THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 280, No. 36, Issue of September 9, pp. 31818–31829, 2005
Printed in U.S.A.

Evolutionary Trace-based Peptides Identify a Novel Asymmetric Interaction That Mediates Oligomerization in Nuclear Receptors*

Received for publication, February 22, 2005, and in revised form, June 22, 2005
Published, JBC Papers in Press, July 1, 2005, DOI 10.1074/jbc.M501924200

Peili Gu‡§, Daniel H. Morgan§¶‖, Minawar Sattar‡§, Xueping Xu‡, Ryan Wagner‡, Michele Raviscioni¶§§, Olivier Lichtarge¶‖**, and Austin J. Cooney‡ ‡‡**

*From the ‡Department of Molecular and Cellular Biology, ¶Department of Molecular and Human Genetics, the ‖Graduate Program in Structural and Computational Biology and Molecular Biophysics, and the **W. M. Keck Center for Computational and Structural Biology, Baylor College of Medicine, Houston, Texas 77030*

Germ cell nuclear factor (GCNF) is an orphan nuclear receptor that plays important roles in development and reproduction, by repressing the expression of essential genes such as *Oct4*, *GDF9*, and *BMP15*, through binding to DR0 elements. Surprisingly, whereas recombinant GCNF binds to DR0 sequences as a homodimer, endogenous GCNF does not exist as a homodimer but rather as part of a large complex termed the transiently retinoid-induced factor (TRIF). Here, we use evolutionary trace (ET) analysis to design mutations and peptides that probe the molecular basis for the formation of this unusual complex. We find that GCNF homodimerization and TRIF complex formation are DNA-dependent, and ET suggests that dimerization involves key functional sites on both helix 3 and helix 11, which are located on opposing surfaces of the ligand binding domain. Targeted mutations in either helix of GCNF disrupt the formation of both the homodimer and the endogenous TRIF complex. Moreover, peptide mimetics of both of these ET-determined sites inhibit dimerization and TRIF complex formation. This suggests that a novel helix 3-helix 11 heterotypic interaction mediates GCNF interaction and would facilitate oligomerization. Indeed, it was determined that the endogenous TRIF complex is composed of a GCNF oligomer. These findings shed light on an evolutionarily selected mechanism that reveals the unusual DNA-binding, dimerization, and oligomerization properties of GCNF.

Nuclear receptors are one of the most diverse families of ligand-activated transcriptional regulators in animals (metazoans), with a central role in regulating human metabolism, hormonal function, homeostasis, and development (1–4). These proteins control transcription through orchestrated interactions with chaperones, ligands (5–8), co-repressors, co-activators, and DNA (7, 9), which they bind to as either monomers, homodimers, or heterodimers (10). The exact state of dimerization of nuclear receptors is determined by specific interactions between the DNA binding domains and between the ligand binding domains (LBD)[1] and frequently varies among family members. The LBD, usually formed by 12 packed α-helices that fold into a three-layer antiparallel α-helical sandwich (4, 7), dimerizes symmetrically in different nuclear receptors through different interaction faces. For example, ERRγ (11), Nurr1 (12), HNF4 (13, 14), and SHP (15) homodimerize via a helix 10/helix 10 interaction, whereas RXR heterodimerizes with RA receptor via a helix 11/helix 11 interaction (16–18). In the case of the peroxisome proliferator-activated receptor γ, dimerization occurs via a helix 10/helix 10 interaction that also involves helices 7 and 9 (18). The glucocorticoid receptor forms homodimers with the β-turn and helices 3 and 4 (19). Yet another dimerization mode has been postulated to exist in the orphan nuclear receptor GCNF (20).

GCNF is expressed in the germ cells of adults and widely during early embryonic development. It is also expressed in differentiating embryonic stem cells and embryonic carcinoma cells (21–25). It represses target gene expression (26) upon binding to the DNA response element DR0, a direct repeat of the sequence AGGTCA (27, 28). GCNF regulates *Oct4* expression during gastrulation (29) and modulates oocyte-specific genes such as BMP15 and GDF9 (30). Surprisingly, when it is translated *in vitro*, GCNF forms a homodimer that is postulated to be mediated by a helix 3/helix 3 interaction (20). However, in differentiated P19 cells and ES cells, endogenous GCNF does not exist as a homodimer but rather forms a large complex called the TRIF complex (23, 29, 31). The role of helix 3 (H3) in dimerization, and the unique presence of an unusual endogenous DNA binding complex, suggested that GCNF may further expand the known repertoire of interactions in the nuclear receptor superfamily, and motivated this study to identify the key amino acids and functional sites that mediate GCNF interactions.

To identify these functional determinants, we turned to the Evolutionary Trace (ET) method. ET is an algorithm that probes the evolutionary record that resides in the divergent amino acid sequences of a large gene family and ranks the importance of each amino acid position by correlating their variations with evolutionary divergences (32, 33). Residues ranked in the top 30th percentile are known to cluster spatially

* This work was supported by National Institutes of Health Nuclear Receptor Signaling Atlas orphan receptor program Grant U19DK6234-01 and in part by National Science Foundation Grant DBI-0318415, National Institutes of Health Grant R01 GM066099, and March of Dimes Grant 1-FY03-93 (to O. L.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ These authors contributed equally to this work.

‡‡ To whom correspondence should be addressed: Dept. of Molecular and Cellular Biology, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030. Tel.: 713-798-6250; Fax: 713-790-1275; E-mail: acooney@bcm.tmc.edu.

§§ Supported by a training fellowship from the W. M. Keck Center for Computational and Structural Biology.

[1] The abbreviations used are: LBD, ligand binding domain; RXR, retinoid X receptor; ET, evolutionary trace; CMV, cytomegalovirus; HA, hemagglutinin; EMSA, electrophoretic mobility shift assay(s); RA, retinoic acid; GCNF, germ cell nuclear factor; TRIF, transiently retinoid-induced factor; H3 and H11, helix 3 and 11, respectively.

in native structures (34) such that they significantly overlap functional sites (35) and predict functional specificity determinants (36) that can then be rationally mutated to reengineer function (37, 38). In this study, ET was applied to the LBD of the NR superfamily and revealed two distinct functional surfaces focused on helices 3 and 11. This led to the rational design of mutations in GCNF that demonstrated that both helix 3 and helix 11 are involved in GCNF homodimerization and TRIF complex formation. In a novel strategy, we used the ET predictions to engineer peptides that aimed to embody the activity of the sites on helix 3 and helix 11. Both were biologically active in inhibiting homodimerization and TRIF complex formation. These data point to a novel interaction mode in nuclear receptors through heterotypic helix 3-helix 11 binding. Heterotypic interactions are permissive to oligomerization, and we show that the TRIF complex is comprised of a GCNF oligomer. In the context of the NR superfamily, this reveals a new mechanism of interaction with implications for transcriptional regulation and cross-talk.

MATERIALS AND METHODS

Evolutionary Trace Analysis (ET)—Homologous sequences were obtained with BLASTp (39) using the NCBI nonredundant data base and aligned with ClustalW (40). After removal of sequences with gaps and outliers of the sequence similarity tree, we obtained an alignment of 231 sequences of the nuclear receptor superfamily. ET analysis was carried out using a similarity matrix (BLOSUM) (62) to allow for substitution of similar amino acids. In the absence of an existing GCNF crystal structure, trace residues were mapped onto a human RXRα structure (Protein Data Bank code 1FBY) (41, 42), which has 29% sequence identity to GCNF.

Peptide Design and Engineering—Functionally important helices H3 and H11 in the GCNF LBD were extracted from the LBD sequence in the form of soluble peptides whose helical propensities were calculated using AGADIR (43). H3 showed a high helicity in the native form (P3L), whereas two different H11 peptides were generated: one with native sequence and low helical propensity (P11L) and a second with nonnative sequence and with a high predicted helicity (P11-2). Only nontrace residues were changed in the H11 engineered peptide to increase helicity (see Table V). All peptides were synthesized by Invitrogen.

Point Mutation of GCNF Expression Vector—The GCNF expression vector was constructed by insertion of full-length GCNF cDNA into sites EcoRI and XhoI of pCMV-HA expression vector (Clontech) or pcDNA3.1(+) (Invitrogen). The point mutations were produced using the QuikChange XL site-directed mutagenesis kit (Stratagene) and performed according to the manufacturer's protocol. The primer sequences used for the point mutations are listed in Table IV. *In vitro* translated human COUP TF I vector (SP6-hCOUP TF I) (44) was mutated using the same methods as described for HA-GCNF.

Expression and Detection of Proteins in COS1 and P19 Cells—Wild type or mutated CMV-HA-GCNF expression plasmids (5 μg) were transfected into 1×10^6 COS1 or P19 cells for 48 h in the presence of 1 μM retinoic acid (Sigma) with the transfection reagent Fugene 6 (Roche Applied Science). Total proteins were extracted in extraction buffer (25 mM HEPES, pH 7.9, 150 mM KCl, 0.4 mM EDTA, 2 mM dithiothreitol, 20% glycerol, 1 mM phenylmethylsulfonyl fluoride, and 1× proteinase inhibitor mixture (Roche Applied Science). The expression of HA-GCNF was detected by Western blot using anti-HA antibody (Santa Cruz Biotechnology, Inc., Santa Cruz, CA) or anti-GCNF antibody (45) and visualized using the ECL detection system (Roche Applied Science).

Limited Proteolytic Digestion of in Vitro Translated Protein—In vitro translated GCNF or hCOUP TF I was synthesized using a T7 or SP6 *in vitro* translation kit (Promega). An aliquot (2 μl) of *in vitro* translated ^{35}S-labeled GCNF was digested with 0.5 μl of 100 μg/ml trypsin (Sigma) for 10 min and denatured by boiling in SDS-loading buffer and separated by 12% SDS-PAGE. The gel was dried and directly exposed to x-ray film.

Cell Culture and Transient Transfection—P19 cells were maintained in Dulbecco's medium supplemented with 10% fetal calf serum. CMV-HA-GCNF expression vector (200 ng) and *Oct4-Luc* reporter vector (400 ng) were co-transfected into P19 cells plated in 6-well plates (2×10^5 cells/well) for 24 h prior to transfection. The RL-Luc reporter (Promega) DNA was also co-transfected as an internal control to correct for differences in transfection efficiency. The empty vector pCMV-HA was used as control. After a 48-h incubation, the cells were harvested, and the luciferase activity was analyzed using the dual luciferase assay kit (Promega) according to the manufacturer's protocols. The *Oct4-Luc* luciferase reporter has been described previously (29). Relative luciferase activity was determined by dividing the activity of the corresponding luciferase reporter activity by that of the control, which was set at 1.

