Supporting Information

Dvir et al. 10.1073/pnas.1114128109

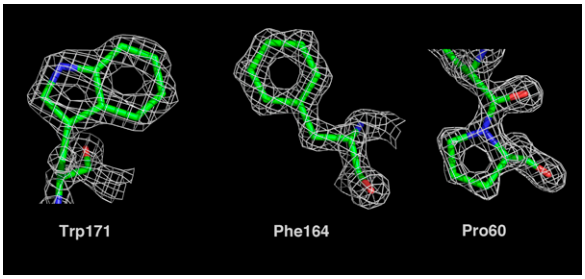

Fig. S1. Representative electron density of the refined structure of the autosomal recessive hypercholesterolemia (ARH)–LDL receptor (LDLR) tail complex. The 2Fo-Fc electron density map (gray mesh) contoured at 1.5 σ around selected ARH residues (green sticks for carbons, red and blue for oxygens and nitrogens, respectively) reveals outstanding atomic resolution.

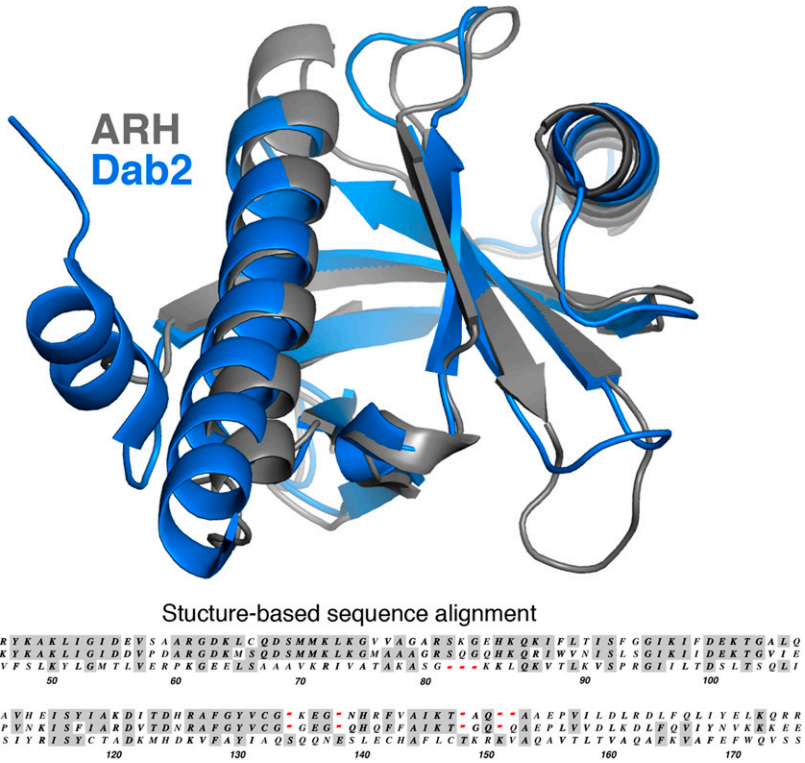

Fig. S2. Superposition of the structure of ARH (gray) and Dab2 (blue) phosphotyrosine-binding (PTB) structures. The overall fold (*Upper*, ribbon representation) is quite similar despite the low sequence identity (shown in the structure-based alignment, *Lower*). The PTB domains of Dab1 and Dab2 are 57.3% identical, but both are less than 18% identical with ARH. Identical residues are shaded.

