

Figure S1. Related to Figure 1. Purification and analysis of mouse GPR56 ECR

(A) (Left) SDS-PAGE analysis of mouse GPR56 ECR purified from insect cells. N-terminal fragment (NTF, black arrow) and *Stachel* (green arrow) are observed. (Right) Gel filtration chromatography of cleaved GPR56 ECR. (B) Multi-Angle Laser Light Scattering (MALLS) analysis of purified GPR56 ECR and calculated molecular weight, which is consistent with purely monomeric glycosylated GPR56 ECR and inconsistent with any higher-order oligomer. (C) MALDI-TOF mass spectrometry of purified GPR56 ECR (blue) and GAIN domain only (red). Calculated molecular weights are displayed above peaks. Calculated molecular weights are consistent with cleaved NTFs (with inhomogeneous glycosylation). (D) Circular Dichroism (CD) spectroscopy temperature melts of mouse GPR56 ECR (top) and GAIN domain only (bottom). (E) Representative image of mouse GPR56 ECR-monobody $\alpha 5$ complex crystals used for x-ray data collection. (F) Representative diffraction pattern for native data set. (G) Representative electron density of GPR56 ECR crystal structure showing lack of density at GPS cleavage site, indicating complete autoproteolysis. (H) Interactions between the GPR56 *Stachel* and the GAIN domain NTF. Water molecules shown as red spheres. (F) Orientation of cleaved *Stachel* in GAIN domain of GPR56 (NTF, gray; *Stachel*, green) and Lphn1 (NTF, pink; *Stachel*, cyan). * indicates autoproteolysis site.

