Supporting Information

Yumoto et al. 10.1073/pnas.1117036108

SI Materials and Methods

Preparation of β-Catenin/Liver Receptor Homolog-1 (LRH-1) Ligand Binding Domain (LBD) Complex. Human full-length LRH-1 complementary DNA (cDNA) was provided by Holly Ingraham's laboratory at the University of California, San Francisco (UCSF). Human full-length β-catenin cDNA was obtained from Open Biosystems. A cDNA fragment encoding human liver receptor homolog-1 (hLRH-1) residues 191–541 corresponding to hLRH-1 LBD with the preceding hinge region, with an additional introduced N-terminal cleavage site for tobacco etch virus (TEV) protease, was cloned into pET32 Xa/ligation independent cloning (LIC) vector (Novagen). A cDNA fragment encoding β-catenin residues 138–663 (corresponding to armadillo repeat region), with introduced N-terminal cleavage site for TEV protease, was cloned into pCDF-2 Ek/LIC vector (Novagen) with an additional introduced cleavage site for TEV protease.The resultant recombinant pET32 Xa/LIC and pCDF-2 Ek/LIC plasmids were cotransformed into BL21Star (DE3) cells (Invitrogen). The *Escherichia coli* cells were grown in 1 L LB medium supplemented with 100 mg/mL ampicillin, and 50 mg/mL streptomycin in a 3 L flask at 37 °C. Expression of the thioredoxin-fusion hLRH-1 hinge-LBD and his-tagged β-catenin (138–663) were expressed with 50 μM (final concentration) isopropyl β-D-thiogalactopyranoside, and incubation continued for 20–24 hours at 16 °C. The cells were collected by centrifugation at 4,500 × *g* for 15 min at 4 °C. The pellet was frozen quickly by liquid nitrogen and stored at −30 °C until use. The cell pellet was resuspended in buffer A [20 mM Tris·HCl (pH 8.0), 1 mM CHAPS, 10% glycerol, 5 mM β-melcaptoethanol, 300 mM NaCl, 20 mM imidazole] with 1 tablet of protease inhibitor cocktail (Roche), and then disrupted by sonication. After centrifugation at 40,000 × *g* for 0.5 h at 4 °C, the supernatant was mixed with Ni-nitrilotriacetate beads (Qiagen), which is equilibrated with buffer A, and then incubated rotating gently for 1 h at 4 °C. The captured proteins were eluted with buffer A containing 400 mM imidazole. The sample was concentrated with Amicon concentrator (Millipore) to 1 mL, mixed with his-tagged TEV protease, and dialyzed against 500 mL buffer B [20 mM Tris·HCl (pH 8.0), 1 mM CHAPS, 10% glycerol, 5 mM DTT, 200 mM NaCl] overnight at 4 °C. The TEV-protease treated sample was diluted with buffer A to reduce DTT concentration below 1 mM, then incubated with Ni-beads, which was equilibrated with buffer A, for 1 h at 4 °C to eliminate the his-tagged proteins. The flow-through fractions of the Ni-beads column was collected, concentrated, and loaded onto gel filtration column Superdex 200 10/30 (GE Healthcare), which was equilibrated with buffer C [20 mM Tris·HCl (pH 8.0), 1 mM CHAPS, 10% glycerol, 10 mM DTT, 150 mM NaCl]. The eluted fractions were analyzed by SDS/PAGE (NuPage, Invitrogen) to assess the purity of the protein complex.

Crystallization of LRH-1/β-Catenin Complex. Prior to crystallization, the protein complex was concentrated to 15 mg/mL in buffer C. Hanging drop vapor diffusion crystallization trials were carried out using mosquito nanoliter drop setter (TTP LabTech) with Nextal PEGs suite (Qiagen). The complex crystallized in a condition, which has 20 mM Tris·HCl (pH 8.5), 40% PEG200 at 20 °C in 1 d. Drops made manually by mixing 1.0 μL protein complex solution and 1.0 μL reservoir solution were equilibrated against 70 μL reservoir solution in Intelli-Plates (Art Robbins Instruments) and incubated at 20 °C. Larger crystals appeared in 2 d.

