Supporting Information

Thomas-Chollier et al. 10.1073/pnas.1316235110

SI Materials and Methods

Plasmids. Reporter plasmids with genomic fragments (approximately 500 bp centered on the summit of the ChIP-seq peak) derived from genomic regions found near the TSS of isoform-specific genes were amplified by PCR and subcloned into pGL3-promoter (Promega). Genomic coordinates (hg19) for these genomic regions: *KLK3*: chr19: 51,357,476–51,357,948; *ITPRIP*: chr10: 106,099,880–106,100,350; and *JPH2*: chr20: 42,815,861–42,816,556. Candidate glucocorticoid receptor (GR) binding sequence (GBS) of the *ITPRIP* reporter (AGAACAggcAGTTAC) was changed to AGAAAAggcCTTTAC to delete the site and to AGAACAgggTGTTCT to mutate the GBS into one resembling the non-differential consensus sequence by site-directed mutagenesis. Other luciferase reporters used have been described (1). Site-directed mutagenesis was used to introduce the extra arginine of GRγ into the lever arm using pcDNA3-based GRα expression constructs with mutations in AF1 (E219K/F220L/W234R), dimerization region (A477T), or AF2 (E773R) (2).

Transient Transfections and Reporter Analysis. U2OS cells were transiently transfected, treated overnight with 100 nM or 1 μM dexamethasone (dex), and harvested, and luciferase activity was measured as described (1).

Cell Lines. Clonal lines derived from parental U2OS cells expressing either GRα or GRγ from rat and clonal lines harboring GRα with point mutations in activation function 1 (AF1) (E219K/F220L/W234R), dimerization region (A477T), and activation function 2 (AF2) (E773R) have been described previously (1, 2). To generate cell lines stably expressing GRγ with point mutations in AF1, dimerization region, or AF2, parental U2OS cells were transfected with the appropriate expression construct. The next day, cells were transferred to 15-cm dishes and resistant clones were selected at 750 μg/mL G418 (Invitrogen). Clonal lines expressing GR mutants at levels similar to those for wild-type GRγ were further analyzed. U2OS cells and derived clonal lines were maintained and propagated as described previously (1).

Immunoblotting. Total protein from equal amounts of cells was separated with SDS/PAGE gels, transferred to membranes, and incubated with either anti-GR (N499) or anti-actin (Sc-1616R; Santa Cruz Biotechnology) antibodies followed by incubation with secondary antibodies conjugated with horseradish peroxidase. Proteins were visualized using an ECL detection system (Amersham Biosciences).

Quantitative Real-Time PCR. RNA isolation, reverse transcription, quantitative (q)PCR, and data analysis were performed as described previously (1). Primer pairs used as described (1) are listed in Table S1.

Microarrays. The arrays were whole genome spotted oligo nucleotide arrays (HEEBO probe set; Invitrogen) printed at University of California San Francisco (UCSF). RNA was isolated from confluent cells that were treated with 1 μM dexamethasone or EtOH for 3 h. For each condition, we analyzed three biological replicates. Microarrays were hybridized in a MAUI hybridization chamber (BioMicro Systems), scanned on a GenePix 4000B scanner (Molecular Devices), and gridded with SpotReader (Niles Scientific). The array data were analyzed with LIMMA package in BioConductor (3). In brief, we applied the normalizeWithinArrays function with the robust multichip average background correction method. For each experiment, we required that a given probe be detected as present by SpotReader in at least two of the three replicate arrays to be included in further analysis. We used lmFit and eBayes with the default parameters, including the Benjamini and Hochberg method to adjust the *P* value for multiple hypothesis testing. To call the regulated genes, we used as cutoff 0.05 for the adjusted *P* value (corresponding to a false discovery rate of 5%) and 1.5 for the fold change value. The annotations of the probes were updated with Ensembl version 61 (assembly GRCh37/hg19).