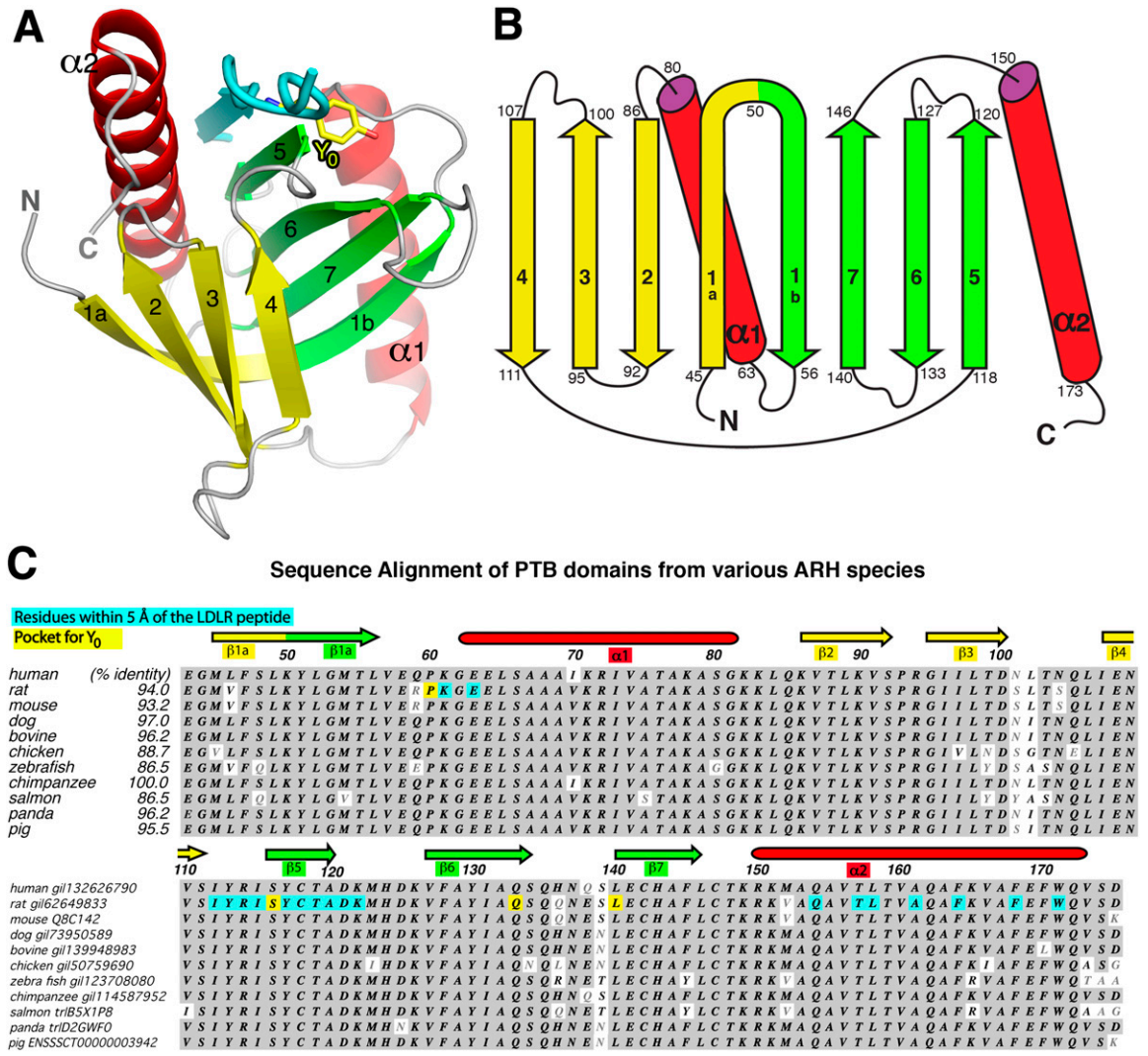

Fig. S3. Topology of the conserved PTB domain of ARH. (*A*) 3D structure of rat ARH (rARH) in complex with LDLR peptide. β-Strands of ARH are represented as green and yellow numbered arrows (β1–β7) and α-helices as red ribbons (α1, α2). The LDLR tail is drawn in cyan and the tyrosine (Y_0) of its FxNPxY_0 motif in yellow. (*B*) 2D topology of the PTB domain of ARH shows that it is made of an array of seven anti-parallel β-strands grouped into two β-sheets (yellow and green). Numbers at the top and bottom of each strand and helix indicate the residues that form them, respectively. (*C*) Sequence alignment of PTB domains from various ARH species. The numbering at the top corresponds to the rARH sequence, whereas the percentage identities to the left of each sequence refer to the human sequence. Identical residues are shaded. Amino acids in the proximity (within 5 Å) of the LDLR tail are highlighted in cyan. Residues lining the pocket for Y_0 of the LDLR FXNPxY_0 motif are highlighted in yellow.