Determination of Crystal Structure of LRH-1/β-Catenin Complex. X-ray diffraction data were measured at −80 °C to 2.8 Å at the BL 8.3.1 ($l = 1.11587$ Å) in Advanced Light Source (Lawrence Berkeley National Laboratory) using a single crystal. Data were integrated and scaled using program HKL-2000 (1). The crystal was of the space group $P2_1$ with cell dimensions of $a = 49.8$ Å, $b = 151.6$ Å, $c = 76.1$ Å, and $\beta = 96.96°$ and contained one LRH-1 LBD/β-catenin complex in the asymmetric unit.

The structure of the complex was determined by the molecular replacement method using the Phaser in Collaborative Computational Project No. 4 (CCP4) suite with starting models derived from the atomic coordinates for human β-catenin fragment (PDB ID: 2Z6H) and hLRH-1 LBD (PDB ID: 1YOK). Electron-density maps based on coefficients $2F_o$-F_c were calculated from the phases of the initial model. Subsequent rounds of model building and refinement were performed using programs COOT and Phenix.refine, respectively. The geometry restraint information was generated using phenix.elbow (2). At later stages of the refinement, the structure was checked using simulated annealing composite omit maps. The final model of the complex is refined to 2.8 Å with R_{free}/R_{work} values of 24.26/19.88. The stereochemical properties were assessed by PROCHECK (3). Superpositions of models were calculated using the program LSQKAB (4) in CCP4 suite. Figures were produced using the program PYMOL (5). Domain motion was analyzed by DynDom (6).

GST Pulldown Assay. ^{35}S-radiolabelled proteins were prepared by in vitro transcription translation (Promega). GST-β-catenin armadillo repeat region (138–663), GST-LRH-1 LBD (294–541), and GST-androgen receptor (AR) LBD (646–919) were bacterially expressed and purified with glutathione-Sepharose 4B beads. GST fusions were mixed with in vitro translated protein in buffer [final concentration 20 mM Tris·HCl (pH 8.0), 100 mM NaCl, or with dihydrotestosterone in the case of AR-LBD] at 4 °C for 90 min. Beads were pelleted, washed three times with 0.5 mL of incubation buffer, and analyzed by SDS/PAGE. Peptides from T-cell factor 4 (TCF-4) (DLGANDELISFKDEGEQEEKSSENSSAERDLADVKSSLVN, >95% purity) and B-cell chronic lymphocytic leukemia/lymphoma 9 (BCL9) (SQEQLEHRERSLQTLRDIQRMLFPDEKE, >95% purity) were purchased from GenScript USA Inc.

Cell Culture and Transfection Assays. CV-1 cells were maintained in Dulbecco's modified Eagle's medium H-21 4.5 g/L glucose, containing 10% steroid depleted fetal bovine serum (Invitrogen), 2 mM glutamine, 50 units/mL penicillin, and 50 mg/mL streptomycin. For transfection, (2.5×10^5) cells per well were plated and incubated overnight. A mixture of typically 200 ng of GAL4 responsive luciferase reporter plasmid, 10 ng of β-actin-β-galactosidase internal control, 10 ng of GAL-LRH-1 or GAL-AR expression vector or empty vector control, and 10–100 ng of β-catenin or empty vector control were mixed with 0.5 μL of transfection reagent from BioRad and incubated for 20 min and then applied to a single well of plated cells for 24 h. Cells were collected, and pellets were lysed in 100 μlL of 100 mM Tris·HCl (pH 7.5) containing 0.1% Triton X-100. luciferase and β-galactosidase activities were measured using the Luciferase Assay System (Promega) and Galacto-Light Plus-galactosidase reporter gene assay system (Applied Biosystems), according to the manufacturer's instructions. Full-length human β-catenin wild type in pCIneo were obtained from Addgene, Inc. (originally from Morin

et al., ref. 7). Mutants of the LRH-1 and β-catenin were made by QuickChange site directed mutagenesis kit (Stratagene). Detection of TCF-4 activity relied on the TOPflash/Fopflash reporter system (Millipore) with cotransfection of TCF-4 expression vector (a kind gift of the David Jablons laboratory, UCSF) and was used to assess effects of β-catenin mutants on TCF-4 interaction. Quantities of reporters and expression vectors were similar to that described above. The GAL-AR-LBD expression vector was described previously (8).

Statistical Analysis. Results are expressed as mean +/− s.e.m. from an appropriate number of samples as indicated in the figure legends. *t*-test was used to determine statistical significance.