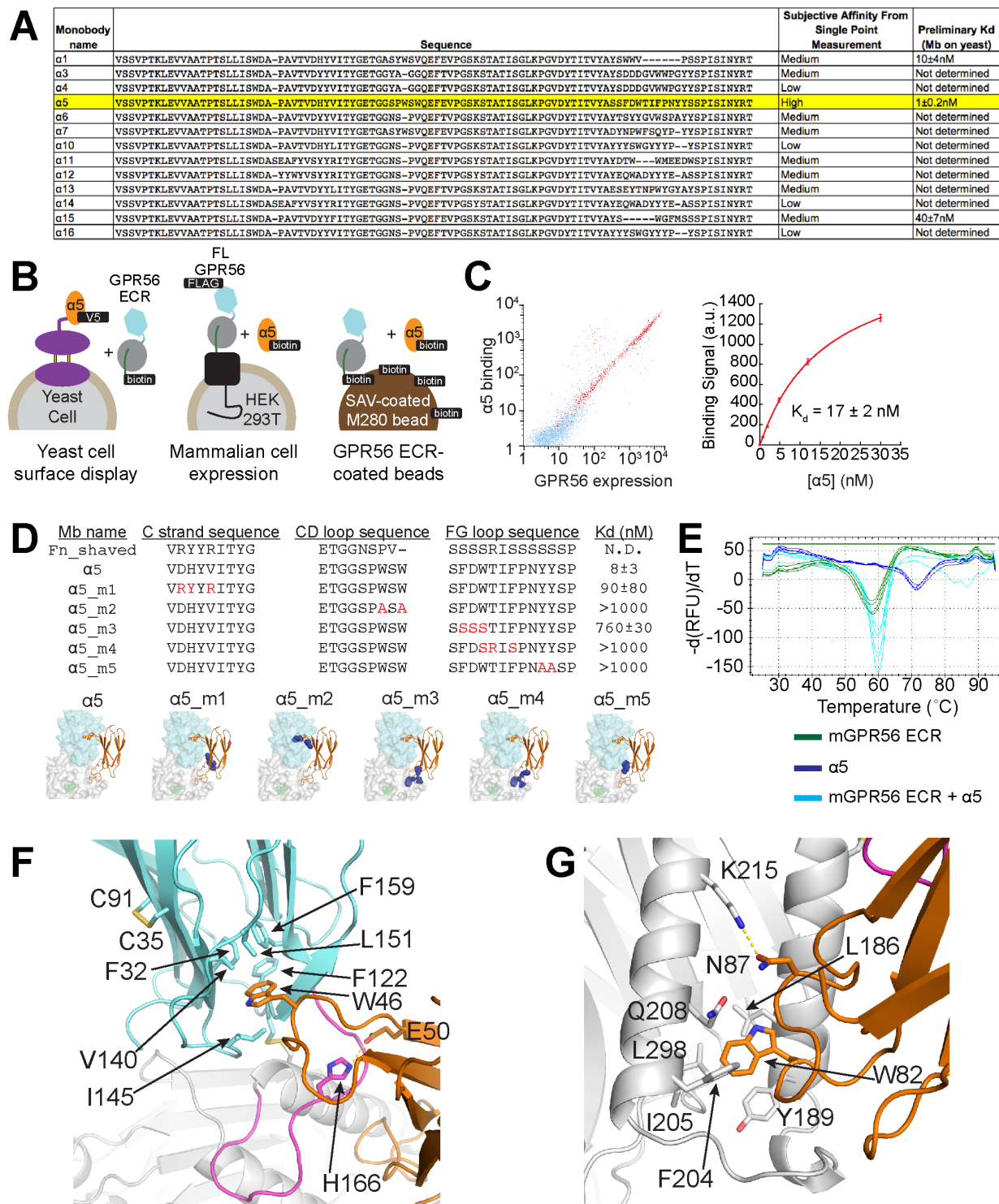


Figure S2

Figure S2. Related to Figure 1. Monobody generation and characterization

(A) Sequences of 13 monobodies selected with affinity for mouse GPR56 ECR. Highest affinity clone, $\alpha 5$ (yellow), used for subsequent experiments. Affinities evaluated by yeast surface display. (B) Schematic of monobody yeast surface display, mammalian cell expression of full-length GPR56, and purified GPR56 ECR-coated M280 beads used for binding affinity measurements. Note: after coating with GPR56 ECR and before $\alpha 5$ binding, M280 beads were blocked with excess biotin. (C) Flow cytometry of HEK293 cells stained with $\alpha 5$. Untransfected and full-length-FLAG-GPR56-transfected cells shown in blue and red, respectively. (Left) Dot plot showing correlation between GPR56 expression (anti-FLAG staining) and $\alpha 5$ binding. (Right) Concentration titration of $\alpha 5$ to measure binding affinity. (D) (Top) $\alpha 5$ point mutations result in decreased affinity for GPR56 ECR by M280 bead assay. (Bottom) Location of $\alpha 5$ point mutations m1-m5 on the crystal structure. Mutated residues shown as blue spheres. (E) Differential scanning fluorimetry temperature melts of mGPR56 ECR ($T_m = 58.13 \pm 0.13^\circ\text{C}$), $\alpha 5$ ($T_m = 71.38 \pm 0.13^\circ\text{C}$), and mGPR56 ECR- $\alpha 5$ complex ($T_m = 59.75 \pm 0.14^\circ\text{C}$). Quadruplicate measurements for each sample are plotted as negative temperature derivative of RFU vs temperature. Peak corresponds to T_m . (F-G) Close-up views of the binding interfaces between $\alpha 5$ and (E) the PLL domain and PLL-GAIN linker and (F) the GAIN domain. Residues at the binding interface are shown as sticks. The PLL domain, PLL-GAIN linker, GAIN domain, and $\alpha 5$ are colored cyan, magenta, gray, and orange, respectively. Polar contacts are indicated by yellow dashes.

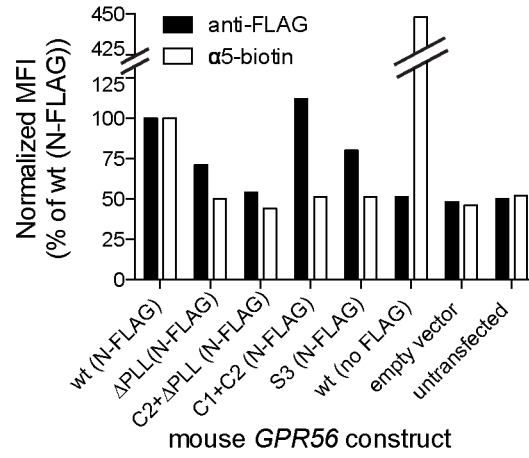


Figure S3. Related to Figure 3. Cell-surface expression of mouse GPR56. Flow cytometry of various *GPR56* constructs stained with anti-FLAG antibody and $\alpha 5$.