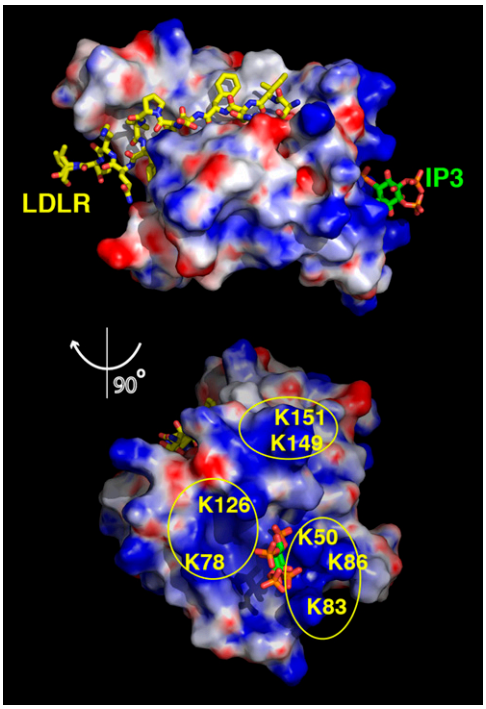

Fig. S4. Proposed binding area for phosphoinosites. Electrostatic surface representation of the structure of the PTB domain of ARH bound to the LDLR (yellow sticks). The related structure of the Dab1 (1) (PDB id 1NU2) in complex with a inositol-triphosphate (IP3) was overlaid on that of ARH according to the protein positions, and the resulting position of the IP3, the phosphorylated head group of PIP2, is displayed green sticks for carbons and orange for the phosphates. Although the ARH side-chains in corresponding positions to those that bind IP3 in Dab1 are not basic, there are three clusters of lysine residues (labeled in yellow, bottom view) in this vicinity that could bind negatively charged phosphate groups. This indicates that the mechanism of binding phosphoinositides by ARH is not identical but similar to that of Dab2. The electrostatic properties were evaluated using APBS (2) and displayed in PyMol, from −7 kT/e (red) to +7 kT/e (blue).

1. Stolt PC, et al. (2003) Origins of peptide selectivity and phosphoinositide binding revealed by structures of disabled-1 PTB domain complexes. *Structure* 11:569–579.
2. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA (2001) Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc Natl Acad Sci USA* 98:10037–10041.