Electrophoretic Mobility Shift Assays (EMSA)—Total protein extracts from COS1 or P19 cells were incubated with ^{32}P-labeled DR0 DNA probe in the 0.5× extraction buffer supplemented with 67 ng/μl poly(dI-dC)·poly(dI-dC) (Roche Applied Science) and separated by 5% native polyacrylamide gel in the 0.5× TBE buffer. The sequence of the DR0 probe, and EMSA, were performed as previously described (29). For the antibody supershift and peptide blocking experiments, 1 μl of antibody or 1 μl of peptide was preincubated with GCNF proteins and then incubated with ^{32}P-labeled DNA probe. The final concentration of peptides in the binding reactions was 10 nM.

EMSA-Western Blot Analysis—Increasing amounts of COS1 extracts and P19 extracts were incubated with the DR0 DNA probe and separated on a 5% native polyacrylamide gel. The native gel was duplicated. One gel, which was used to separate the radioactive labeled DNA-protein complexes, was dried for detection of the radioactive signal produced by the ^{32}P-labeled DR0 DNA probe. The other gel, which was used to separate the proteins, was transferred onto polyvinylidene difluoride membranes for detection of GCNF protein by Western analysis. The radioactive signal produced by the bound DNA probe was measured by a PhosphorImager (Storm 860A; Amersham Biosciences) and quantitated with ImageQuant 5.2 (Amersham Biosciences); the amount of ECL signal produced by bound GCNF was determined by densitometry (Personal Densitometer SI; Amersham Biosciences) and quantitated with ImageQuant 5.2 (Amersham Biosciences).

RESULTS

Homodimerization and TRIF Complex Formation of GCNF Is DR0-dependent—To probe the impact of DNA binding on GCNF interactions, we compared the formation of the homodimer and the TRIF complex on the DR0 response element. Both the GCNF homodimer in COS1 cells and the TRIF complex in P19 cells bound to the *Oct4* DR0 probe (Fig. 1*A*, *lanes 2* and *4*). In addition, COUP TF I homodimers and SF1 monomers also bound to the DR0 probe (Fig. 1*A*, *lanes 6* and *8*) as previously shown (44, 46–48). Western blot analysis with GCNF-specific antibodies showed that the DR0 probe detected GCNF at the same position as the homodimer in COS1 cells and the TRIF complex in P19 cells (Fig. 1*B*, *lanes 2* and *4*), which confirms the presence of GCNF in the TRIF complex, as previously described (23, 29). Strikingly, when the DR0 probe was withheld from the reaction mix, GCNF was not detected in the Western blots at these positions in either COS1 cells or P19 cells; rather, it remained in the well, presumably as a very large molecular weight complex (Fig. 1*B*, *lane 3*). This is in sharp contrast to COUP TF I and SF1, which, under the same conditions and using the appropriate antibodies, are homodimeric (Fig. 1*B*, *lanes 5* and *6*) (44, 49) and monomeric (Fig. 1*B*, *lanes 7* and *8*) (50), respectively, independent of binding to the DR0 probe. Thus, in the absence of DNA binding, the GCNF LBD alone is insufficient to form a homodimer or the TRIF complex.

Oligomerization of GCNF in the TRIF Complex—Since the TRIF complex is a high molecular weight assembly that contains GCNF but not the nuclear receptor co-regulators, NCoR or SMRT, which are known to interact with GCNF (29, 51), it is possible that a GCNF multimer is the major, or even perhaps the sole, component of the TRIF complex. To support this hypothesis, we determined the molecular weight of the TRIF complex (52). Standard protein molecular weight markers were separated on the same native gel with COS1 and P19 extracts (Fig. 1*C*), and the relationship of molecular weight to the distance migrated was plotted (Fig. 1*D*). The distance migrated by the GCNF homodimer lies within the standard curve of the molecular weight markers, whereas the distance migrated by the TRIF complex lies outside of the standard curve. Its molecular weight was therefore estimated by extrapolation. From

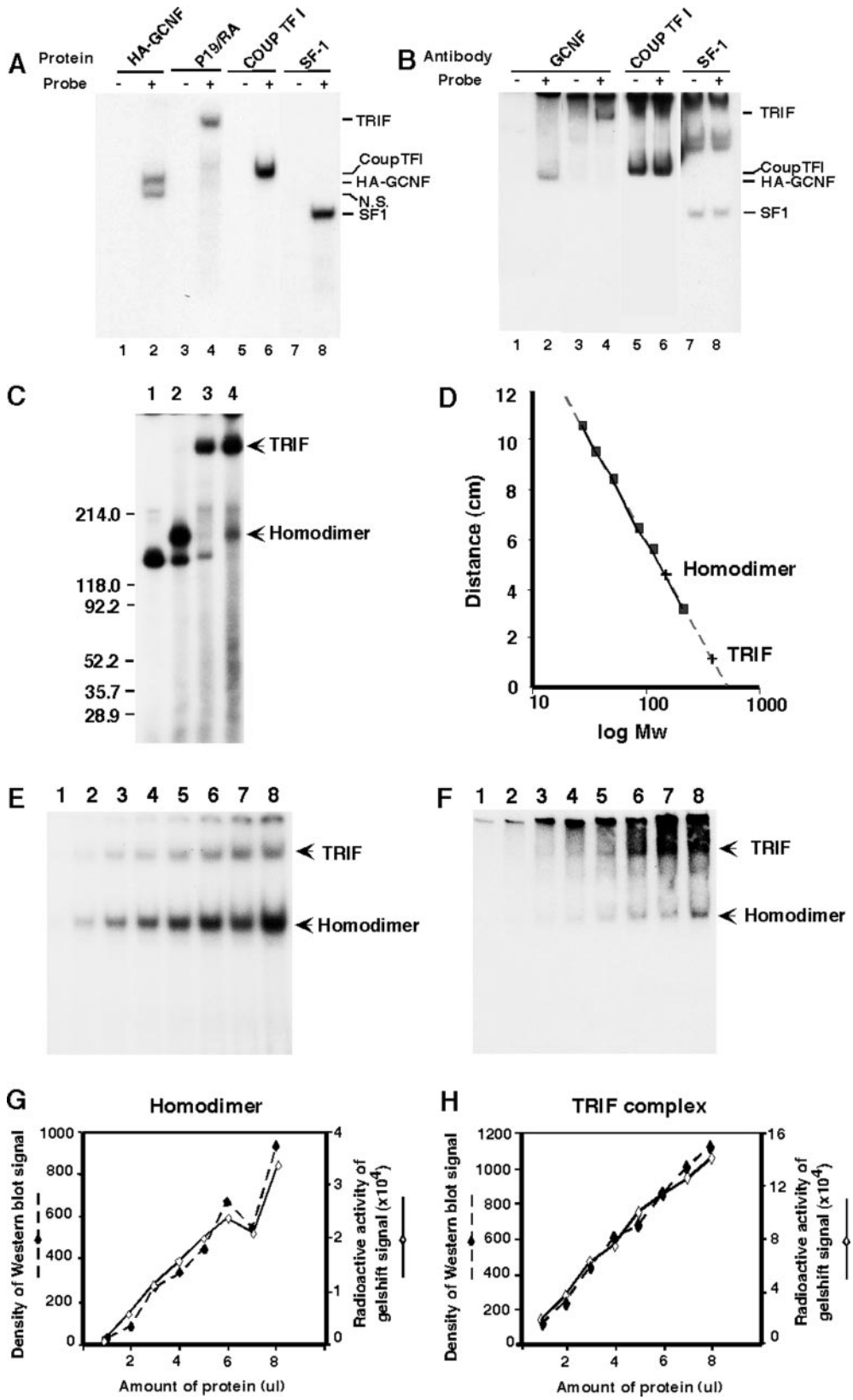

FIG. 1. **Homodimerization of recombinant GCNF and oligomerization of endogenous GCNF in P19 cells is dependent on binding to the DR0 DNA probe.** *A*, EMSA analysis of recombinant GCNF expressed in COS1 cells (*lanes 1* and *2*), endogenous GCNF in differentiated P19 cell extract (*lanes 3* and *4*), human COUP TF I-expressed COS1 cells (*lanes 5* and *6*), and mouse SF1-expressed COS1 cells (*lanes 7* and *8*) in the absence (*lanes 1*, *3*, *5*, and *7*) or presence (*lanes 2*, *4*, *6*, and *8*) of radiolabeled DR0 probe. *B*, analysis of GCNF, COUP TF I, and SF1 expression by Western blot of transferred duplicated EMSA gels. Oligomerization of GCNF in differentiated P19 cell extracts is shown. *C*, EMSA shows the different mobilities of the GCNF homodimer and the TRIF complex. Total protein extract from COS1 cells (*lane 1*), HA-GCNF-transfected COS1 cells (*lane 2*), differentiated P19 cells (*lane 3*), and HA-GCNF-transfected differentiated P19 cells (*lane 4*), were incubated with DR0 DNA probe. The molecular weights of standard protein markers (Bio-Rad) are indicated on the *left*. *D*, estimation of the molecular weight of the GCNF homodimer expressed in COS1 cells and the TRIF complex in P19 cells relative to the standard molecular weight curve. *E*, EMSA results demonstrate the amount of bound DR0 probe by GCNF homodimer and TRIF complex. Increasing amounts of HA-GCNF-transfected COS1 cell

TABLE I
Estimation of molecular weight of different GCNF/DNA stoichiometries

Complex[a]	DR0·GCNF[b]	Calculated molecular weight of the DRO·GCNF complex[c]
Dimer	1:2	144.5 (149.1)
Tetramer	1:4	263.3 (272.5)
	2:4	289.0 (298.2)
Hexamer	1:6	382.1 (395.9)
	2:6	407.8 (421.6)
	3:6	433.5 (447.3)

[a] Based on the number of GCNF molecules in a complex.
[b] Molecular ratio of DR0 DNA probe to GCNF.
[c] Calculated molecular weight of DR0·HA-GCNF complex is shown in parentheses.

TABLE II
Ratio of radioactive signal to ECL signal in GCNF binding complexes

Ratio	Amount of protein								Average
	1	2	3	4	5	6	7	8	
				μl					μl
Homodimer	336	595	407	453	437	351	377	360	415 ± 84
TRIF	133	146	136	120	143	130	123	124	132 ± 10

the standard curve, the estimated molecular weight of the homodimer is 157, and that of the TRIF complex is 373. According to the molecular weight of the DR0 DNA probe and GCNF or HA-GCNF, the calculated molecular weights of the DNA·GCNF complexes formed by different combinations of the DR0 probe and GCNF subunits are listed in Table I. When compared, the estimated and calculated molecular weights of the GCNF COS-1 homodimer are very similar. The estimated molecular weight of the TRIF complex is very similar to that predicted for a hexamer of GCNF but considerably different from that predicted for a homodimer or tetramer.