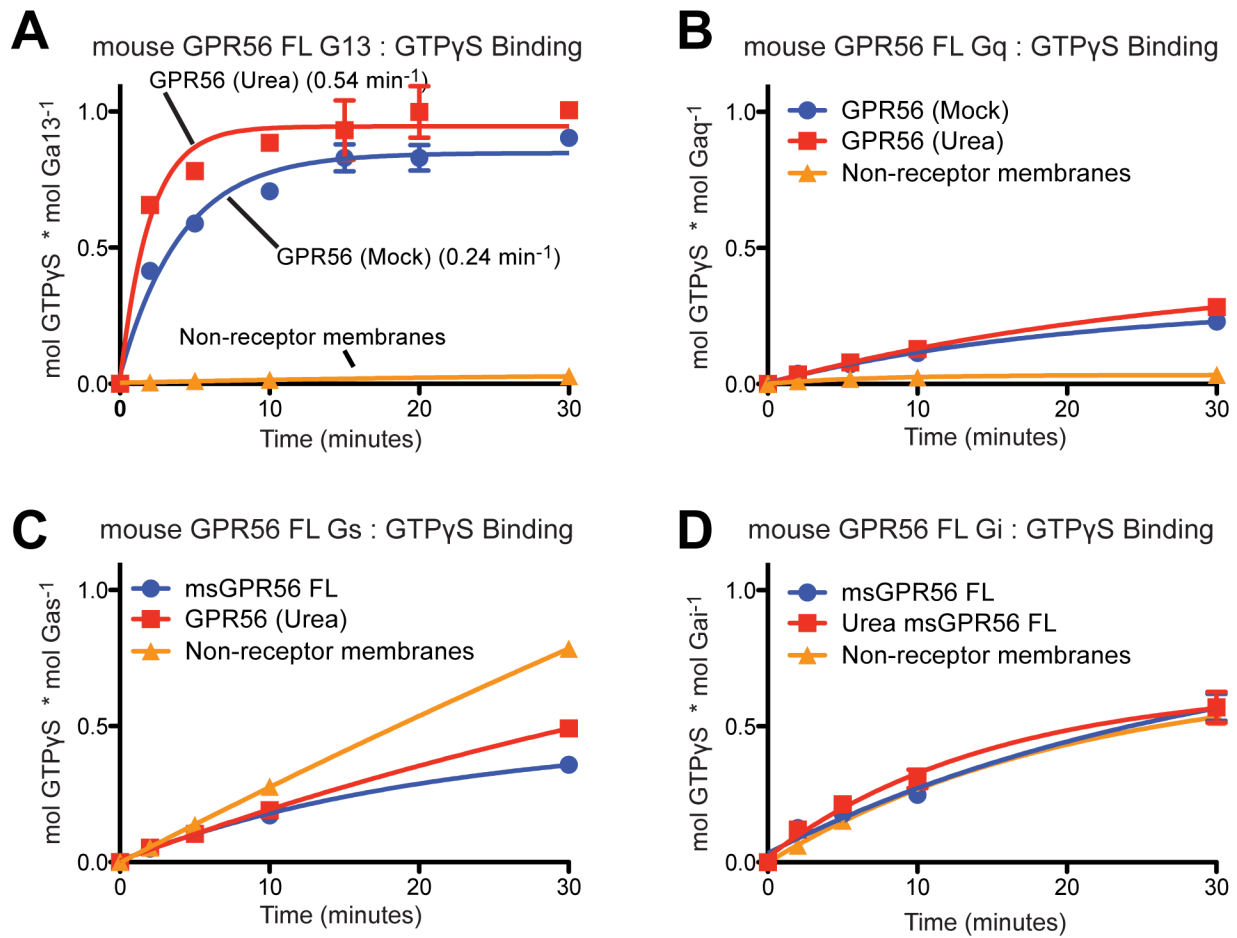


Figure S4. Related to Figure 3. G protein signaling of mouse GPR56. High-Five membranes with full-length mouse GPR56 were subject to mock or urea treatment to induce NTF-shedding and compared to membranes with no receptor. Membranes were reconstituted with G proteins (α and $\beta\gamma$). The receptor-mediated G protein activation kinetics were measured using the $[^{35}\text{S}]$ -GTP γ S binding assay. (A) GPR56 and $G\alpha_{13}$. (B) GPR56 and $G\alpha_q$. (C) GPR56 and $G\alpha_s$. (D) GPR56 and $G\alpha_i$.