1. Otwinowski Z, Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode, *Methods in Enzymol*, Macromolecular Crystallography, Part A, eds Carter, Jr CW, Sweet RM (Academic, New York) Vol. 276, pp. 307–326.
2. Moriarty NW, Grosse-Kunstleve RW, Adams PD (2009) Electronic ligand builder and optimization workbench (eLBOW): A tool for ligand coordinate and restraint generation. *Acta Crystallogr D Biol Crystallogr* 65:1074–1080.
3. Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) PROCHECK: A program to check the stereochemical quality of protein structures. *J Appl Cryst*, 26:283–291.
4. Kabsch W (1976) A solution for the best rotation to relate two sets of vectors *Acta Crystallogr A* 2:922–923.
5. DeLano WL (2002) *The PyMOL Molecular Graphics System* (DeLano Scientific, San Carlos, CA).
6. Poornam G, Matsumoto A, Ishida H, Hayward S (2009) A method for the analysis of domain movements in large biomolecular complexes. *Proteins* 76:201–212.
7. Morin PJ, et al. (1997) Activation of beta-catenin-Tcf signaling in colon cancer by mutations in beta-catenin or APC. *Science* 275:1787–90.
8. Estébanez-Perpiñá E, et al. (2005) The molecular mechanisms of coactivator utilization in ligand-dependent transactivation by the androgen receptor. *J Biol Chem* 280:8060–8068.

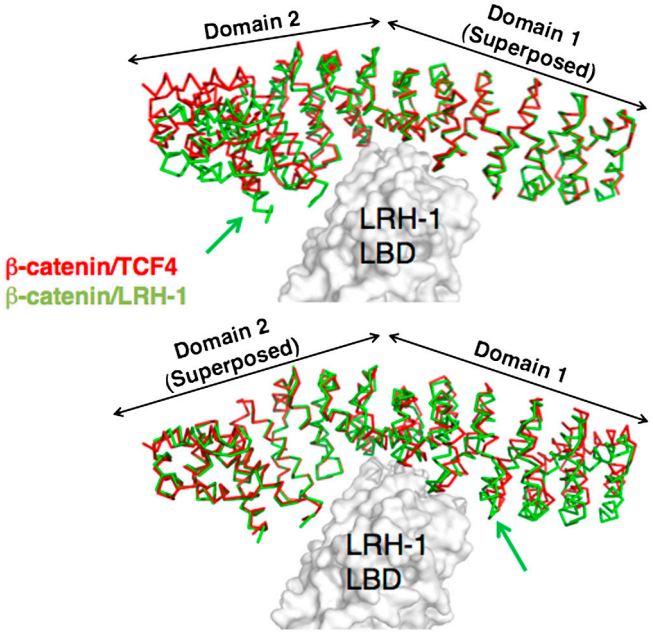

Fig. S1. Structural comparison of β-catenin complexed with LRH-1 LBD and with TCF-4. (*A*) Ribbons in green and red represent β-catenin in complexed with LRH-1 LBD (shown in gray surface) and with TCF-4, respectively. Structures of domain 1 (upper) and domain 2 (lower), defined by DynDom, were superposed by LSQKAB in CCP4 suite one domain by one domain.

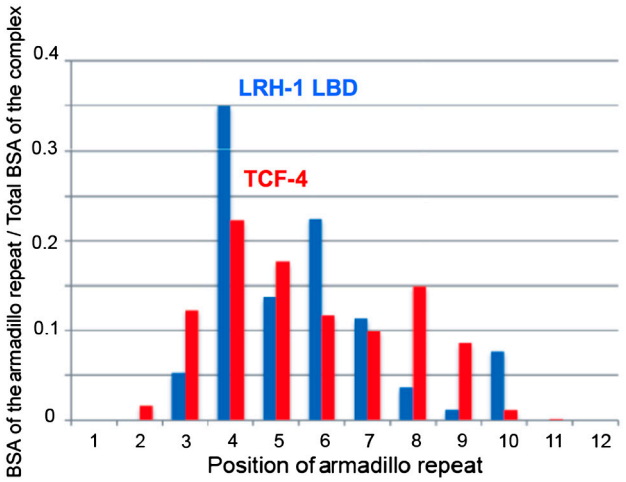

Fig. S2. Ratios of buried surface area (BSA) in each armadillo repeat against the total BSA were calculated and shown in the figure. Blue and red bars show the data on LRH-1 and TCF-4 interactions, respectively.