To determine the stoichiometry of DR0 probe to GCNF in the TRIF complex, we designed an experiment to combine the EMSA and Western blot analyses. We mixed COS1 cell extracts, P19 cell extracts, and radiolabeled DNA probe. Each sample was divided in two; half was used for EMSA, and the rest was used for Western blot analysis. The probe-homodimer complex was separated from the probe-TRIF complex on the same native gel (Fig. 1, *E* and *F*). The amount of radioactive signal in the shifted complexes was determined by phosphor image analysis, and the amount of GCNF protein in the shifted bands was determined simultaneously from the ECL signal of the native gel subjected to Western blot. The combination of the two analyses measured the relative amounts of GCNF protein and DR0 DNA in each shifted complex. When the amount of bound probe *versus* the amount of GCNF in the COS1 cell homodimer and the P19 cell TRIF complex was plotted (Fig. 1, *G* and *H*), it was found that the ratio between the two was consistently 3 (Table II). Since the stoichiometry of GCNF homodimer to DR0 probe is 2:1, we estimated that the TRIF complex stoichiometry is 6:1, or 3 times larger. Thus we propose that the TRIF complex is a GCNF hexamer that binds a single DR0 probe.

Evolutionary Trace Predicts Functional Surfaces on Helix 3 and Helix 11—The unique characteristics of GCNF DNA binding activity, and the formation of multimers in the TRIF complex, led us to mine the structural information hidden in the amino acid sequence of GCNF and the nuclear receptor family as a whole. To probe the evolutionary origin of GCNF/GCNF interactions, we turned to the ET method to identify possible LBD functional sites. The results of ET analysis of 231 NR LBD sequences were mapped onto an RXRα structure (Protein Data Bank code 1FBY) (Fig. 2), because the GCNF LBD has not been crystallized (41, 42). ET analysis identified the ligand binding pocket, as expected, and also two distinct and large surface clusters on the solvent-accessible faces of helix 3 and helix 11, which are statistically significant (z-score = 3.2) at a coverage of 18% (Fig. 2*A*). The six surface trace residues on H3 of RXRα are Asp^{273}, Lys^{274}, Phe^{277}, Val^{280}, Lys^{284}, and Arg^{284}, corresponding to the cognate residues Asp^{307}, Glu^{308}, Phe^{311}, Ile^{314}, Lys^{318}, and Lys^{319} of GCNF (Tables III and IV and Fig. 2*B*). The four trace residues on H11 of RXRα are Leu^{422}, Arg^{426}, Leu^{430}, and Glu^{434}, corresponding to cognate residues Leu^{459}, Arg^{463}, Gly^{467}, and Asn^{471} of GCNF (Tables III and IV and Fig. 2*B*). Trace residues in H11 are generally ranked higher than their counterparts in H3, consistent with a more widespread pattern of homodimerization and heterodimerization via helix 10/11 interactions (16, 17, 53, 54). The lower, still significant, ranked residues of H3 are consistent with the more diverse functions among different NR subfamilies in that helix (20, 28, 55–57). ET analysis suggests that both H3 and H11 have important biological roles among the nuclear receptors analyzed, including GCNF, and that mutations of trace residues in these helices should disrupt their normal interactions and reveal their specific function.

Repression Function of GCNF Point Mutants on Oct4 Reporter Gene Expression—To test the validity of the ET predictions, we generated targeted point mutations in H3 and H11 surface trace residues and assayed their effect on GCNF repression of *Oct4* reporter gene expression (29). Trace residue substitutions introduced amino acids not observed among the natural evolutionary variations of cognate residues in our alignment of 231 NR sequences in order to have the greatest chance of generating a loss of function mutation for GCNF (Table III). For example, residue Phe^{311} in H3 of mouse GCNF occasionally varies to valine in other NRs but never to glycine, which was thus our choice of mutation (Table IV). As controls, we introduced two sets of double mutations in the hinge region and helix 12, which had been previously reported as important to GCNF function, and for interaction with the corepressor NcoR (29, 51). HA-tagged GCNF point mutant expression vectors were co-transfected with an *Oct4* luciferase reporter in undifferentiated and retinoic acid (RA) differentiated P19 cells. There were no significant differences in the expression levels between the GCNF mutants and the wild type HA-GCNF in P19 cells (Fig. 3, *C* and *D*). As expected, wild type HA-GCNF caused a 65% reduction in *Oct4* reporter gene activity compared with controls (empty vector pCMV-HA). Of the eight mutations targeted at trace residues, five completely impaired the ability of GCNF to repress transcription. Specifically, GCNF mutants, E308A, A/K318, and K319W in helix 3 and L459K and R463E in helix 11, did not repress *Oct4*-luciferase reporter gene activity in either undifferentiated or differentiated P19 cells. In contrast, mutants F311G, K318A, and N471W substantially reduced *Oct4-Luc* activity, comparable

extracts, and RA-treated P19 cell extracts, were incubated with same amount of DR0 probe, respectively, and mixed together for loading. *Lanes 1–8* indicate the added protein amount from 1–8 μl of COS1 cell extracts (2.5 μg/μl total protein) or P19 cell extracts (7.5 μg/μl total protein). *F*, detection of GCNF protein transferred from duplicated EMSA gels with anti-GCNF antibody by Western blot analysis. *G* and *H*, plot of radioactive signal and Western blot signal relative to amount of the added protein. The bound radioactive signal was quantitated by phosphorimaging, and GCNF protein levels were measured by the densitometry of the Western blots. The *arrows* mark the migration of the GCNF homodimer COS1 cell extracts and the TRIF complex in differentiated P19 cells.

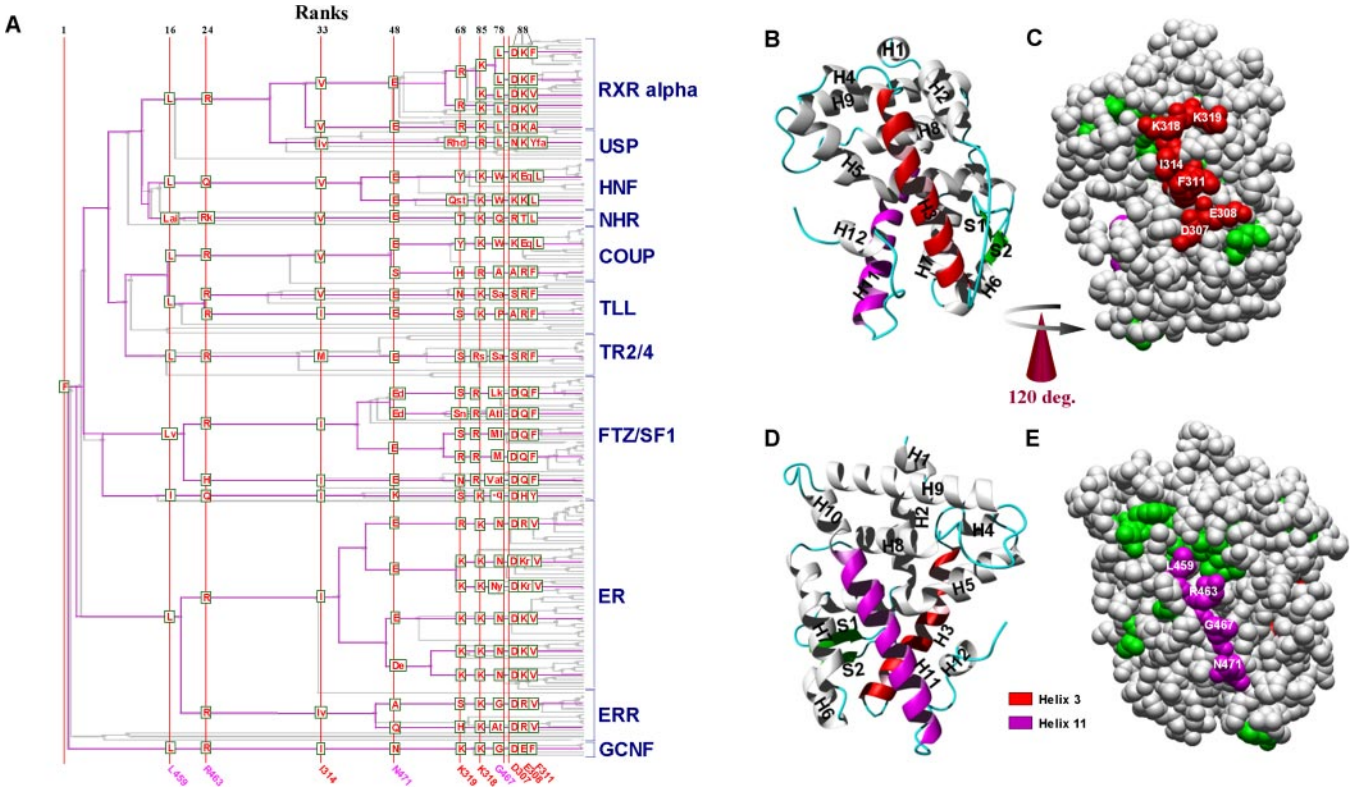

FIG. 2. **Evolutionarily important surface residues in helices 3 and 11 of GCNF mapped onto a RXRα structure (Protein Data Bank code 1FBY) (41).** *A*, the evolutionary tree of 231 nuclear receptors. Evolutionary ranks are labeled at the *top* of each *vertical line*, with corresponding H3 or H11 mouse GCNF surface trace residues annotated at the *bottom*. H3 surface trace residues are shown in *red*, whereas H11 surface trace residues are shown in *magenta*. The *large letters* at the *nodes* represent the most frequent amino acids in that branch, and *small letters* represent the less frequent. *B* and *C*, helix 3 and helix 11 of RXRα LBD are shown by *ribbon representation* in *red* and *magenta*, respectively. *D* and *E*, evolutionarily important amino acids are shown to cluster on the surface of helix 3 and helix 11 and are mapped onto the RXRα LBD in *red* and *magenta*, respectively. All other trace residues are shown in *green*. The surface trace residues of H3 and H11 are labeled with their mGCNF cognates on the RXRα structure.

TABLE III
Variability of traced amino acids on the surface within the nuclear receptor superfamily
Values represent percentage variability from an alignment of 231 sequences of nuclear receptor.