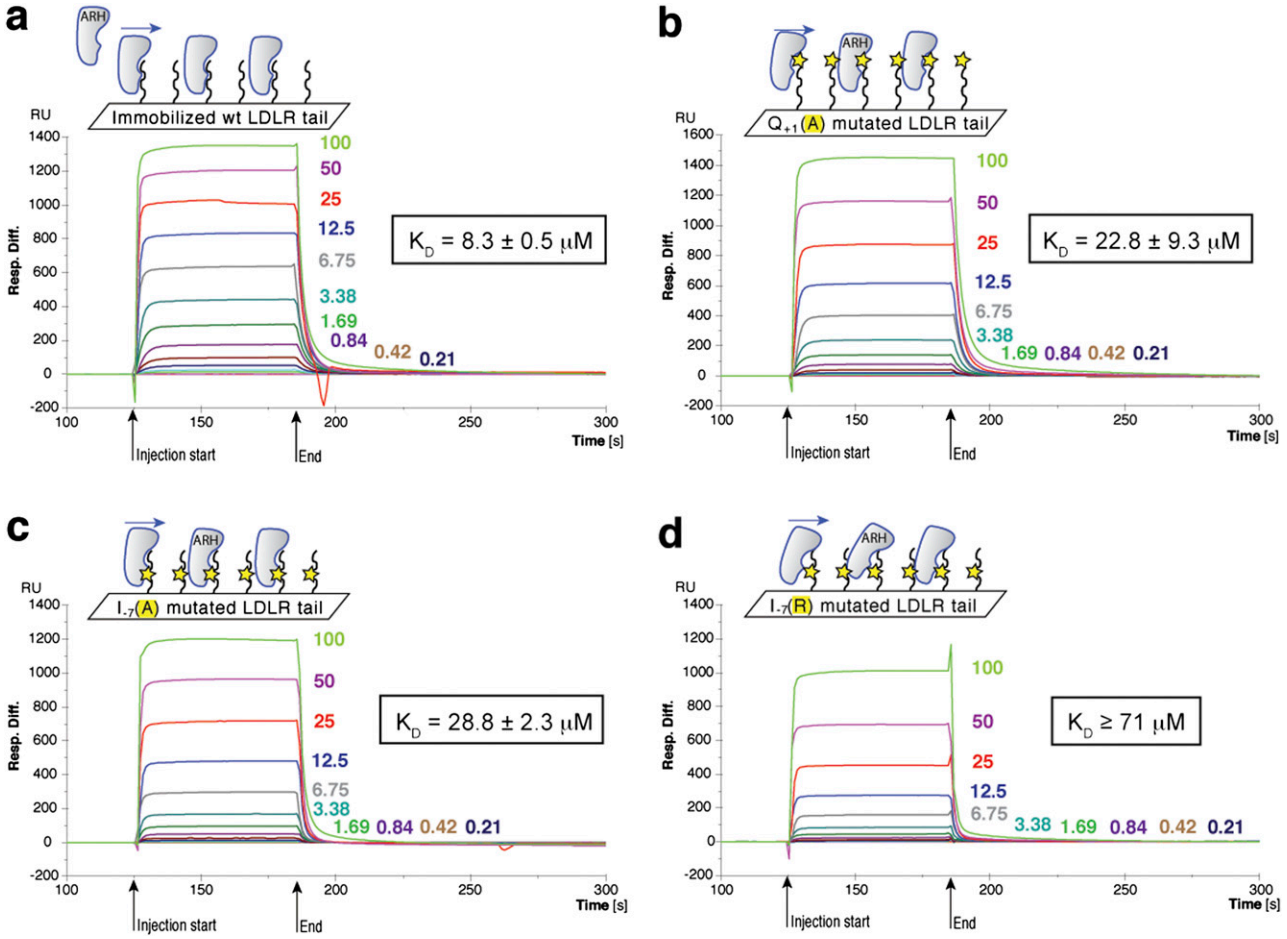

Fig. S5. Steady-state representative surface plasmon resonance sensorgrams from which Fig. 5 was generated.

Table S1. List of protein sequences for interaction with ARH, which resulted from searching a sequence database (motif.genome.jp/MOTIF2.html)