Electrophoretic Mobility Shift Assays. Electrophoretic mobility shift assays (EMSAs) were performed as described (1) with the following exceptions: poly(deoxyinosinic-deoxycytidylic) concentration was 50 instead of 100 ng/μL and gels were scanned with a Fuji FLA-5100R scanner to quantify free versus total DNA. K_Ds were determined as described (1). Sequence of the 5′ Cy-5 end-labeled KIAA GBS was as follows: TCGACGGACAAAATGTTCTGTAC.

ChIP and ChIP-seq. ChIP assays were essentially done as described (4) with the following exceptions for samples that were subsequently analyzed by deep sequencing: Approximately 10 million cells were treated with 0.01% ethanol vehicle or 1 μM dexamethasone for 1.5 h. Chromatin was sheared with a Bioruptor water bath sonicator (Diagenode) to produce fragments of ∼100–200 bp. Protein-G–coupled magnetic beads (Active-Motif) were preincubated for 1 h with GR-antibody (N499) before chromatin was added and incubated for an additional 2–4 h while rotating at 4 °C. Subsequently beads were washed three times with 10 mM Tris·HCl pH 8.0, 1 mM EDTA, 500 mM NaCl, 5% (vol/vol) glycerol, 0.1% sodium deoxycholate, 0.1% SDS, 1% Triton X-100, 0.5 mg/μL BSA, followed by three additional washes with 20 mM Tris, pH 8.0, 1 mM EDTA, 250 mM LiCl, 0.5% Nonidet P-40, and 0.5% sodium deoxycholate.

ChIP-seq libraries were prepared from 10 ng of ChIP DNA as described (5).

NMR. ***Protein expression and purification.*** Expression and purification of rat GRα and GRγ DNA binding domain, residues 440–525, (rGR-DBD) was performed as described previously (1) with the following exceptions: BL21 Gold cells (Stratagene) were grown in 50 mL LB media to optical density of ∼0.6, then pelleted and resuspended in 1 L minimal media containing 2 g/L $^{15}NHCl_4$ as the only nitrogen source. Expression was induced at optical density ∼0.7 with 0.5 mM IPTG for 16 h at 20 °C. In the final purification step, protein was eluted from a 16/60 Superdex75 gel filtration column in NMR buffer (20 mM sodium phosphate pH 6.7, 100 mM NaCl, and 1 mM DTT).

Protein–DNA complex formation. Single-stranded GBS oligos (IDT) were purified by MonoQ Column (GE Healthcare) equilibrated with 10 mM NaOH, 100 mM NaCl, and eluted by linear gradient reaching ∼600 mM NaCl. Purified oligos were dialyzed into H_2O, lyophilized, and resuspended at ∼2 mM in H_2O. Complementary oligos were annealed in 20 mM sodium phosphate, 100 mM NaCl, 5 mM $MgCl_2$ in a boiling waterbath, and cooled slowly to room temperature. DNA duplexes were diluted in NMR buffer, combined with rGR-DBD at a ∼35% excess DNA to GR-DBD dimer to ensure that all protein was bound, and concentrated slowly at 4 °C using a 3K MWCO Centrifugal Filter (Amicon) to 300 μM complex. Concentrated complexes were

filtered to remove any precipitate using Ultrafree PVDF 0.22-μm columns (Millipore).

FKBP5: gtacAGAACAgggTGTTCTtcgac,

ITPRIP: gtacAGAACAggcAGTTACtcgac.

NMR and chemical shift perturbation analysis. ^{1}H-^{15}NHSQC spectra were acquired on a Bruker 800-MHz spectrometer at 35 °C. Data were processed in NMRPipe (National Institutes of Health) and analyzed in Sparky (UCSF). Peak assignments for the GRα-DBD:DNA complex (6) were transferred to the GRα-DBD:*FKBP5* and GRα-DBD:*ITPRIP* complexes. Chemical shift difference between GRα and GRγ complexes was determined by measuring the distance from each assigned GRα peak to the nearest peak in the GRγ spectrum using the formula: combined chemical shift $\Delta\delta = [(\Delta H\ ppm)^2 + (\Delta N\ ppm/5)^2]^{-2}$.