Helix 3 trace residues					Helix 11 trace residues		
Asp^{307} (88)[a]	Glu^{308} (88)	Phe^{311} (88)	Lys^{318} (78)	Lys^{319} (68)	Leu^{459} (16)	Arg^{463} (24)	Asn^{471} (48)
Ala 15.28%	Glu 6.55%	Ala 1.75%	Lys 65.94%	Ala 0.44%	Ala 0.44%	His 3.49%	Ala 2.62%
Asp 62.88%	Gly 0.44%	Asp 0.44%	Leu 1.75%	Asp 1.31%	Ile 3.06%	Lys 1.31%	Asp 10.48%
Glu 0.44%	His 1.31%	Glu 0.44%	Arg 32.31%	His 6.99%	Leu 94.32%	Gln 8.30%	Glu 75.98%
His 0.87%	Lys 32.75%	Phe 46.29%		Lys 24.45%	Val 2.18%	Arg 86.90%	Lys 1.31%
Lys 6.55%	Pro 0.44%	His 0.87%		Leu 0.44%			Asn 2.62%
Leu 0.44%	Gln 21.4%	Ile 4.37%		Met 1.31%			Gln 3.93%
Asn 5.24%	Arg 34.93%	Leu 8.30%		Asn 12.66%			Ser 3.06%
Arg 1.31%	Ser 0.44%	Val 34.50%		Gln 0.87%			
Ser 5.68%	Thr 1.75%	Tyr 3.06%		Arg 22.27%			
Thr 0.87%				Ser 22.27%			
GAP 0.44%				Thr 1.75%			
				Tyr 5.24%			

[a] Mouse GCNF residues are shown in the heading, followed by ET rank in parentheses.

with wild type GCNF. The control mutations in the hinge region and helix 12 did not affect the repression function of GCNF (Fig. 3, *A* and *B*). Dose-dependent repression assays further confirmed the loss of repression function for mutations (E308A, A/K318, K319W, L459K, and R463E) (Fig. 3*C*). In summary, mutation of many of the trace residues in both helix 3 and helix 11 caused a loss of GCNF repression function *in vivo*, which confirmed their functional relevance.

Disruption of GCNF Homodimerization and TRIF Complex by H3 and H11 Mutations—To probe a possible link between loss of transcriptional activity and GCNF DNA binding and dimerization, we asked whether trace residue mutations also affect homodimer and TRIF complex formation. All mutants were shown, by Western blot analysis, to be highly expressed at similar levels in COS1 cells, suggesting that the various point mutations had not altered the stability of the protein (Fig. 4*A*). The mutations, S246G,Y247G in the hinge region, F311G in helix 3, and K489A,T490A, in helix 12 did not affect DNA binding or dimerization; however, mutations E308A, A/K318, and K319W in helix 3 and L459K in helix 11 completely abolished binding of the GCNF homodimer formed in COS1 cells to the DR0 (Fig. 4*B*). Mutations, K318A in helix 3 and R463E and N471W in helix 11, significantly weakened the DNA binding affinity of the GCNF homodimer (Fig. 4*B*).

TABLE IV
Point mutations of amino acids of mouse GCNF and human COUP TF I corresponding to human RXRα

ET Rank	Human RXRα	Mutations in mGCNF	Sequence of sense primer for mutations	Region
		S246G,Y247G	5′-CATACCACACCTTTTTGGCGGTTCTGGCCACTCACCAC-3′	Hinge
88	Lys^{274}	E308A	5′-GCCGCCTGGCCGACGCGTTGCTCTTTAGGC-3′	Helix 3
88	Phe^{277}	F311G	5′-GGCCGACGAGTTGCTCGGTAGGCAGATTGCCTGG-3′	
78	Lys^{284}	K318A	5′-GGCAGATTGCCTGGATCGCGAAGCTGCCTTTCTTCTGC-3′	
		A/K318	5′-GGCAGATTGCCTGGATCGCGAAGAAGCTGCCTTTCTTCTGC-3′	
68	Lys^{285}	K319W	5′-GGCAGATTGCCTGGATCAAGTGGCTGACTTTCTTCTGC-3′	
		K319A	5′-GGCAGATTGCCTGGATCAAGGCGCTGACTTTCTTCTGC-3′	
16	Leu^{422}	L459K	5′-CCTGATCTTATGATGTGCAAGCCAGAGATCCGATACATCGC-3′	Helix 11
		L459A	5′-CCTGATCTTATGATGTGCGCGCCAGAGATCCGATACATCGC-3′	
24	Arg^{426}	R463E	5′-GCTTGCCAGAGATCGAATACATCGCAGGCAAGATGG-3′	
		R463A	5′-GTGCTTGCCAGAGATCGCATACATCGCAGGCAAGATGG-3′	
48	Glu^{434}	N471W	5′-CGCAGGCAAGATGGTGTGGGTGCCCCTGGAGCATG-3′	
		N463A	5′-CGCAGGCAAGATGGTGGCTGTGCCCCTGGAGCATG-3′	
		K489A,T490A	5′-GGTGCTGCACTCCTGCGCGGCAAGTACGGTGAAGG-3′	Helix 12
		Mutations in hCOUP TF I		
88	Lys^{274}	R225A	5′-CTGCGAGCTGGCCGCGGCCCTGCTCTTCAGCGCCG-3′	Helix 3
68	Lys^{285}	N276W	5′-CCGTCGAGTGGGCCCGCTGGATCCCCTTCTTCCCGG-3′	
16	Leu^{422}	L374K	5′-GGCAAACTGCTGCTGCGAAAGCCCTCGCTGCGCACC-3′	Helix 11

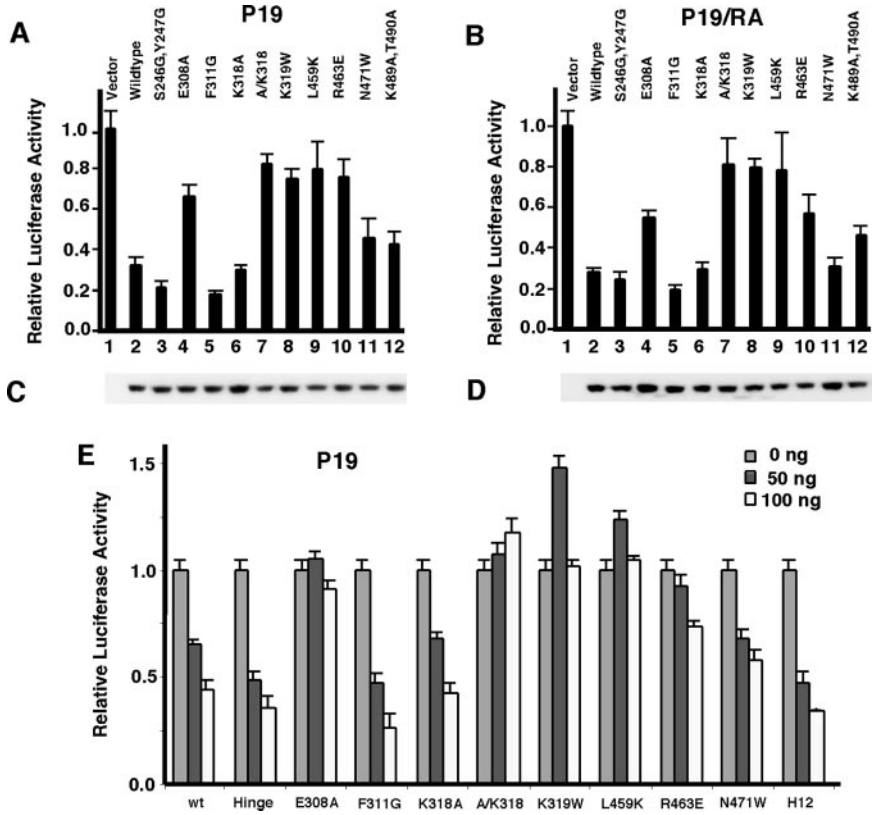

FIG. 3. GCNF helix 3 and helix 11 point mutants lose their ability to repress the *Oct4* promoter in undifferentiated (*A*) and differentiated (*B*) P19 cells. The luciferase reporter activity was determined relative to the CMV-HA empty vector. Transfected wild-type and mutated HA-GCNF expression levels in undifferentiated (*C*) and differentiated (*D*) P19 cells were detected with anti-HA antibody by Western blot analysis. *E*, dose-dependent repression activity of GCNF mutants on the *Oct4* promoter in undifferentiated P19 cells. Relative luciferase activity was measured according to the ratios of *Oct4* luciferase activity to the internal control RL-luciferase activity, and the empty vector CMV-HA co-transfected relative luciferase activity was set at 1. The experiments were repeated three times, and the S.E. values in triplicate points are shown as *error bars* for each *column*.

As a control of specificity for the disruptive effects of H3 mutations on dimerization and DNA binding, point mutations were introduced into H3 and H11 of COUP TFI, since it binds to the same DR0 DNA probe as GCNF but has not been found to form a TRIF-like complex in RA-induced P19 cells. Three cognate point mutations were produced in H3 and H11 of COUP-TFI (Table IV). EMSA analysis demonstrated that the two mutations R225A and N276W in helix 3 had no effect on DNA binding of *in vitro* translated COUP TF I homodimers (Fig. 4, *E* and *F*, *lanes 2* and *3*). As anticipated, the mutation L374K in helix 11 completely abolished DNA binding of COUP TF I (Fig. 4*F*, *lane 4*). Thus, as expected, helix 11 mutations affect COUP TF dimerization, but those in H3 do not, highlighting the view that H3 has distinctly differing functions in GCNF and COUP TF.

Interestingly, when we performed supershift analysis of COS1 extracts with an anti-GCNF antibody, DNA binding functionality was rescued in those GCNF mutants that had lost DNA binding (Fig. 4*C*). Thus, the GCNF antibody, in contrast to the HA tag antibody (Fig. 4*D*), was able to rescue the DNA binding mutants, which is important for two reasons. First, it confirmed that the nonfunctional GCNF mutants were not grossly altered in their folding and retained their potential to bind DNA, and, second, cross-linking with the GCNF antibodies rescued DNA binding, supporting the contention that the disruption of GCNF/GCNF interaction within the LBD led to loss of DNA binding.