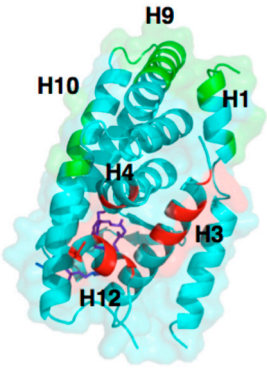

Fig. S3. Protein–protein interaction sites mapped on the ribbon schematic of LRH-1 LBD are about 30–40 Å apart. The β-catenin binding site (this study) and the steroid receptor coactivator-2 third box peptide binding site are colored green and red.

```
hLRH-1  SIPHLILELLKCEPDE---------------------------------------------- 315
mLRH-1  SIPHLILELLKCEPDE---------------------------------------------- 334
cLRH-1  SIPHLILELQKCEPDE---------------------------------------------- 275
hSF-1   NVPELILQLLQLEPDE---------------------------------------------- 237
mSF-1   NVPELILQLLQLEPEE---------------------------------------------- 238
hAR     TARDHVLPIDYYFPPQKTCLICGDEASGCHYGALTCGSCKVFFKRAAEGKQKYLCASRND 600
        .  . :* :    * :

hLRH-1  ATLNNLMSHAQELVAKLRSLQFDQREFVCLKFLVLFSLD-VKNLENFQLVEGVQEQVNAA 480
mLRH-1  VAFNNLLSLAQELVVRLRSLQFDQREFVCLKFLVLFSSD-VKNLENLQLVEGVQEQVNAA 499
cLRH-1  ATLNNLMSHAQELVAKLRSLQFDLREFVCLKFLVLFSLD-VKNLENFQLVEGVQEQVNAA 440
hSF-1   SLLHSLVLRAQELVLQLLALQLDRQEFVCLKFIILFSLD-LKFLNNHILVKDAQEKANAA 400
mSF-1   SLLHSLVLRAQELVLQLHALQLDRQEFVCLKFLILFSLD-VKFLNNHSLVKDAQEKANAA 401
hAR     --MYSQCVRMRHLSQEFGWLQITPQEFLCMKALLLFSIIPVDGLKNQKFFDELRMNYIKE 837
          : .     :.*  .:  **:  :**:*:* ::***   :. *:*  :..  : :

hLRH-1  LLDYTMCNYP---QQTEKFGQLLLRLPEIRAISMQAEEYLYYKHLNGDVP---YNNLLIE 534
mLRH-1  LLDYTVCNYP---QQTEKFGQLLLRLPEIRAISKQAEDYLYYKHVNGDVP---YNNLLIE 553
cLRH-1  LLDYTMCNYP---QQTDKFGQLLLRLPEIRAISMQAEEYLYCKHLNGDVP---CNNLLIE 494
hSF-1   LLDYTLCHYP---HCGDKFQQLLLCLVEVRALSMQAKEYLYHKHLGNEMP---RNNLLIE 454
mSF-1   LLDYTLCHYP---HCGDKFQQLLLCLVEVRALSMQAKEYLYHKHLGNEMP---RNNLLIE 455
hAR     LDRIIACKRKNPTSCSRRFYQLTKLLDSVQPIARELHQFTFDLLIKSHMVSVDFPEMMAE 897
        *     *:          :* **   * .::.:: : .:: :   : ..:      ::: *
```

Fig. S4. Amino acid sequence alignment of LRH-1s, steroidogenic factor 1s (SF-1s), and AR in region of β-catenin-binding site in LRH-1. Sequences were obtained from Uniprot database for human LRH-1(hLRH-1, O00482), mouse LRH-1(mLRH-1, P45448), chicken LRH-1(cLRH-1, O42101), human SF-1(hSF-1, Q13285), mouse SF-1(mSF-1, P33242), and human AR(hAR, P10275). The sequences were aligned with Clustal W2 (1).