Computational Analysis. ***Microarray analysis.*** To identify the isoform-specific regulated genes from the microarray data, we first calculated the principal component of the $\log^2$ fold change of GRα versus GRγ probes (represented by a gray line in Fig. 1*B*). The isoform-specific probes were identified as the points that deviate of at least 2 SDs in the direction of the second principal component. The probes showing opposite regulation (e.g., up-regulated for GRα and down-regulated for GRγ) were assigned to the isoform showing the highest fold change absolute value. The probes were finally linked to their associated genes, thereby constituting the list of isoform-specific genes.

ChIP-seq analysis. The obtained sequence reads were 30 bp and 34 bp in length for the GRα and GRγ, respectively. The reads were mapped with Bowtie parameter setting (-v 2 -m 1) (7) on the human GRCh37/hg19 assembly. The peak calling step was performed with model-based analysis for ChIP-seq (MACS) 1.4 (–bw 300,–mfold 5,30 –pvalue 1e-5) (8) using as control the input DNA from the same U2OS cell line. MACS was also run on the input DNA alone; the resulting peaks serve as filters to remove artifactual peaks. A stringent cutoff of false discovery rate 0.2 was applied before processing the peak list with PeakSplitter (9) to subdivide the peak regions into individual enriched regions. To identify the isoform-differential peaks, the number of non-identical reads was calculated for each peak region in the GRα and GRγ datasets. The two linear regression lines were calculated and the points deviating of at least 2 SDs from these lines were considered as isoform specific (Fig. S1).

Motif discovery and clustering. The DNA binding motifs of GR were extracted from the ChIP-seq peaks with de novo motif discovery approaches embedded in the program peak motifs from RSAT (10). In brief, this program searches for globally overrepresented motifs (or spaced motifs) in the peak dataset, for local over-representation in various sizes of windows and for motifs with global positional bias. The peak sequences were therefore centered on their summit before analysis to ensure that these last two algorithms perform well. Motif discovery was conducted on three different peak datasets (nondifferential and GRα-specific and GRγ-specific binding regions). The nonspecific dataset was created by randomly selecting 1,000 peaks among all nonspecific peaks. Each analysis was performed with and without repeat masking, and both on the complete peak length or ±50 bp around the peak summit.

All resulting motifs were manually filtered to extract the GR motifs. Next, these motifs (in the form of count matrices) were grouped by dataset (GRα specific, GRγ specific, and non-specific). To obtain a condense view of the found motifs in each dataset, we clustered these motifs using STAMP (11). One familial profile was thereby obtained for each dataset, allowing motif comparison.

Correlating genomic binding and regulation. The regulated genes were separated into three datasets (nondifferential, GRα-specifically regulated, and GRγ-specifically regulated), and the coordinates of the region ±10 kb around the TSS were retrieved. The overlap between these coordinates and the peak coordinates was calculated using intersectBed from the BEDTools suite (12), and the number of genes showing such overlap was calculated and used to produce Fig. 2*A*. The overlap does not reflect the number of peaks found in the 20-kb region; rather, the following rules were used: a gene is assigned to GRα-specific binding if there is at least one GRα-specific peak in the region. Conversely, GRγ-specific binding is assigned if there is at least one GRγ-specific peak found. The case with both GRα-specific and GRγ-specific peaks was never encountered. If only nondifferential peaks are found, the gene is assigned to nondifferential binding.