To test the affect of the point mutations on TRIF complex formation, each mutant was transfected into P19 cells. Transfected HA-GCNF was detected by anti-GCNF antibody in differentiated P19 cells (Fig. 5*A*). The expressed HA-GCNF is incorporated efficiently into the TRIF complex (Figs. 4*C* and 5*B*), as confirmed by anti-HA antibody supershift (Fig. 5*D*). Significantly, the same point mutants that affected GCNF ho-

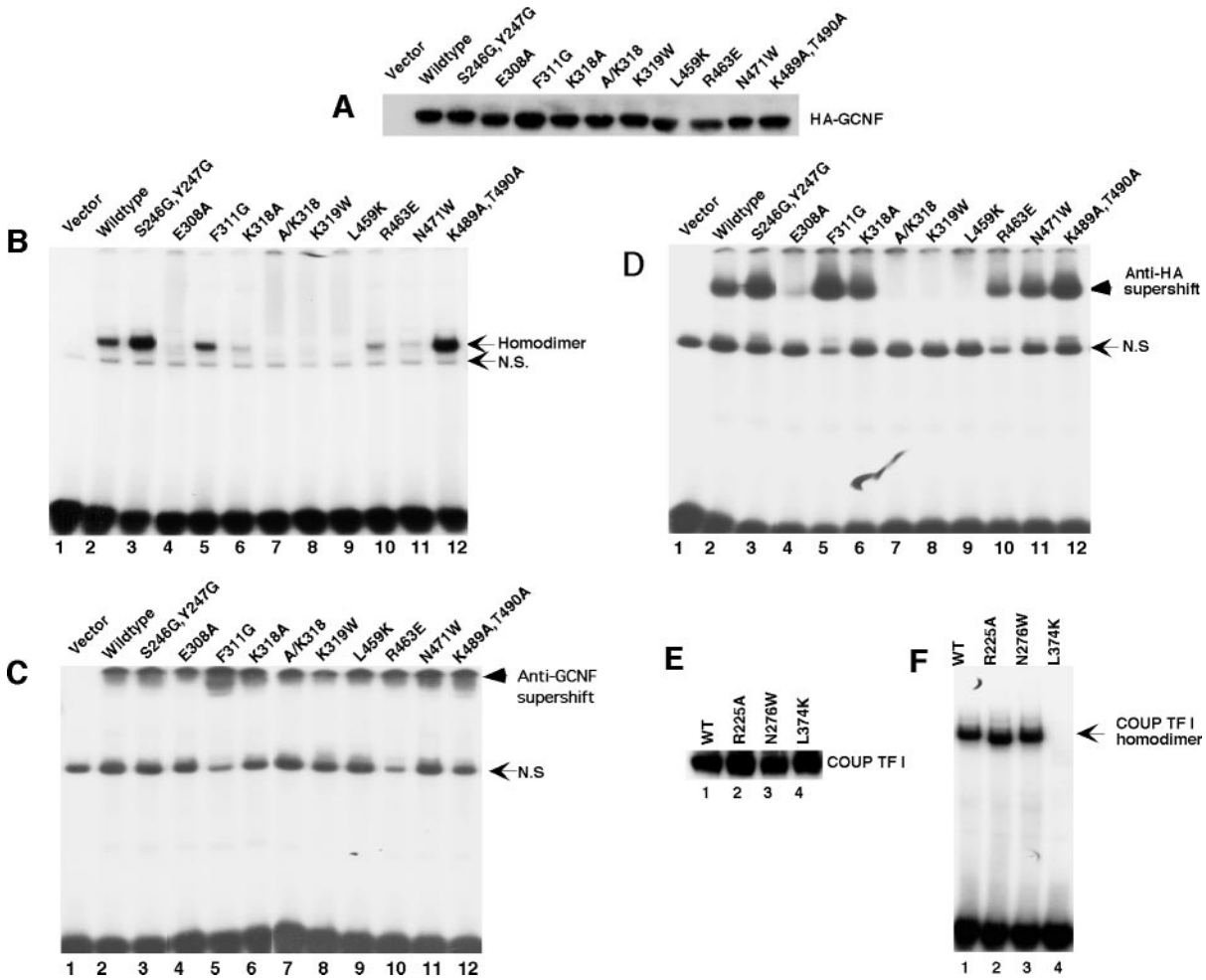

FIG. 4. **Disruption of GCNF homodimerization by helix 3 and helix 11 point mutations in GCNF.** *A*, detection of HA-GCNF point mutant expression in COS1 cells by Western blot with anti-HA antibody. *B*, EMSA analysis of DNA binding of HA-GCNF point mutants expressed in COS1 cells. *C* and *D*, supershift of HA-GCNF homodimers expressed in COS1 cells using anti-GCNF antibody (*C*) and anti-HA antibody (*D*). *E*, detection of mutant COUP TF expression by Western blot analysis. *F*, EMSA analysis of dimerization and DNA binding of COUP TF mutants synthesized by *in vitro* translation. The *arrows* indicate the migration of the GCNF homodimer. The *arrowheads* indicate the supershift produced by anti-GCNF and anti-HA antibody. *N.S.*, nonspecific signals.

modimerization in COS1 cells also disrupted the TRIF complex in P19 cell extracts. Meanwhile, the GCNF mutant controls that did not affect homodimerization in COS1 cells also did not affect HA-GCNF incorporation into the TRIF complex (Fig. 5*D*).

Intact Conformation in Mutant GCNF—The choice of amino acid substitutions was based on the evolutionary trace; thus, some of the substitutions significantly change the physiochemical character of amino acids at the particular position (*e.g.* K319W, L459K, R463E, and N471W) and have the potential to cause misfolding of the GCNF LBD and loss of DNA binding activity. To address this possibility, alanine substitutions were introduced at these sites. The effects of these alanine substitutions on the *Oct4* reporter activity and DNA binding were compared with their original amino acid substitutions (Fig. 6, *A* and *B*). It was found that the replacement of Lys^{319} and Leu^{459} with alanine resulted in the total loss of repression function on the *Oct4* reporter in undifferentiated and differentiated P19 cells, similar to the original replacements of K319W and L459K. Similarly, R463A and N471A both partially lost repressive function as R463E and N471W (Fig. 6, *A* and *B*). The dimerization and DNA binding activities of alanine mutants were detected and compared with their original mutations (Fig. 6*C*). The results showed that the different amino acid substitutions caused the same effect on the DNA binding and dimerization functions. K319A and L459A mutants could not form the dimer as did K319W and L459K. The dimerization abilities of R463A and N471A were similar to R463E and N471W and much weaker than the wild type GCNF.

To further confirm that the amino acid substitutions did not affect the entire molecular conformation or folding of the LBD, the *in vitro* translated GCNF mutants were partially digested with trypsin. Partial proteinase digestion is a sensitive assay for nuclear receptor LBD conformation (58, 59). Nuclear receptor LBDs tend to be resistant to proteinase digestion, and if they become misfolded, will be rapidly degraded like the rest of the molecule. As shown in Fig. 6*D*, *panel b*, the LBD of wild type GCNF was resistant to trypsin digestion (~23 kDa band). All of the GCNF mutants (either alanine substitution or other amino acid substitutions) were resistant to the trypsin treatment in their LBDs and produced the same size of LBD fragment as wild type GCNF. However, deletion of helix 12 resulted in loss of the 23-kDa resistant fragment of GCNF LBD and suggested a significant change of conformation (20). Therefore, the amino acids substitutions did not disrupt the entire molecular folding or conformation of the LBD but rather specifically interfered with the GCNF/GCNF interaction at their interaction surfaces.

Disruption of GCNF Homodimerization and TRIF Complex by H3 and H11 Peptides—Our data thus suggested that both H3 and H11 contribute to GCNF dimerization, TRIF complex

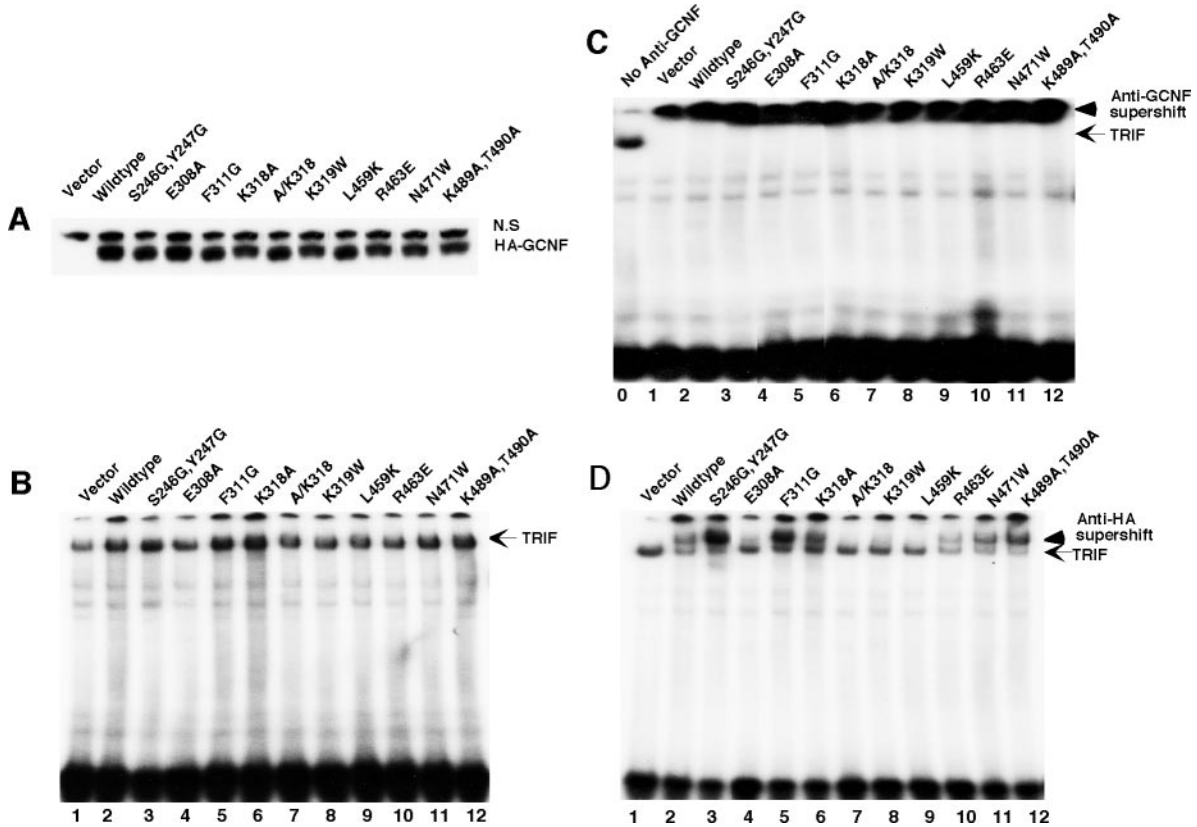

FIG. 5. **Disruption of TRIF complex incorporation by point mutations in helix 3 and helix 11 of GCNF.** *A*, detection of HA-GCNF point mutant expression in RA-differentiated P19 cells by Western blot with anti-HA antibody. *B*, EMSA analysis of DNA binding of HA-GCNF point mutants incorporated into the TRIF complex in P19 cells. *C* and *D*, supershift of HA-GCNF incorporated into the TRIF complex in P19 cell extracts using anti-GCNF antibody (*C*) and anti-HA antibody (*D*). The *arrows* indicate the migration of the GCNF TRIF complexes. The *arrowheads* indicate the supershift produced by anti-GCNF and anti-HA antibody.

formation, and DNA binding. This is in direct contradiction to deletion experiments that showed that helix 3 alone in the LBD is responsible for GCNF dimerization (20), suggesting a number of alternative hypotheses illustrated in Fig. 7*A*: either H3 is solely responsible for dimerization and the effect of the mutation of H11 residues is purely allosteric, or likewise, H11 is solely responsible for dimerization and the effect of mutations of H3 residues is purely allosteric, or alternatively, both H3 and H11 are involved in a direct interaction. The major distinction between the three mechanisms is that the former two are examples of homotypic interaction, whereas the latter would be a heterotypic interaction that may lead to oligomer formation.