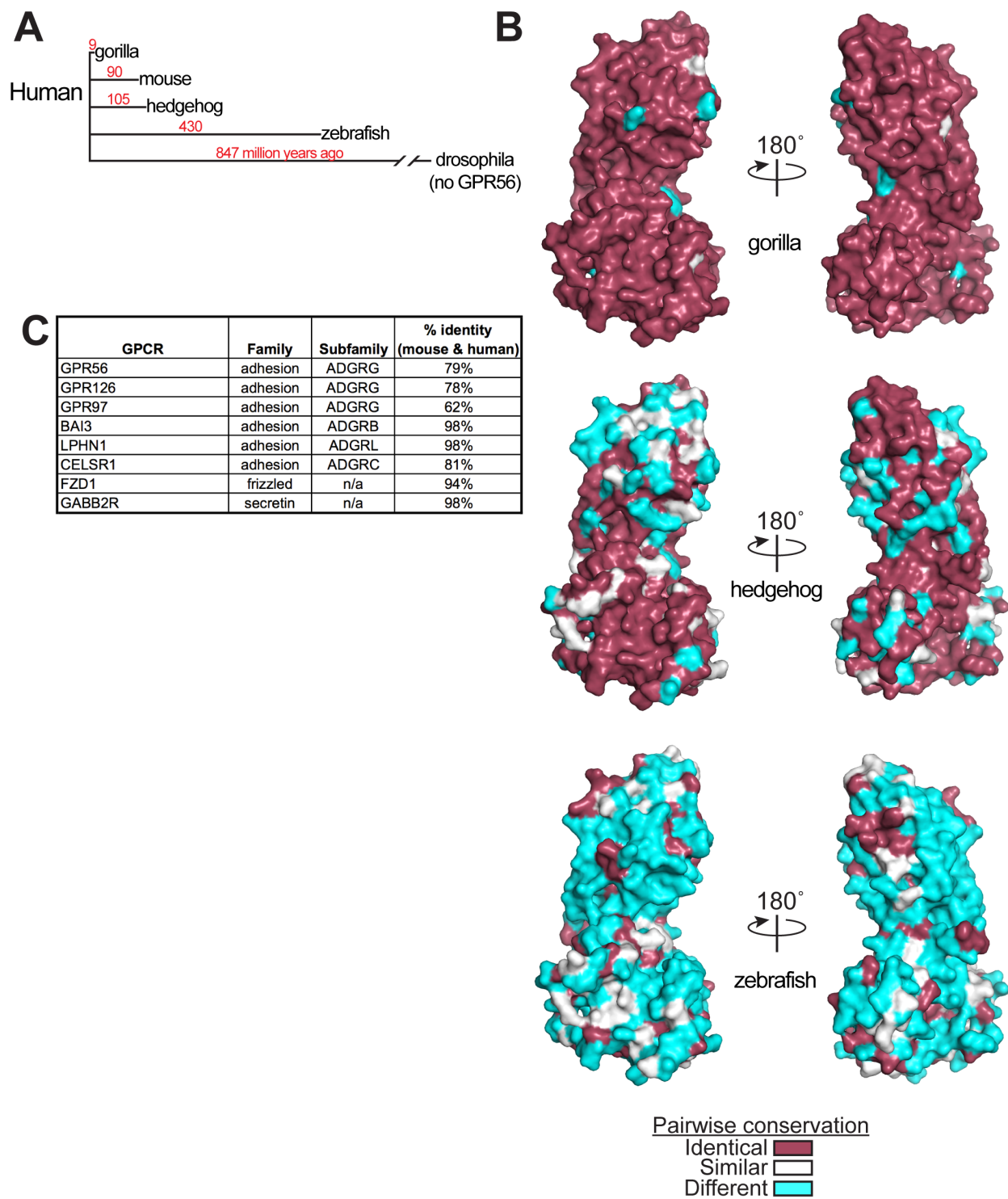


Figure S5. Related to Figure 4. Pairwise surface conservation analysis of PLL domain

(A) Cartoon illustrating approximate time since last common ancestor with human. (B) Pairwise surface conservation analysis of human GPR56 ECR vs gorilla, hedgehog, and zebrafish mapped to the structure of mouse GPR56 ECR. (C) Percent identity between full length human and mouse protein sequences of GPCRs from several families.