1. Meijsing SH, et al. (2009) DNA binding site sequence directs glucocorticoid receptor structure and activity. *Science* 324(5925):407–410.
2. Rogatsky I, et al. (2003) Target-specific utilization of transcriptional regulatory surfaces by the glucocorticoid receptor. *Proc Natl Acad Sci USA* 100(24):13845–13850.
3. Smyth GK, Yang YH, Speed T (2003) Statistical issues in cDNA microarray data analysis. *Methods Mol Biol* 224:111–136.
4. So AY, Chaivorapol C, Bolton EC, Li H, Yamamoto KR (2007) Determinants of cell- and gene-specific transcriptional regulation by the glucocorticoid receptor. *PLoS Genet* 3(6):e94.
5. Holdorf MM, Cooper SB, Yamamoto KR, Miranda JJ (2011) Occupancy of chromatin organizers in the Epstein-Barr virus genome. *Virology* 415(1):1–5.
6. Watson LC, et al. (2013) The glucocorticoid receptor dimer interface allosterically transmits sequence-specific DNA signals. *Nat Struct Mol Biol* 20(7):876–883.
7. Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 10(3):R25.
8. Zhang Y, et al. (2008) Model-based analysis of ChIP-Seq (MACS). *Genome Biol* 9(9): R137.
9. Salmon-Divon M, Dvinge H, Tammoja K, Bertone P (2010) PeakAnalyzer: Genome-wide annotation of chromatin binding and modification loci. *BMC Bioinformatics* 11:415.
10. Thomas-Chollier M, et al. (2011). RSAT 2011: Regulatory sequence analysis tools. *Nucleic Acids Res* 39(Web Server issue):W86–W91.
11. Mahony S, Benos PV (2007) STAMP: A web tool for exploring DNA-binding motif similarities. *Nucleic Acids Res* 35(Web Server issue):W253–W258.
12. Quinlan AR, Hall IM (2010) BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* 26(6):841–842.

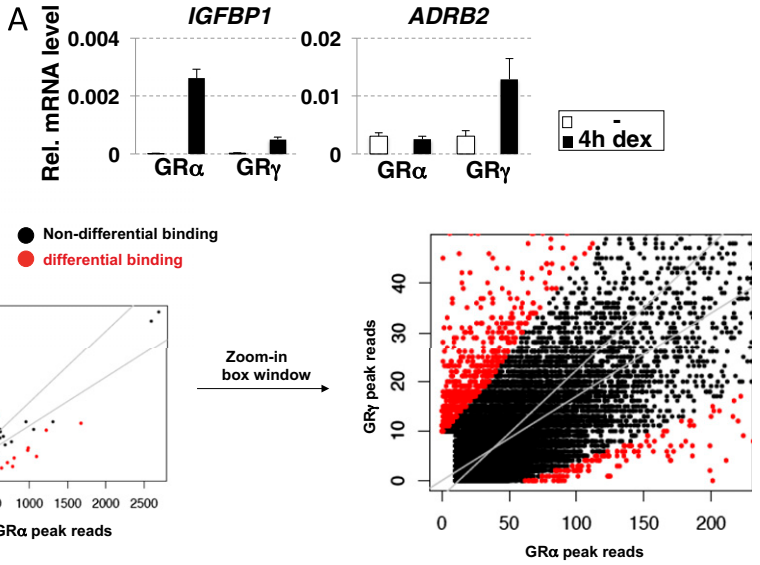
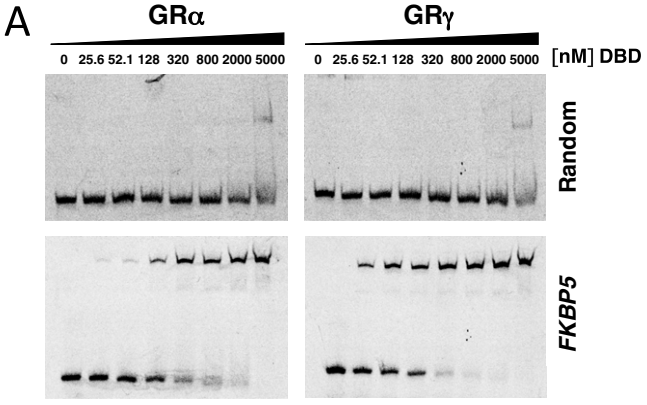