To distinguish between these competing mechanisms, we designed peptides to mimic each of the putative binding surfaces identified by ET on H3 and H11. The rationale being that both peptides should competitively inhibit heterotypic interaction via H3 and H11, but that only one or the other peptide would inhibit homotypic interaction, depending on whether they were mediated by H3 or H11. In contrast, both of the peptides would inhibit the heterotypic interaction in homodimer and TRIF complex formation (illustrated in Fig. 7*A*). Such ET-based rational design is facilitated by the fact that the trace residues concerned are on one face of the respective helices. Using the program AGADIR (43) to predict the helical propensity of peptides, we designed peptides based on H3 and H11 that preserved the relative positioning of identified surface trace residues within their respective helices. Substitutions at other positions were introduced to enhance α-helix formation in P11, as detailed in Table V. Thus, when the first peptide, P3, had no effect on the binding of the GCNF homodimer or TRIF complex to the DR0 probe (Fig. 7*B*, *lanes 2* and *6*), we added five amino acids at the N terminus to create P3L, a second peptide predicted to have greatly increased helicity. P3L specifically disrupted the binding of the GCNF homodimer and TRIF complex to the DR0 probe (Fig. 7*B*, *lanes 3* and *7*). To further confirm the specific effect of P3L on GCNF homodimerization and TRIF complex formation, the same mutant amino acids used in the point mutation experiments were introduced into the peptides. As expected, these modified versions of peptide P3LmDE and P3LmKK failed to disrupt either the GCNF homodimer or TRIF complex binding to DR0 (Fig. 7*B*, *lanes 4*, *5*, *8*, and *9*).

When we tested the first H11 peptide, P11L, it had no effect on GCNF homodimer or TRIF complex formation. The P11 peptide was reengineered based on the trace results for H11. The helical score of the new peptide, P11-2, was drastically increased (76%) compared with the native P11L peptide (<2%) by introducing helix-forming amino acids at nontrace residue positions in the sequence of H11 (Table V). As with P3L, P11-2 disrupted both binding of the GCNF homodimer and the TRIF complex to the DR0 probe (Fig. 7*C*, *lanes 3* and *7*). Significantly, P3 and P11-2 disruption of TRIF complex binding occurred without the generation of a homodimer, meaning that it is impossible for GCNF to exist in this interaction, H3/H3 and H11/H11, in the TRIF complex as illustrated in Fig. 7*A*. When the point mutation L459K was introduced in the P11-2 peptide, the mutant peptide P11-2mL lost the ability to disrupt homodimerization and TRIF complex formation (Fig. 7*C*, *lanes 4* and *8*). Therefore, the peptide blocking experiments demonstrated that mutations in both H3 and H11 impaired the blocking activities of peptide 3 and 11 on dimerization and TRIF complex formation suggesting that both H3 and H11 are involved in the GCNF/GCNF interaction. Conclusively, both P3 and P11 peptides disrupted homodimer and TRIF complex

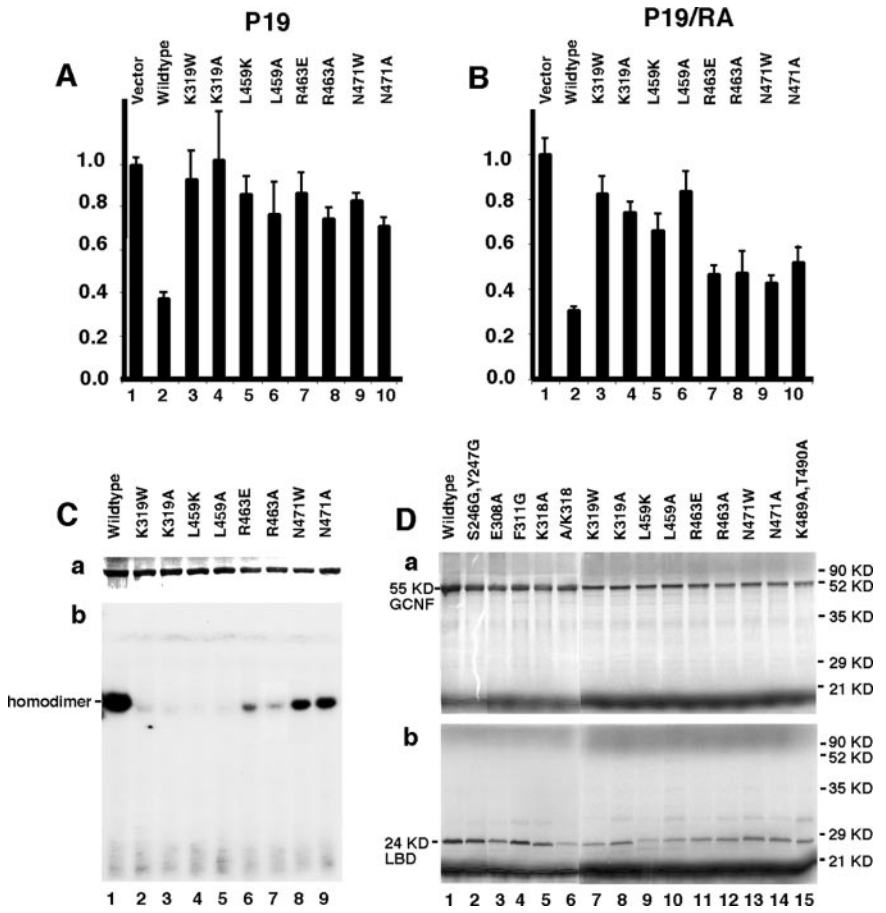

FIG. 6. **Comparison of the effect of different amino acid substitutions at the same site on GCNF function and proteinase resistance.** *A* and *B*, loss of repression activity on *Oct4* reporter in undifferentiated P19 cells (*A*) and differentiated P19 cells (*B*). The detailed descriptions are given in the legend to Fig. 3. *C*, *a*, expression of *in vitro* translated GCNF mutants; *b*, EMSA analysis of homodimer binding activity of GCNF mutants. *D*, partial trypsin digestions of *in vitro* translated and ^{35}S-labeled GCNF mutants were separated on 12% SDS-PAGE without trypsin treatment (*a*) or with 100 μg/ml trypsin treatment for 10 min (*b*).

formation, and neither peptide generated a homodimer, which is consistent with the results predicted for a H3/H11 heterotypic interaction (Fig. 7*A*).

DISCUSSION

DR0-dependent Dimerization and TRIF Complex Formation—Within the nuclear receptor evolutionary tree, GCNF forms its own subfamily (NR6A) (1). Although GCNF is closely related to the SF-1/LRH-1/FTZ-F1 (NR5A) subfamily and COUP TF I/II (NR2F) in DNA binding properties (*i.e.* binding to the DR0 element), there are clear mechanistic differences between these factors. In contrast to COUP-TF, which forms dimers in the absence of DNA, GCNF forms a homodimer or TRIF complex, which is dependent on binding the DR0 element. This finding of DNA-dependent dimerization of GCNF is consistent with our failure to detect direct interaction between GCNF LBDs using yeast two-hybrid, mammalian two-hybrid, and co-immunoprecipitation assays (data not shown). Thus, our data show that 1) in the absence of binding DNA, GCNF forms neither a monomer, homodimer, nor TRIF complex; 2) the interaction with DNA in the form of a DR0 induces GCNF homodimerization and TRIF complex formation. Thus, GCNF may offer a unique insight into nuclear receptor dimerization and complex formation. In comparison, a two step-model for dimeric binding of RXR heterodimers to DNA has been proposed. The first step is formation of a heterodimer through the interaction surfaces contained in the LBD, and the second step involves induced dimerization of the DNA binding domains upon binding to DNA (53). The DNA binding mechanism of GCNF is clearly different, in that the two-step process appears to be reversed. The first step is DNA binding domain dimerization and DNA binding, and the second step is induced dimerization of the LBDs. The unusual DNA binding and dimerization properties of GCNF and the differences between the recombinant and endogenous GCNF complexes led us to further investigate the novel mechanism underlying these phenomena.

Asymmetric Dimerization and Oligomerization of GCNF—Our results show that two evolutionarily privileged surfaces on H3 and H11 of the NR LBD are concurrently involved in GCNF DNA binding by forming a direct interaction in recombinant GCNF homodimers and in the endogenous GCNF TRIF complex. The alternative models for homotypic association necessarily require all of the following conditions to be simultaneously met: 1) dimerization dependent on symmetrical H3-H3 or H11-H11 association; 2) mutations of nearly all surface trace residues in H11 allosterically affect H3, so it cannot dimerize, and *vice versa*; 3) H11 mutations at nontrace residues have no such allosteric effect on H3-H3 associations, and *vice versa*; 4) the P11-2 peptide binds some part of GCNF and allosterically inhibits H3-H3 binding, and the P3L peptide allosterically affects H11-H11 binding. The simplicity of the H3-H11 model *versus* the many prerequisites of the H3-H3 or H11-H11 models argues in favor of the former. Additionally, the asymmetric H3-H11 heterotypic interaction model also provides for a novel multimerization mechanism for GCNF. Thus, we favor the hypothesis that both helix 3 and helix 11 play a direct role in GCNF homodimerization and TRIF complex formation and that the interaction surface is asymmetric in nature.