Protein name	Database entry	Sequence
LDLR	P01130	INFDNPVYQ
DAB2 interaction protein	Q5VWQ8	LSFQNPVYQ
LRP8 (ApoER2)	Q14114	MNFDNPVYR
LRP4 human	O75096	LTYSNPSYR
MAP kinase kinase 2 (MAPKK 2)	Q07192	LNFANPAFK
Tryptophan synthase α chain	Q8TLP4	MTYYNPVFR
Ras GTPase-activating protein nGAP	Q9UJF2	LSFQNPVYH
Ras guanine nucleotide exchange factor K	Q54FF3	IRLTNPVFH
Fusion glycoprotein F0	Q2Y2M3	VTIDNPVYQ
DNA-directed RNA polymerase subunit β	A1SEK1	LSFENPVFY
RING finger protein 31	Q96EP0	VKFNNPVFR
Tyrosine-protein phosphatase YVH1	Q02256	VDFDNPAYK
Lipophorin receptor - *Aedes aegypti* (yellow fever mosquito)	GenBank: AAQ16410.1	MNFDNPVYR
Parkinson disease 7 domain-containing protein 1	Q3B7H1	FNLCNPVFN
Cell cycle control protein 50C (extracellular domain)	Q2T9P5	VKFQNPAFQ
Minor capsid protein L2 (HPV type 67)	Swiss O90729.1	ITFDNPAFQ
L2 protein - HPV 97	GenBank: ABO27082.1	ITYDNPAFN
Carm1-pending protein	XP_003144122	LDLKNPLFR
MEGF7, multiple EGF-like motifs (LRP4/LRP10)	GenBank: BAA32468.1	LTYSNPSYR
Reticuline oxidase	P30986	LSIQNPLFQ
Bifunctional aspartokinase/homoserine dehydrogenase	P44505	VDLNNPLYK
2′,3′-Cyclic-nucleotide 2′-phosphodiesterase	P47376	INIKNPLYN
Epithelial discoidin domain-containing receptor 1	Q08345	LLLSNPAYR
β-1,3-GalNAc-T2	Q5M900	LNITNPTFK
Rhizobactin siderophore biosynthesis protein rhbC	Q9Z3R0	LQITNPLFE
Putative translation initiation factor eIF-2B subunit 2	Q57586	IKIRNPAFD
Probable cation-transporting ATPase 13A4	Q4VNC1	VSYSNPVFE
1,4-α-D-glucan glucanohydrolase	P53354	FSLSNPLFN
LRP1B	Q9NZR2	VEIGNPSYN
Diadenosine tetraphosphatase	A3N0C2	FKLDNPLFH
β-1,3-*N*-acetylgalactosaminyltransferase II	Q5M900	LNITNPTFK
Transcription factor castor	Q7M3M8	FQLQNPLFY
Ceroid-ipofuscinosis neuronal protein 5 homolog	Q553W9	YNISNPVYQ
CCR4-NT transcription complex subunit 1	A0JP85	YELDNPSFQ
Protein dltD	P39578	FHIDNPVYK
DNA polymerase	P87503	FALPNPAYK
Formmide amidohydrolase	Q50228	YGIKNPVFQ
Probable Ras GTPase-activating protein	Q8T498	LAFKNPSYQ
High-affinity potassium transporter (yeast)	P50505	IQIANPSFN
Spheroidin (viral)	P23061	VYITNPSFN
GTPase-activating protein SynGAP	Q96PV0	LSFQNPLFH
Uncharacterized protein L442 (*Acanthamoeba polyphaga* mimivirus)	Q5UQN6	MGFHNPSYN
EGFR(Receptor tyrosine-protein kinase erbB-1)	P00533	GSVQNPVYH
LRP1B	Q9NZR2	VEIGNPSYN
LRP1 (ApoER)	Q07954	VEIGNPTYK
Hepatocyte growth factor receptor	P08581	IYVHNPVFK
Probable phospholipid-transporting ATPase VB	O94823	GQLSNPTFY

The structure-based ARH specificity formula $\Phi_{-7}x\Phi_{-5}xNP\psi[F/Y]_0\Omega_{+1}$ was used. Φ stands for a bulky aliphatic or aromatic residue such as Ile, Met, or Phe; x for mostly polar (these are the noncontacting residues that contribute least to the binding); ψ for small aliphatic residues such as Val, Ala, or possibly Thr; and Ω for long polar residues with hydrogen-donating group such as Gln, Asn, His, Tyr, Lys, or Arg. To keep all of the possible sequences resulting from this search, the whole raw list is presented and no filtering for specific organisms was applied. Only sequences of cytoplasmic domains are shown. Uniprot database entries are listed unless otherwise indicated.