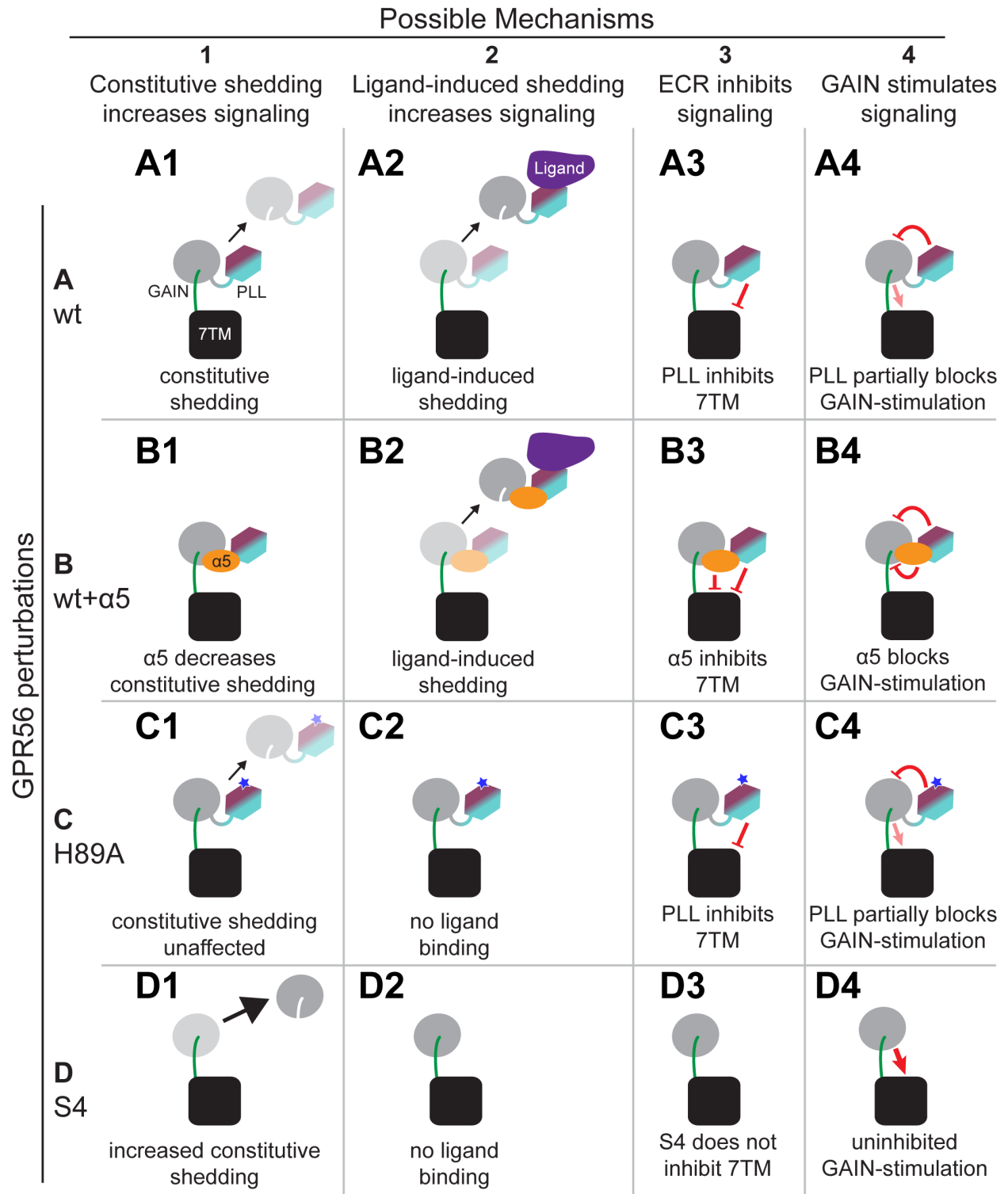


Figure S6

Figure S6. Related to Figure 7. Detailed possible mechanisms leading to experimental observations of GPR56 activity

Four possible mechanisms are proposed (numbered 1-4) that explain the observed phenomena. We do not rule out the possibility of additional mechanisms nor do we deny that more than one proposed mechanism may occur. In our experiments, four systems were tested, each corresponding to a GPR56 perturbation (labeled A-D). Black arrows represent NTF shedding. Pointed (stimulating G-protein signaling) and flat-ended (inhibiting G-protein signaling) red arrows may represent conventional binding interactions, weak and transient interactions, or random collisions.

Proposed mechanism 1 relies on the hypothesis that, even in the absence of ligand, there exists some small but significant constitutive level of shedding in GPR56. By this mechanism, the higher the level of constitutive shedding, the more G protein signaling occurs, due to the observation that the CTF of the receptor is highly active. (A1) wt GPR56 undergoes some level of constitutive shedding leading to the observed basal activity of the receptor. (B1) Upon addition of $\alpha 5$, the constitutive shedding decreases, leading to the observed decrease in G protein signaling. (C1) The H89A mutation on the conserved face of PLL domain does not alter constitutive shedding, resulting in the observation that H89A has similar basal activity to wt. (D1) The ECR of S4, corresponding to only the GAIN domain, has an increased level of constitutive shedding, leading to the observed high basal activity of the Δ PLL construct.

Proposed mechanism 2 relies on the hypothesis that a ligand may bind the conserved face of the PLL domain, exert mechanical force, and thereby cause NTF shedding. Similar to mechanism 1, by this mechanism, with more ligand-induced shedding, more G protein signaling occurs, due to the observation that the CTF of the receptor is highly active. (A2) wt GPR56 undergoes some level of ligand-induced shedding. (B2) Upon addition of $\alpha 5$, the ligand-induced shedding is not affected. (C2) The H89A mutation on the conserved face of PLL domain abolishes ligand binding, and therefore blocks ligand-induced shedding. (D2) The ECR of S4 lacks the ligand-binding site on the PLL domain, and therefore does not undergo ligand-induced shedding.

Proposed mechanism 3 relies on the hypothesis that domains in the ECR directly or indirectly inhibit signaling by interacting in some way with the 7TM. (A3) The PLL domain in wt GPR56 partially inhibits 7TM. (B3) Together, the PLL domain and $\alpha 5$ strongly inhibit the 7TM. (C3) The H89A mutation on the conserved face of PLL domain does not affect PLL-mediated inhibition of the 7TM. (D3) The ECR of S4 lacks the inhibitory PLL domain, and therefore has higher basal activity than wt.

Proposed mechanism 4 relies on the hypothesis that the GAIN domain stimulates signaling by directly or indirectly interacting with the 7TM. (A4) The PLL domain in wt GPR56 partially inhibits the stimulatory GAIN domain. (B4) Together, the PLL domain and $\alpha 5$ strongly inhibit the stimulatory GAIN domain. (C4) The H89A mutation on the conserved face of PLL domain does not affect PLL-mediated inhibition of the GAIN domain. (D4) The ECR of S4 lacks the inhibitory PLL domain, and therefore has higher basal activity than wt.