The biologically active form of GCNF is the TRIF complex, and the dimeric form is only detected with overexpressed, recombinant GCNF (20, 28). Previous studies conducted with *in vitro* translated GCNF showed that H3, but not H11, is involved in GCNF dimerization (20). Greschik *et al.* (20) made a series of C-terminal deletions in GCNF and showed that the

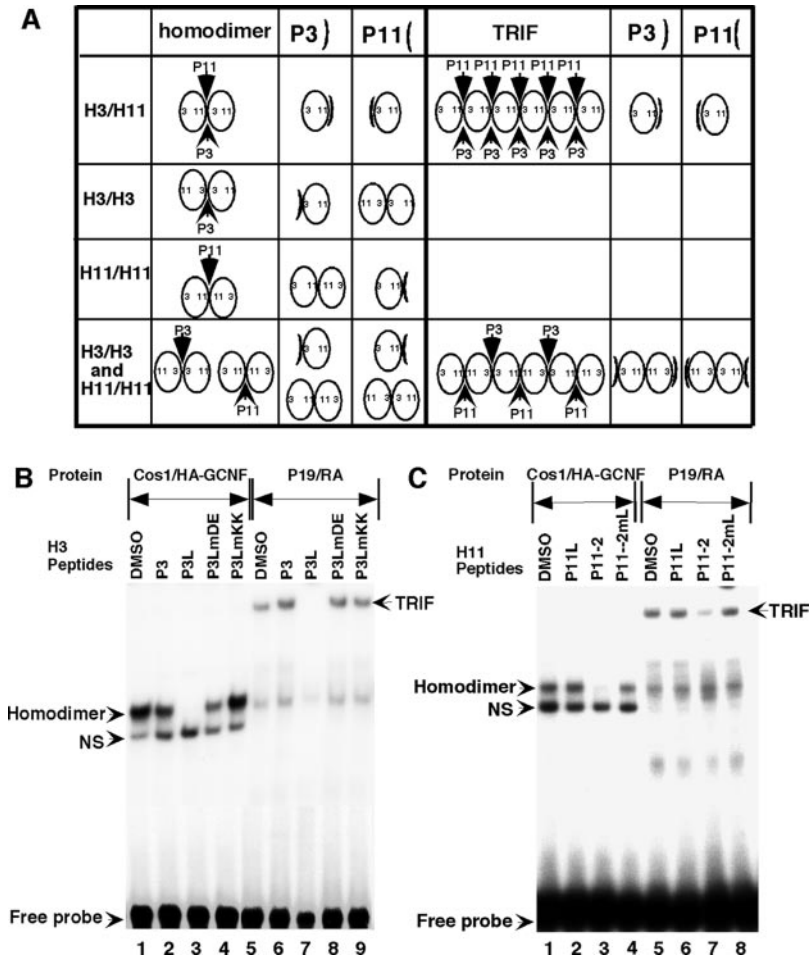

FIG. 7. **Disruption of GCNF homodimer or TRIF complex binding to DRO probes by helix 3- and helix 11-derived peptides.** *A*, illustration of three possible interaction surfaces in GCNF homodimer and TRIF complex and corresponding predicted peptide disruption results. *B*, EMSA analysis of competition with H3-derived peptides for COS1-expressed GCNF homodimer and P19 cell TRIF complex DNA binding. *C*, EMSA analysis of competition with H11-derived peptides for COS1-expressed GCNF homodimer and P19 cell TRIF complex DNA binding. Peptides were diluted with binding buffer and preincubated with protein extracts to a final concentration of 10 nM before the addition of the DNA probes. The amino acid sequences of peptides are listed in Table V.

TABLE V
Amino acid sequence and predicted helicity of peptides used for blocking

Peptides	Amino acid sequence	Helicity
P3 (wild type)	A**DE**LLFRQIAWI**KK**	35%
P3L (wild type)	ALLCRLA**DE**LLFRQIAWI**KK**	65%
P3LmDE	ALLCRLA**WA**LLFRQIAWI**KK**	86%
P3LmKK	ALLCRLA**DE**LLFRQIAWI***AA***	54%
P11L (wild type)	MMC**L**PE***IRY***I***AGK***MV***N***VPLE	<2%
P11-2	***L***AL***IRY***L***AGK***LL***N***ALFALL	76%
P11-2mL	****K***AL***IRY***L***AGK***LL***N***ALFALL	72%

[a] The mutated amino acids are in boldface italic type, and their original amino acids are in boldface type.

truncated forms of GCNF lacking H11 retained the ability to bind DNA through H3-mediated dimerization. The computational identification of functional sites within both H3 and H11 of GCNF and the experimental verification that both of these surfaces are indeed involved in dimerization and oligomerization lead us to believe that upon deletion of H11, H3-H3 becomes the prevalent, yet not necessarily physiologically relevant, interaction. In our experimental model, H11 and H3 were intact and minimally disrupted by single point mutations targeted by the ET analysis to the residues found to be the most important during evolution. The minimal nature of the disruptions caused by the point mutations was proven when the GCNF antibody could cross-link the mutant GCNFs and restore DNA binding. Different amino acid substitution at the same site produced the same effects on the repression function, binding activity, and proteinase resistance. Furthermore, the key functional determinants identified by the ET were designed into small soluble peptides that proved sufficient to competitively and specifically inhibit the protein-protein interaction. With both approaches, we show that a specific evolutionarily selected interaction exists between H3 and H11 and that such interaction is necessary for GCNF repression of the *Oct4* promoter.

Evolution-directed Analysis of Protein-Protein Interaction—From a technical perspective, we note that the evolutionary trace efficiently identified key protein-protein interfacial residues that enabled rational mutational targeting, and in a first for ET, rational design of peptide inhibitors of a protein-protein interaction. These trace residues allowed us to uncover and characterize a new protein-protein interaction interface. In fact, they were shown to be specifically necessary and sufficient molecular determinants of asymmetric GCNF oligomerization. The scalability of ET (33, 35) suggests that many other protein interfaces will lend themselves to the design of such peptides. Thus, the combined extraction of trace residue determinants of binding from the context of the whole protein, and the generation of peptide mimetics, represent a powerful tool for the targeted study and disruption of protein-protein interactions that take place in cellular networks, such as we show here in transcriptional control of gene expression.

Oligomerization of Endogenous GCNF—GCNF oligomerization and TRIF complex formation represent a new mechanism of regulation of transcriptional repression not previously described within the nuclear receptor superfamily. Although homodimers and heterodimers are the major DNA binding configurations for nuclear receptors, tetramerization has been detected. Human RXRα forms homodimers or heterodimers in the presence of ligand, but it can also exist as a tetramer in solution in the absence of ligand (60–63). According to structural analyses, the

RXRα tetramer is mediated through two symmetrical interactions of H3/H3 and H11/H11, and it represents an inactive or autorepressive form, which dissociates into RXRα homodimers upon 9-*cis*-retinoic acid treatment (61).

For GCNF, the functional consequence of oligomerization in differentiated P19 cells has yet to be determined. However, the GCNF oligomer is not the inactive form, as in the case of the RXRα tetramer, because the transiently transfected GCNF in differentiated P19 cells also incorporates into the TRIF complex causing repression of the *Oct4* reporter gene. It is possible that oligomerization of GCNF recruits different co-repressors or partners simultaneously to form a large repression complex. We have tried to understand the nature of the TRIF complex based on the interaction of GCNF with other co-regulators, however, although several GCNF interacting factors have been identified using yeast two-hybrid screens, including the co-repressors NCoR and SMRT, we failed to detect their presence in the TRIF complex (data not shown). We have used fast protein liquid chromatography to partially purify the TRIF complex and found that the estimated molecular weight of the TRIF complex from gel filtration columns was much larger than that detected by native gel electrophoresis. Thus, the endogenous GCNF in P19 cells may interact with other partners to form a large repression complex that simultaneously brings them to a promoter for rapid and coordinate repression of gene expression. Another possible function of GCNF oligomers is that the GCNF complex may bind to more than one DR0 element spatially distributed in a promoter, permitting efficient sequestration from transcriptional activators (52).

The mechanics of how the switch between the GCNF homodimer and the TRIF complex is flipped are still unknown. A conformational change induced by an endogenous ligand, interaction partner, and/or post-translational modification may change the conformation of helix 3 and/or helix 11, regulating how they interact with each other and form the GCNF oligomer in differentiated P19 cells.

A New Mechanism of NR Regulation—The asymmetric oligomerization interface between H3 and H11 may represent a more general mode of NR regulation not restricted to GCNF, since the ET analysis is based on the evolution of the whole protein family and points to the widespread importance of both H3 and H11 among all nuclear receptors. Thus, a similar association mechanism may be at play in other receptors. Currently, helix 11 is known to be widely important for homo- and heterodimerization of NRs. For example, deletion of helix 10/11 results in the loss of homodimerization of ERRγ (11). Mutation of Leu^{331} of HNF4α, similar to Leu^{459} in GCNF, results in the loss of dimerization (13). Mutation of Leu^{562} of Nurr1, corresponding to Leu^{459} in GCNF, also causes disruption of heterodimerization of Nurr1 with RXRα. Here, however, we show with multiple computational and experimental approaches that helix 3 is also involved in GCNF dimerization. H3 is widely known to be directly involved in co-activator binding (64) in many other nuclear receptors; however, we cannot exclude the possibility that the same interface is also involved in different interactions, as in the case of GCNF.

In conclusion, homodimerization and TRIF complex formation of GCNF are DNA-dependent. Evolutionary trace analysis identified two surfaces on opposing sides of GCNF that have dimerization features. Point mutation analyses demonstrated that a novel intermolecular interaction between helix 3 and helix 11 of GCNF not only mediates the formation of homodimers of recombinant GCNF but also mediates the formation of the endogenous GCNF TRIF complex. The heterotypic interaction permits oligomerization, and indeed the endogenous GCNF expressed in the differentiated P19 cells forms a GCNF hexamer. Our findings reveal unusual DNA binding, dimerization, and oligomerization properties for GCNF and suggest a novel mechanism by which GCNF acts as a transcriptional repressor that has implications for other members of the NR superfamily.

Acknowledgments—We thank Drs. Bin He and Jiemin Wong for constructive suggestions.

