. Cell. Mol. Med.*, **25**, 2404–2417.
53. Hu,R., Wu,W., Niles,E.G. and LoVerde,P.T. (2006) SmTR2/4, a *Schistosoma mansoni* homologue of TR2/TR4 orphan nuclear receptor. *Int. J. Parasitol.*, **36**, 1113–1122.
54. Xu,R.X., Lambert,M.H., Wisely,B.B., Warren,E.N., Weinert,E.E., Waitt,G.M., Williams,J.D., Collins,J.L., Moore,L.B., Willson,T.M. *et al.* (2004) A structural basis for constitutive activity in the human CAR/RXRalpha heterodimer. *Mol. Cell.*, **16**, 919–928.
55. Liu,N.C., Lin,W.J., Kim,E., Collins,L.L., Lin,H.Y., Yu,I.C., Sparks,J.D., Chen,L.M., Lee,Y.F. and Chang,C. (2007) Loss of TR4 orphan nuclear receptor reduces phosphoenolpyruvate carboxykinase-mediated gluconeogenesis. *Diabetes*, **56**, 2901–2909.
56. Kim,E., Xie,S., Yeh,S.D., Lee,Y.F., Collins,L.L., Hu,Y.C., Shyr,C.R., Mu,X.M., Liu,N.C., Chen,Y.T. *et al.* (2003) Disruption of TR4 orphan nuclear receptor reduces the expression of liver apolipoprotein E/C-I/C-II gene cluster. *J. Biol. Chem.*, **278**, 46919–46926.
57. 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.
58. Devarakonda,S., Harp,J.M., Kim,Y., Ozyhar,A. and Rastinejad,F. (2003) Structure of the heterodimeric ecdysone receptor DNA-binding complex. *EMBO J.*, **22**, 5827–5840.
59. Lou,X., Toresson,G., Benod,C., Suh,J.H., Philips,K.J., Webb,P. and Gustafsson,J.A. (2014) Structure of the retinoid X receptor α–liver X receptor β (RXRα–LXRβ) heterodimer on DNA. *Nat. Struct. Mol. Biol.*, **21**, 277–281.
60. Shaffer,P.L. and Gewirth,D.T. (2002) Structural basis of VDR–DNA interactions on direct repeat response elements. *EMBO J.*, **21**, 2242–2252.
61. Osz,J., McEwen,A.G., Poussin-Courmontagne,P., Moutier,E., Birck,C., Davidson,I., Moras,D. and Rochel,N. (2015) Structural basis of natural promoter recognition by the retinoid X nuclear receptor. *Sci. Rep.*, **5**, 8216.
62. Tian,Z., Li,X., Li,M., Wu,W., Zhang,M., Tang,C., Li,Z., Liu,Y., Chen,Z., Yang,M. *et al.* (2020) Crystal structures of REF6 and its complex with DNA reveal diverse recognition mechanisms. *Cell Discov.*, **6**, 17.
63. Rohs,R., West,S.M., Sosinsky,A., Liu,P., Mann,R.S. and Honig,B. (2009) The role of DNA shape in protein–DNA recognition. *Nature*, **461**, 1248–1253.
64. Choi,H., Park,S.S., Kim,S.J. and Kim,E. (2020) Beta-catenin inhibits TR4-mediated lipid accumulation in 3T3-L1 adipocytes via induction of Slug. *Cell Biosci.*, **10**, 119.
65. Lin,S.J., Zhang,Y., Liu,N.C., Yang,D.R., Li,G. and Chang,C. (2014) Minireview: pathophysiological roles of the TR4 nuclear receptor: lessons learned from mice lacking TR4. *Mol. Endocrinol.*, **28**, 805–821.
66. Lee,J.W., Lee,Y.C., Na,S.Y., Jung,D.J. and Lee,S.K. (2001) Transcriptional coregulators of the nuclear receptor superfamily: coactivators and corepressors. *Cell. Mol. Life. Sci.*, **58**, 289–297.
67. Osz,J., McEwen,A.G., Bourguet,M., Przybilla,F., Peluso-Iltis,C., Poussin-Courmontagne,P., Mély,Y., Cianférani,S., Jeffries,C.M., Svergun,D.I. *et al.* (2020) Structural basis for DNA recognition and allosteric control of the retinoic acid receptors RAR–RXR. *Nucleic Acids Res.*, **48**, 9969–9985.
68. Maletta,M., Orlov,I., Roblin,P., Beck,Y., Moras,D., Billas,I.M. and Klaholz,B.P. (2014) The palindromic DNA-bound USP/EcR nuclear receptor adopts an asymmetric organization with allosteric domain positioning. *Nat. Commun.*, **5**, 4139.
69. Peluso-Iltis,C., Osz,J. and Rochel,N. (2020) DNA recognition by retinoic acid nuclear receptors. *Methods Enzymol.*, **637**, 235–260.
70. Lee,Y.F., Shyr,C.R., Thin,T.H., Lin,W.J. and Chang,C. (1999) Convergence of two repressors through heterodimer formation of androgen receptor and testicular orphan receptor-4: a unique signaling pathway in the steroid receptor superfamily. *Proc. Natl Acad. Sci. USA*, **96**, 14724–14729.
71. Huang,Y.H., Liao,C.H., Chen,R.N., Liao,C.J. and Lin,K.H. (2010) Human testicular orphan receptor 4 enhances thyroid hormone receptor signaling. *J. Cell. Physiol.*, **222**, 347–356.
72. Lee,C.H., Chinpaisal,C. and Wei,L.N. (1998) A novel nuclear receptor heterodimerization pathway mediated by orphan receptors TR2 and TR4. *J. Biol. Chem.*, **273**, 25209–25215.
73. Shyr,C.R., Hu,Y.C., Kim,E. and Chang,C. (2002) Modulation of estrogen receptor-mediated transactivation by orphan receptor TR4 in MCF-7 cells. *J. Biol. Chem.*, **277**, 14622–14628.
74. Khorasanizadeh,S. and Rastinejad,F. (2001) Nuclear-receptor interactions on DNA-response elements. *Trends Biochem. Sci.*, **26**, 384–390.
75. Wurtz,J.M., Bourguet,W., Renaud,J.P., Vivat,V., Chambon,P., Moras,D. and Gronemeyer,H. (1996) A canonical structure for the ligand-binding domain of nuclear receptors. *Nat. Struct. Biol.*, **3**, 206.
76. McKenna,N.J. and O'Malley,B.W. (2002) Combinatorial control of gene expression by nuclear receptors and coregulators. *Cell*, **108**, 465–474.
77. Heery,D.M., Kalkhoven,E., Hoare,S. and Parker,M.G. (1997) A signature motif in transcriptional co-activators mediates binding to nuclear receptors. *Nature*, **387**, 733–736.