Fig. S1. Comparison of gene regulation and genomic binding by GRα and GRγ. (*A*) Examples of isoform-specific target genes. Relative transcript levels of treated (4 h, 100 nM dex) or untreated cells expressing GRα or GRγ were analyzed by real-time quantitative PCR. Averages ± SEM are shown ($n = 3$) for *IGFBP1* (*Left*) and for *ADRB2* (*Right*). (*B*) Quantitative comparison of genomic binding for GRα and GRγ. Reads mapped to binding regions are plotted for GRα and GRγ showing an overview (*Left*) or a zoom-in of the box window (*Right*). To identify the isoform-specific peaks, the number of nonidentical reads was calculated for each peak region in the GRα and GRγ datasets. The two linear regression lines were calculated and the points deviating ≥2 SDs from these lines were considered as isoform specific.



B

GBS	Sequence	K_D (nM) GRα	K_D (nM) GRγ
CGT	AGAACAtttTGTACG	91±37	47±23
FKBP5	AGAACAgggTGTTCT	224±79	22±4
GILZ	AGAACAttgGGTTCC	468±116	207±140
Pal	AGAACAaaaTGTTCT	23±3	15±5
SGK	AGAACAtttTGTCCG	342±97	73±17
ITPRIP	AGAACAggcAGTTAC	4347±1685	1860±150
Random	-	>5000	>5000

Fig. S2. Specific and nonspecific DNA binding by GRα and GRγ. (*A*) Electrophoretic mobility shift assays showing binding of the DBD of (*Left*) GRα and (*Right*) GRγ (human-GR 380–540) to (*Lower Panels*) the *FKBP5* GBS (AGAACAgggTGTTCT) or to (*Upper Panels*) randomized sequences. (*B*) Binding affinities for GRα and GRγ (human GR DBD 380–540) for each GBS ± SEM are shown ($n \geq 3$). GBS derived from GRα-specific gene is in red, GBS from GRγ-specific gene, in blue.

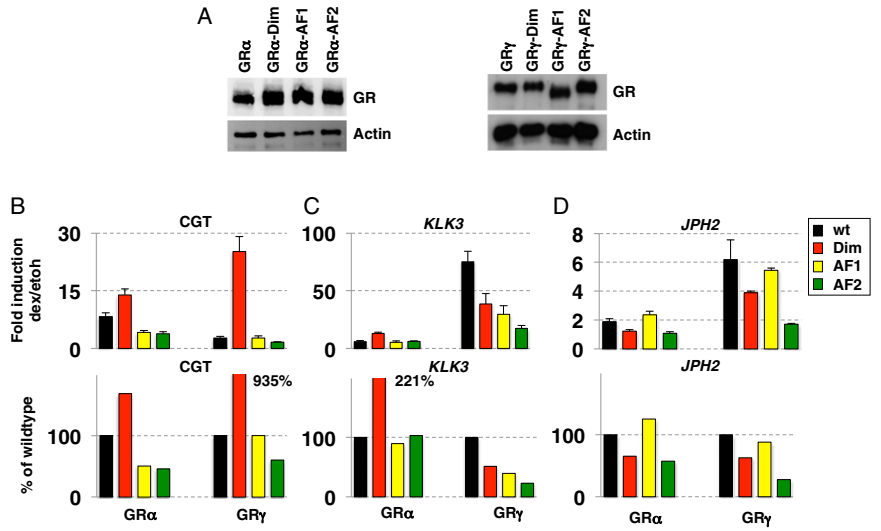


Fig. S3. GR isoform-specific patterns of domain requirements. (*A*) Protein expression levels of U2OS cells stably expressing (*Left*) GRα or (*Right*) GRγ with mutations in Dim, AF1, and AF2. Total cell extracts were analyzed by Western blotting. (*B–D*) The effect of point mutations in Dim, AF1, and AF2 in GRα and GRγ was tested in U2OS cells transiently cotransfected with GR expression vectors as indicated and luciferase reporters. The reporters carried either (*B*) a single GBS (CGT) or (*C* and *D*) a 500-bp GBR fragment (*KLK3* and *JPH2*). Fold induction by treatment with dex ± SEM (*Upper*) and percentage of activity relative to the wild-type isoform as indicated on the *x*-axis (*Lower*) are shown ($n \geq 3$).