REFERENCES

1. Robinson-Rechavi, M., Escriva Garcia, H., and Laudet, V. (2003) *J. Cell Sci.* **116,** 585–586
2. Zhang, Z., Burch, P. E., Cooney, A. J., Lanz, R. B., Pereira, F. A., Wu, J., Gibbs, R. A., Weinstock, G., and Wheeler, D. A. (2004) *Genome Res.* **14,** 580–590
3. Gronemeyer, H., Gustafsson, J. A., and Laudet, V. (2004) *Nat. Rev. Drug Discov.* **3,** 950–964
4. Greschik, H., and Moras, D. (2003) *Curr. Top. Med. Chem.* **3,** 1573–1599
5. Aranda, A., and Pascual, A. (2001) *Physiol. Rev.* **81,** 1269–1304
6. Whitfield, G. K., Jurutka, P. W., Haussler, C. A., and Haussler, M. R. (1999) *J. Cell. Biochem. Suppl.* **32**, 110–122
7. Kumar, R., and Thompson, E. B. (1999) *Steroids* **64,** 310–319
8. Ribeiro, R. C., Kushner, P. J., and Baxter, J. D. (1995) *Annu. Rev. Med.* **46,** 443–453
9. Khorasanizadeh, S., and Rastinejad, F. (2001) *Trends Biochem. Sci.* **26,** 384–390
10. Moras, D., and Gronemeyer, H. (1998) *Curr. Opin. Cell Biol.* **10,** 384–391
11. Hentschke, M., Susens, U., and Borgmeyer, U. (2002) *Eur. J. Biochem.* **269,** 4086–4097
12. Aarnisalo, P., Kim, C. H., Lee, J. W., and Perlmann, T. (2002) *J. Biol. Chem.* **277,** 35118–35123
13. Bogan, A. A., Dallas-Yang, Q., Ruse, M. D., Jr., Maeda, Y., Jiang, G., Nepomuceno, L., Scanlan, T. S., Cohen, F. E., and Sladek, F. M. (2000) *J. Mol. Biol.* **302,** 831–851
14. Dhe-Paganon, S., Duda, K., Iwamoto, M., Chi, Y. I., and Shoelson, S. E. (2002) *J. Biol. Chem.* **277,** 37973–37976
15. Seol, W., Chung, M., and Moore, D. D. (1997) *Mol. Cell. Biol.* **17,** 7126–7131
16. Bourguet, W., Ruff, M., Chambon, P., Gronemeyer, H., and Moras, D. (2000) *Mol. Cell* **5,** 289–298
17. Bourguet, W., Ruff, M., Chambon, P., Gronemeyer, H., and Moras, D. (1995) *Nature* **375,** 690–697
18. Gampe, R. T., Jr., Montana, V. G., Lambert, M. H., Miller, A. B., Bledsoe, R. K., Milburn, M. V., Kliewer, S. A., Willson, T. M., and Xu, H. E. (2000) *Mol. Cell* **5,** 545–555
19. Bledsoe, R. K., Montana, V. G., Stanley, T. B., Delves, C. J., Apolito, C. J., McKee, D. D., Consler, T. G., Parks, D. J., Stewart, E. L., Willson, T. M., Lambert, M. H., Moore, J. T., Pearce, K. H., and Xu, H. E. (2002) *Cell* **110,** 93–105
20. Greschik, H., Wurtz, J. M., Hublitz, P., Kohler, F., Moras, D., and Schule, R. (1999) *Mol. Cell Biol.* **19,** 690–703
21. Chen, F., Cooney, A. J., Wang, Y., Law, S. W., and O'Malley, B. W. (1994) *Mol. Endocrinol.* **8,** 1434–1444
22. Susens, U., Aguiluz, J. B., Evans, R. M., and Borgmeyer, U. (1997) *Dev. Neurosci.* **19,** 410–420
23. Heinzer, C., Susens, U., Schmitz, T. P., and Borgmeyer, U. (1998) *Biol. Chem.* **379,** 349–359
24. Bauer, U. M., Schneider-Hirsch, S., Reinhardt, S., Benavente, R., and Maelicke, A. (1998) *FEBS Lett.* **439,** 208–214
25. Zhang, Y. L., Akmal, K. M., Tsuruta, J. K., Shang, Q., Hirose, T., Jetten, A. M., Kim, K. H., and O'Brien, D. A. (1998) *Mol. Reprod. Dev.* **50,** 93–102
26. Cooney, A. J., Hummelke, G. C., Herman, T., Chen, F., and Jackson, K. J. (1998) *Biochem. Biophys. Res. Commun.* **245,** 94–100
27. Yan, Z. H., Medvedev, A., Hirose, T., Gotoh, H., and Jetten, A. M. (1997) *J. Biol. Chem.* **272,** 10565–10572
28. Borgmeyer, U. (1997) *Eur. J. Biochem.* **244,** 120–127
29. Fuhrmann, G., Chung, A. C., Jackson, K. J., Hummelke, G., Baniahmad, A., Sutter, J., Sylvester, I., Scholer, H. R., and Cooney, A. J. (2001) *Dev. Cell* **1,** 377–387
30. Lan, Z. J., Gu, P., Xu, X., Jackson, K. J., DeMayo, F. J., O'Malley, B. W., and Cooney, A. J. (2003) *EMBO J.* **22,** 4070–4081
31. Schmitz, T. P., Susens, U., and Borgmeyer, U. (1999) *Biochim. Biophys. Acta* **1446,** 173–180
32. Lichtarge, O., Bourne, H. R., and Cohen, F. E. (1996) *J. Mol. Biol.* **257,** 342–358
33. Mihalek, I., Res, I., and Lichtarge, O. (2004) *J. Mol. Biol.* **336,** 1265–1282
34. Madabushi, S., Yao, H., Marsh, M., Kristensen, D. M., Philippi, A., Sowa, M. E., and Lichtarge, O. (2002) *J. Mol. Biol.* **316,** 139–154
35. Yao, H., Kristensen, D. M., Mihalek, I., Sowa, M. E., Shaw, C., Kimmel, M., Kavraki, L., and Lichtarge, O. (2003) *J. Mol. Biol.* **326,** 255–261
36. Lichtarge, O., Yamamoto, K. R., and Cohen, F. E. (1997) *J. Mol. Biol.* **274,** 325–337
37. Madabushi, S., Gross, A. K., Philippi, A., Meng, E. C., Wensel, T. G., and Lichtarge, O. (2004) *J. Biol. Chem.* **279,** 8126–8132
38. Quan, X. J., Denayer, T., Yan, J., Jafar-Nejad, H., Philippi, A., Lichtarge, O., Vleminckx, K., and Hassan, B. A. (2004) *Development* **131,** 1679–1689
39. Altschul, H., Madden, T., Schaffer, A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. (1997) *Nucleic Acids Res.* **25,** 3389–3402
40. Thompson, J., Higgins, D., and Gibson, T. (1994) *Nucleic. Acids Res.* **22,** 4673–4680
41. Egea, P. F., Mitschler, A., Rochel, N., Ruff, M., Chambon, P., and Moras, D. (2000) *EMBO J.* **19,** 2592–2601

42. Bourguet, W., Ruff, M., Chambon, P., Gronemeyer, H., and Moras, D. (1995) *Nature* **375,** 377–382
43. Munoz, V., and Serrano, L. (1997) *Biopolymers* **41,** 495–509
44. Cooney, A. J., Tsai, S. Y., O'Malley, B. W., and Tsai, M. J. (1992) *Mol. Cell. Biol.* **12,** 4153–4163
45. Hummelke, G. C., Meistrich, M. L., and Cooney, A. J. (1998) *Mol. Reprod. Dev.* **50,** 396–405
46. Burbach, J. P., Lopes da Silva, S., Cox, J. J., Adan, R. A., Cooney, A. J., Tsai, M. J., and Tsai, S. Y. (1994) *J. Biol. Chem.* **269,** 15046–15053
47. Barnea, E., and Bergman, Y. (2000) *J. Biol. Chem.* **275,** 6608–6619
48. Fuhrmann, G., Sylvester, I., and Scholer, H. R. (1999) *Cell Mol. Biol.* (*Noisy-Le-Grand*) **45,** 717–724
49. Butler, A. J., and Parker, M. G. (1995) *Nucleic Acids Res.* **23,** 4143–4150
50. Ikeda, Y. (1996) *Acta Paediatr. Jpn.* **38,** 412–419
51. Yan, Z., and Jetten, A. M. (2000) *J. Biol. Chem.* **275,** 35077–35085
52. Yasmin, R., Yeung, K. T., Chung, R. H., Gaczynska, M. E., Osmulski, P. A., and Noy, N. (2004) *J. Mol. Biol.* **343,** 327–338
53. Mangelsdorf, D. J., and Evans, R. M. (1995) *Cell* **83,** 841–850
54. Lee, W. Y., and Noy, N. (2002) *Biochemistry* **41,** 2500–2508
55. Suino, K., Peng, L., Reynolds, R., Li, Y., Cha, J. Y., Repa, J. J., Kliewer, S. A., and Xu, H. E. (2004) *Mol. Cell* **16,** 893–905
56. Shan, L., Vincent, J., Brunzelle, J. S., Dussault, I., Lin, M., Ianculescu, I., Sherman, M. A., Forman, B. M., and Fernandez, E. J. (2004) *Mol. Cell* **16,** 907–917
57. 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., and Moore, J. T. (2004) *Mol. Cell* **16,** 919–928
58. Allan, G. F., Leng, X., Tsai, S. Y., Weigel, N. L., Edwards, D. P., Tsai, M. J., and O'Malley, B. W. (1992) *J. Biol. Chem.* **267,** 19513–19520
59. Zeng, Z., Allan, G. F., Thaller, C., Cooney, A. J., Tsai, S. Y., O'Malley, B. W., and Tsai, M. J. (1994) *Endocrinology* **135,** 248–252
60. Kersten, S., Kelleher, D., Chambon, P., Gronemeyer, H., and Noy, N. (1995) *Proc. Natl. Acad. Sci. U. S. A.* **92,** 8645–8649
61. Kersten, S., Reczek, P. R., and Noy, N. (1997) *J. Biol. Chem.* **272,** 29759–29768
62. Chen, Z. P., Iyer, J., Bourguet, W., Held, P., Mioskowski, C., Lebeau, L., Noy, N., Chambon, P., and Gronemeyer, H. (1998) *J. Mol. Biol.* **275,** 55–65
63. Gampe, R. T., Jr., Montana, V. G., Lambert, M. H., Wisely, G. B., Milburn, M. V., and Xu, H. E. (2000) *Genes Dev.* **14,** 2229–2241
64. Darimont, B. D., Wagner, R. L., Apriletti, J. W., Stallcup, M. R., Kushner, P. J., Baxter, J. D., Fletterick, R. J., and Yamamoto, K. R. (1998) *Genes Dev.* **12,** 3343–3356