Construct	Description of mutation	Surface expression (HEK293T)	SRE assay signaling (not normalized for expression)	function in zebrafish
wild-type	wt FL mGPR56	3	3	0
wild-type (N-FLAG)	wt FL mGPR56 with N-FLAG	1	0	ND
R33E	R33N is a human disease mutation	2	ND	ND
G36D+Q37A	S36P is a human disease mutation	0	ND	ND
R38Q	human disease mutation	0	ND	ND
N39A	delete N39-linked glycosylation	2	ND	ND
T41A	delete N39-linked glycosylation	2	ND	ND
Y88S	Y88C is a human disease mutation	1	ND	ND
Y88C	human disease mutation	0	ND	ND
H89A	conserved, surface-exposed patch	3	4	3
C91S	human disease mutation	0	ND	ND
G106R	conserved, surface-exposed patch	1	ND	ND
C121S+C177S	delete interdomain disulfide	4	4	0
A137R+S139R	conserved, surface-exposed patch	2	ND	ND
N148A	delete N148-linked glycosylation	3	ND	ND
S150A	delete N148-linked glycosylation	3	4	0
H381S	autoproteolysis-null	2	2	1
T383G	autoproteolysis-null	4	2	ND
S3	splice variant 3	0	ND	ND
S4	splice variant 4 (Δ signal peptide + Δ PLL)	0	0	ND
Δ PLL	S4 with signal peptide	1	4	ND
CTF	T383M,Y384,etc. through native C-terminus	1	4	ND
7TM	H401M,Y402,etc. through native C-terminus	0	0	ND

Construct	Description of mutation	Yield in insect cells (purified ECR)
wild-type	wt FL mGPR56	3
R38Q	human disease mutation	0
R38W	human disease mutation	0
Y88C	human disease mutation	1
Y88S	Similar to Y88C	1
C121S+C177S	delete interdomain disulfide	4

Scale:	
none	0
little bit	1
less than wt	2
comparable to wt	3
more than wt	4
not determined	ND

Table S1. Related to Figure 3. Expression, signaling, and myelination phenotypes of all mouse GPR56 mutants tested.

Alignment S1. Multiple sequence alignment of full-length GPR56 from 14 species. Conservation score is calculated for positions with at least some detectable conservation (9 is most conserved). ECR domain structure is annotated based on the crystal structure of the ECR, while 7TM and intracellular region (ICR) are annotated based on sequence-based predictions: s, signal peptide; P, PLL; l, linker; G, GAIN; 7, 7TM; I, ICR. Secondary structure is also annotated for the ECR: α , α -helix; β , β -strand based on the crystal structure. Cysteine residues involved in the intra-PLL domain disulfide bond are colored red. Cysteine residues involved in the interdomain (PLL-GAIN) disulfide bond are colored blue. The conserved pentraxin motif on the PLL domain, beginning with mouse H89, is highlighted yellow. The conserved N-linked glycosylation site (glycan is adjacent to the conserved patch on the PLL domain) is highlighted green. Autoproteolysis occurs between the two residues highlighted in black. Bolded residues have been found mutated in human diseases.

```

Conservation: 94 4 4 64 446 4744449 4 69966964666 63 646 94 3 494
cuckoo 1 MK-----VLLLLLSPLOGVANGDQEEFRFCGDRNQ-TQNSSVIYEH-SP-ANIS
chicken 1 MK-----VLLLLLSPLOGVGVSG-RQEDFRFCGDRNQ-TENSSVIYEH-SS-ANIS
zebrafish 1 MKQNPAKTARMWIIICLLFLVLQQA-TD-NDRDFKMGKWLHGIAPQLEYDLKTGCERIE
bat 1 MAAPVQVTGPLLLGLLLLVQQAQAGG-FREDFRFCGQRNQ-TQKSSLYER-WTQLRIS
bovine 1 MTAQVLLQMPFLLLG-LFLVPGAHGGG-PREDFRFCGQRNQ-TQNSSLHYKQ-ASQLHIS
elephant 1 MAAQGL-QTALFPLVLLIQTLCGG-LQEDFRFCGQRNQ-TQTSNLYEQ-TAVLHIS
mouse 1 MAVQVLRQMVYFLLSLFSLVQGAHSGS-PREDFRFCGQRNQ-TQOSTLHYDQ-SSEPHIF
rat 1 MAVQVLLQMVYFLLTLFLVQGAHGAS-PREDFRFCGQRNQ-TQOSTLHYDQ-TSEPHIF
gorilla 1 MTAQSLLQTTFLLLFLVQGAHGRG-HREDFRFCSQRNQ-THRSSLHYKP-TADLRIS
monkey 1 MTAQSLLQTTFLLLFLVQGAHGRG-HREDFRFCSQRNQ-THISSLHYKF-TPDLRIS
chimp 1 MTAQSLLQTTFLLLFLVQGAHGRG-HREDFRFCSQRNQ-THRSSLHYKP-TPDLRIS
human 1 MTPQSLLQTTFLLLFLVQGAHGRG-HREDFRFCSQRNQ-THRSSLHYKP-TPDLRIS
consensus 1 M aq llqt lflsllflvqGahg g reDfrfCqgrnq tq sslhYe t elrIs
ss ββββββββββ ββββββ
domain sssssssssssssssssssssssssssPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP