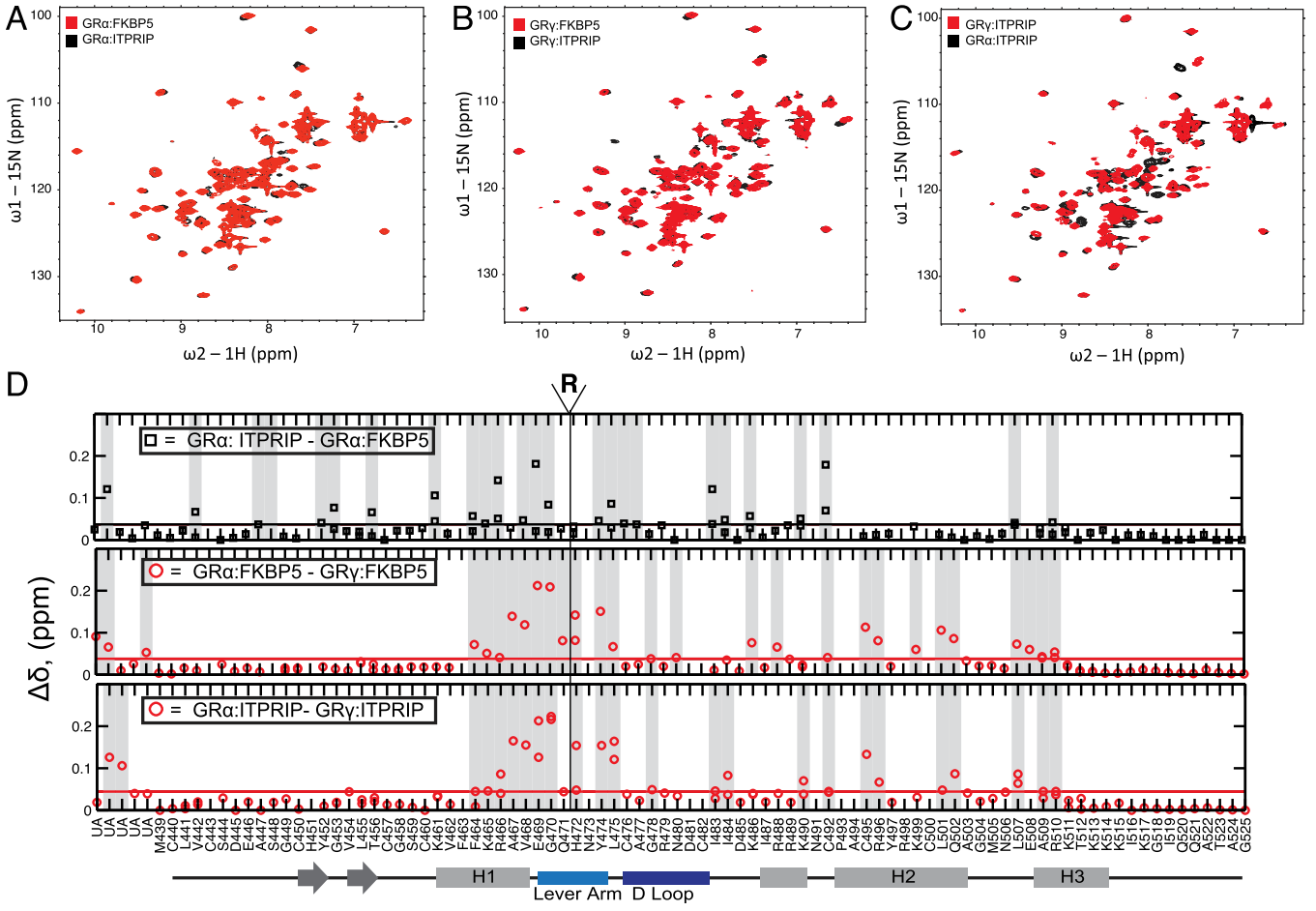


Fig. S4. Comparison of ^{1}H-^{15}N HSQC for GRα-DBD and GRγ-DBD bound to *FKBP5* and *ITPRIP*. (*A*) GRα bound to *FKBP5* (red) versus *ITPRIP* (black). (*B*) GRγ bound to *FKBP5* (red) versus *ITPRIP* (black). (*C*) GRγ (red) and GRα (black) bound to *ITPRIP*. (*D*) Chemical shift difference analysis comparing ^{1}H-^{15}N HSQC spectra of GRγ bound to *ITPRIP* versus *FKBP5* (*Top*), GRα and GRγ bound to *FKBP5* (*Middle*), and GRα and GRγ bound to *ITPRIP* (*Bottom*) for all assigned GRα residues (*x*-axis). Residues with split peaks in the spectra are plotted as two data points for the same residue. Gray bars highlight residues