1 Chenna R, et al. (2003) Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Res* 31:3497–3500.

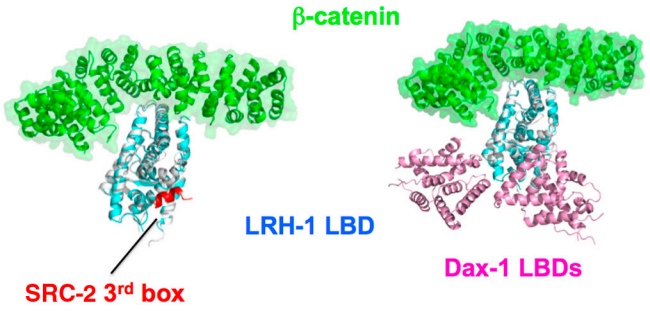

Fig. S5. (*A*) Superposed structures of LRH-1/β-catenin LBD and LRH-1/SRC-2 third box peptide complexes. LRH-1 LBD in LRH-1/SRC-2 third box peptide complex is colored in white. The peptide is colored in red. (*B*) Superposed structures of LRH-1/β-catenin LBD and mouse LRH-1 LBD/dosage-sensitive sex reversal, adrenal hypoplasia critical region, on chromosome X, gene 1 (Dax-1) LBD dimer. The Dax-1 dimer is colored in pink.

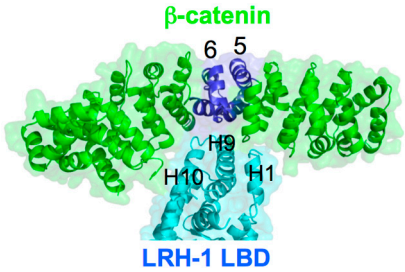

Fig. S6. (*A*) Image of SDS/PAGE gel showing that bacterially expressed LRH-1 LBD competes for AR-LBD binding to GST-β catenin. The numbers for LRH-1 LBD show the molar ratio against AR-LBD. (*B*) AR interaction site on β-catenin surface is mapped onto LRH-1/β-catenin structure. Numbers show that the positions of ARM-5 and -6.

Table S1. Crystallographic data

X-ray source Advanced Light Source BL 8.3.1	
Wavelength (Å)	1.116
Space group	$P2_1$
Unit-cell parameters (Å)	$a = 49.8$, $b = 151.6$, $c = 76.1$ $\beta = 96.96°$
No. of complexes/asym. Unit	1
No. of unique reflections	28946
Resolution range (A)	47.02–2.75 (2.80–2.75)
R_{sym} (%)	7.4
Completeness (%)	99.9 (99.7)
Data redundancy	15.2 (15.4)
$\langle I\rangle/\langle s(I)\rangle$	55.0/5.4
Refinement	
Resolution (Å)	49.457–2.755
No. reflections	28,882
Completeness for the range (%)	99.76
R_{work}/R_{free}	19.88/24.26
No. atoms	5,644
B-factors (Å^2)	
Protein	74.9
Ligand of LRH-1 LBD	79.3
Ramachandran plot (%)	favored 93.0; allowed 7.0; outlier 0
Rms deviations	
Bond lengths (Å)	0.008
Bond angles (°)	1.155

Values in parentheses are for highest-resolution shell. Data for the LRH-1/β-catenin structure were collected from a single crystal. $R_{sym} = \Sigma_{hkl}\Sigma_i|I_i(hkl) - \langle I(hkl)\rangle|/\Sigma_{hkl}\Sigma_i I_i(hkl)$. Asym., asymmetric.

Table S2. Summary of analyses of conformational changes in β-catenin from DynDom

LRH-1 complex/TCF-4 complex (1JDH)		
	Domain 1	Domain 2
Size (residues)	261	238
Region	150–388, 394–409, 414–419	389–392, 410–413, 420–659
Rmsd (Å)	1.18	0.7
Rotation angle (°)	12.4	
Translation (Å)	-0.1	
Closure (%)	24.4	
Bending residues (Region of ARMs)	388–396, 409–414, 419–420 (ARM-6, ARM-7)	

Rotation angle relates to the angle change of the principal axes of the two domains. Twist axis is parallel to the line connecting the centers of mass of the two domains, and closure axis is perpendicular to the twist axis (1). Closure relates to a percentage measure of the degree of closure motion, which is defined from the square of the projection on the closure axis. Maximal possible domain closure is set at 100% and varies between β-catenin complexes. The domain definitions and configurations are shown for β-catenin when bound to LRH-1 compared with TCF-4. ARMs, armadillo-repeat regions.

Hayward S, Berendsen HJ. (1998) Systematic analysis of domain motions in proteins from conformational change: new results on citrate synthase and T4 lysozyme. *Proteins* 30:144–154.