```

```

Conservation: 6 644 49 6 6643 3 46 4646699449 44 4 9 6 699 4
cuckoo 50 IENTAQALIIKRPFLPNRRSSYYKYT----LPPTLGRYRFCIYWFKANRTLRLV-YGKQS
chicken 49 IENTAQALVIRSPFLADRTTPSYQYS----LPTTLGRYRFCIYWFKANRTLWLA-YGKKS
zebrafish 59 ISANESTLSIQGRITAKCTQ---SSSIQLDSNPHQNSHFCVFWEPLLDLLIVEVNGKNH
bat 58 IENSEEALTVHAPFPGVPGA---SHF----FPDPRGLYHFCLSWNRHAGELHLL-YGKND
bovine 57 IRNSEEALTVHAPFPGVQSA---SWP----FPLPRGLYHFCLYWNRHAGKLHLR-YGKKD
elephant 57 IKNSEGALTVHAPFPAAGA---SRV----LPDRRGLYHFCLYWSRHTGKLHLR-YGKND
mouse 58 VWNTEETLTIRAPFLAADI---PRF----FPEPRGLYHFCLYWSRHTGRLHLR-YGKHD
rat 58 VWNTEESLTIRAPFPAADI---PYF----FPEPRGLYHFCLYWSRHTGRLHLR-YGKND
gorilla 58 IENSEEALTVHAPFPAARPA---SQS----FPDPRGLYHFCLYWNRHAGRLHLL-YGKHD
monkey 58 IENSEEALTVHAPFPAHPA---SRS----FPHPRGLYHFCLYWDRHAGRLHLL-YGKHD
chimp 58 IENSEEALTVHAPFPAHPA---SQS----FPDPRGLYHFCLYWNRHAGRLHLL-YGKRD
human 58 IENSEEALTVHAPFPAHPA---SRS----FPDPRGLYHFCLYWNRHAGRLHLL-YGKRD
consensus 61 ienseeaLtihapfpaar a s s fpdprglyHFClyW rhagrLhlv yGKkd
ss βββ βββββ ββββββββ ββββββ β
domain PPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP

```

```

Conservation: 4464 44 444 444 4 4444436666 4 3 49396 66 4364 46
cuckoo 105 FLLGRDPSNSITQGKESQKERTSTSIFNVS-YVLKGGKNTSLASASEYSFSLQG-----
chicken 104 FFLGHDPADGIARGLEK-----TKASILNVS-YVFKGQKNTSLESASEYLFPVPP-----
zebrafish 116 TLCKPNGLQGTCTCTDLSQGVQ-----DNAHMYGIVNGSVKGD--IITGDLKGNYIF
bat 110 FLLSDQASGLLCFQQEPILA-QGPRMLATSVSSWWRPQNTSLPSAAGFTFSFHGSGEP
bovine 109 FVLSDQALDLLCFRHQEETLV-PGPPLFATSVSSWSPQNTSLPSAASFIFSFHN---PP
elephant 109 FLLSDQASGLLCFRHQEESLV-QGPPLATSVSSWSPQNTSLPGATSFTFSFHK---PP
mouse 110 YLSSQASRLLCFQKQESLK-QGAPLIATSVSSWQIPQNTSLPGAPSFIFSFHN---AP
rat 110 YLSSRASRLLCFRKQEESLK-QGAPLVATSVSSWQSPQNTSLPGAPSFIFSFHN---AP
gorilla 110 FLLSDKASLLCFQHQEESLA-QGPPLATSVTSWSPQNTSLPSAASFTFSFHS---PP
monkey 110 FLLSDQASLLCFQHQEESLA-QGPPLFATSVTSWSPQNTSLPSAASFTFSFHS---PP
chimp 110 FLLSDKASLLCFQHQEESLA-QGPPLATSVTSWSPQNTSLPSAASFTFSFHS---PP
human 110 FLLSDKASLLCFQHQEESLA-QGPPLATSVTSWSPQNTSLPSAASFTFSFHS---PP
consensus 121 fllsdnas llcfqhqeesl qgpplatsvsswspqNtSlpsa sftfsfh pp
ss ββββ ββββ ββββ βββββ
domain PPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP111111