with chemical shift greater than the mean chemical shift of all peaks in the comparison (black or red line). UA, unassigned peaks. R on top indicates site of arginine insertion for GRγ.

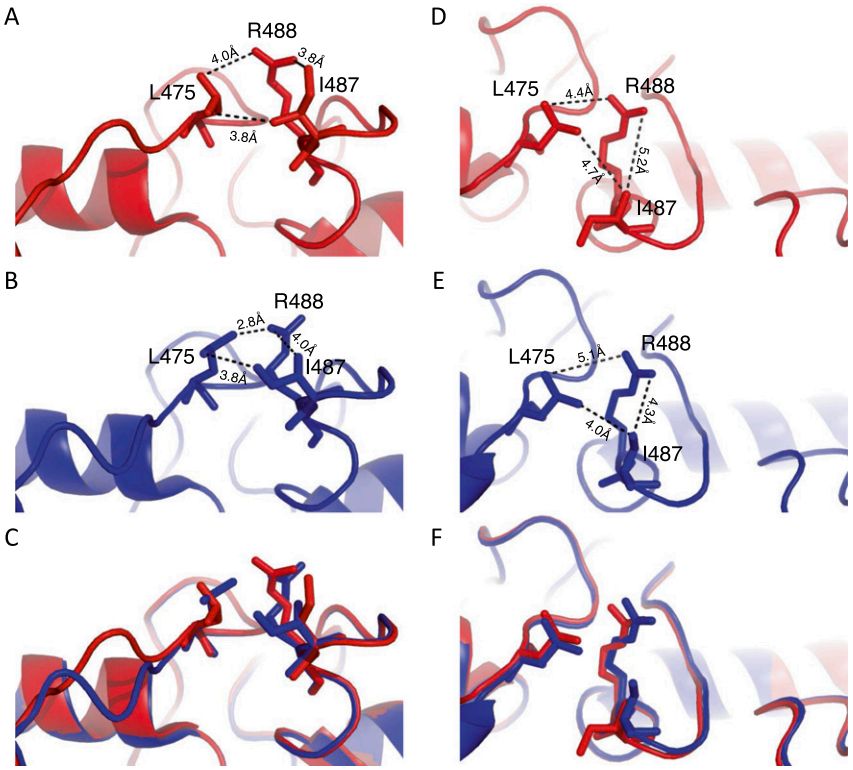


Fig. S5. GRγ lever arm insertion results in conformational changes in the dimerization interface. (*A–C*) Side chains positioning and distances between L475 of chain B and R488 and I487 of chain A for (*A*) GRα (red, PDB ID 3G6U), (*B*) GRγ (blue, PDB ID 3G6T), and (*C*) overlay of GRα and GRγ. (*D–F*) Side chain contacts and distances between L475 of chain A and R488 and I487 of chain B for (*D*) GRα, (*E*) GRγ, and (*F*) overlay of GRα and GRγ.

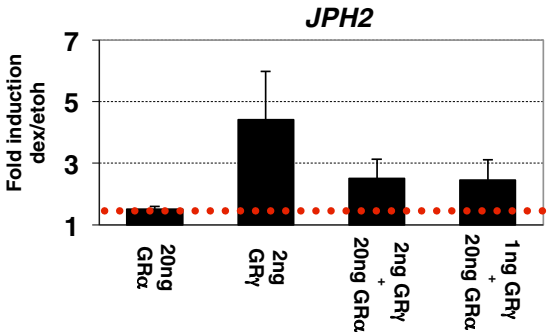


Fig. S6. Excess of GRα reduced but did not abolish GRγ transcriptional activity. U2OS cells were transfected with the GRγ-specific *JPH2*-luciferase reporter construct together with GRα and GRγ expression constructs as indicated. Fold induction upon dex treatment ± SEM is shown ($n = 3$). Dotted red line indicates induction level observed for GRα in the absence of GRγ.

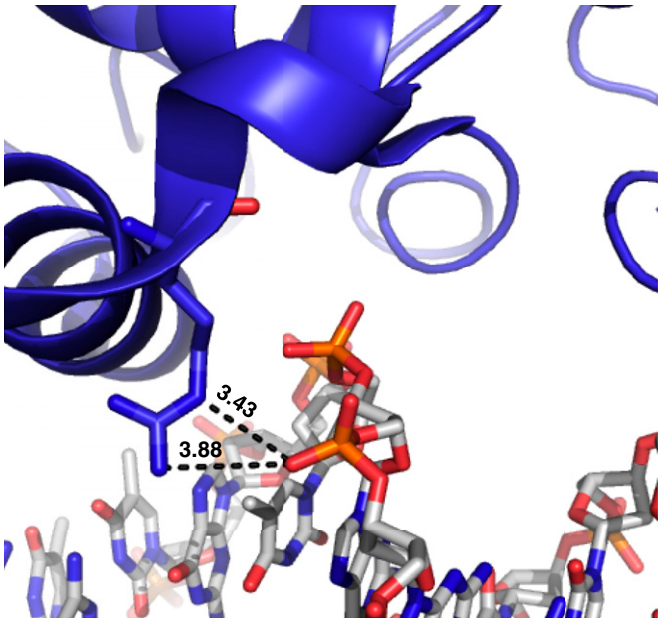


Fig. S7. DNA backbone contacts of R471 of the lever arm of GRγ. Distances of epsilon amine and terminal amides of R471 of the GRγ:*FKBP5* complex (3G6T) relative to the phosphate–oxygen backbone of the DNA are shown.

Table S1. Primers used for qPCR analysis

Gene	Forward primer	Reverse primer
ADRB2	cttccattgatgtgctgtgc	atggcaaagtagcgatccac
IGFBP1	tcacagcagacagtgtgagac	agacccagggatcctcttc
ITPRIP	agtctggcccagttcagaga	cgctggagtgaggaactagg
PNLIP	gatgttggggacttgcagat	aacttcctccctgacggttt
SPINK5L3	ctgcctttccccacaagatt	ttgaggcacaaacaggtgct
*ADRB2*_GBR	aagtacttgacagcgagtgtgc	tcacccgagcactgataatg
*IGFBP1*_Intron	ccaggaggtgtttggaatgt	tggcaatgaatggaagtgaa
*ITPRIP*_GBR	aaccctcatgtctggattgg	ttccctgacttcttgactcctc

Other Supporting Information Files

Dataset S1 (XLSX)