```



```

Conservation: 64444666 6 4
cuckoo      669 VKLQP--NSRESHLG
chicken     655 VKLQP--NSSQSHPG
zebrafish   637 SKQHM---LQTNEKS
bat         676 TRLPITSGSTSSSCI
bovine      672 ARLPISSGSTSSSRI
elephant    673 ARIPINSGSTSSGRI
mouse       673 AKLPISSGSTSSSRI
rat         673 AKLPISSGSTSSSRI
gorilla     673 ARLPISSGSTSSSRI
monkey      673 ARLPISTGSTSSSRI
chimp       673 ARLPISSGSTSSSRI
human       679 ARLPISSGSTSSSRI
consensus   721 arlpissgstsssri
domain      IIIIIIIIIIIIIIIII
    
```

protein:	Uniprot ID:
cuckoo: cuckoo GPR56	(A0A091FQY3)
chicken: chicken GPR56	(E1C0Q2)
zebrafish: zebrafish GPR56	(F1QZM9)
bat: <i>Myotis lucifugus</i> GPR56	(G1PD76)
bovine: bovine GPR56	(A4IF70)
elephant: elephant GPR56	(G3T7E2)
mouse: mouse GPR56	(Q8K209)
rat: rat GPR56	(Q8K3V3)
gorilla: gorilla GPR56	(Q50DM6)
monkey: Rhesus macaque GPR56	(Q50DM8)
chimp: chimpanzee GPR56	(Q50DM7)
human: human GPR56	(Q9Y653)

Alignment S2 (p. 1 of 2). Multiple sequence alignment of PTX and LNS domains. Canonical Pentraxin (PTX) and Laminin/Neurexin/Sex hormone-binding globulin (LNS) domains were identified including all PTX and LNS domains annotated to be present in all aGPCR ECRs and aligned with the PLL domain from GPR56/ADGRG1. Bolded sequences have well-defined PTX domain features (including an intradomain disulfide bond, cysteine residues colored red, and the conserved PTX motif, highlighted yellow). GPR133/ADGRD1 and GPR98/ADGRV1 are annotated to have a PTX domain, though they lack the intradomain disulfide bond and therefore do not have a complete PTX motif, highlighted cyan.

```

hCRP          47 PL-KAFTVCLHFYTE-LS-S----TRGYSIFS-YATKRQDNEI-LIF-WS---KDIGYS-
hSAP          48 PL-QNFTLCFRAYSD-LS-----RAYSLFS-YNTQGRDNEL-LVY-KE---RVGEYS-
mGPR56        28 P-REDFRFCGORN---Q-----TQOSTLHY---DOSSEPHI-FV--WN---TEETLT-
hGPR126       178 PELSAFTLCFEATKV-GH-E----DSDWTAFS-YSNASF-TQL-LSF-GK---AKSGYF-
hGPR112       50 PELSRFTACIDLVFM-DD-N----SRYWMAFS-YITNN----ALLGR-ED---IDLGLAG
hGPR144       138 PELAALTACTHVQWD-CA-S----PDPAALFS-VAAPALPNAL-QLRAFAEPGGVVRAA-
hGPR133       97 PEQCGPEGVTFSFFWKTQ-G---EQSRPIPSAYGGQVISNGF-KV--CS-SGGRGSVE-
hGPR98       1346 PSR--NNTIANFTFS-AWVMPNANTNGFIIAKDDGNGSIYYGV-KI--QT---NESHVT-
hSHBG         65 KI---TKTSSSFEVR-TW-----DPEGVIFYG-DTN-PKDDWFMLGL-RD---GRPEIQ-
bNRX1a        316 PI-QSSSDEITLSFK-TL-----QRNGLMLH--TGK-S-ADYVNLAL-KN---GAVSLV-
mNRX3b        103 PS--TRSDRLAVGFS-TT-----VKDGILVRI-DSAPGLGDFLQLHI-EQ---GKIGVV-
mLAMa2        2953 K--VGLDLLVEFEFR-TT-R----PTGVLLG--ISSQK-MDGMGIEM-ID---EKLMFH-

```

```

hCRP          93 --FT--VG--GSEIL--FEVP-EV-TV-----APVHICTSWESASGIVEFWVDGKP
hSAP          92 --LY--IG--RHKVT--SKVI-EK-FP-----APVHICVSWESSSGIAEFWINGTP
mGPR56        67 --I-----R----APFLAAPDIPRF--PEPRGLYHFCLYWSRHTGRLHLRYGKHD
hGPR126       224 --LS--IS--DSKCL--LNNALPVKEDI--FA-ESFEQLCLVWNNSLGSIGVNFKRNY
hGPR112       95 DHQQ--LI--L--YR-LGKTF--SIRH--HLAS--FQWHTICLIWDGVKGKLELFLNKER
hGPR144       189 --LV--VR--GQHAP--FLAA-FR-AD-----GRWHHVCATWEQRGGRWALFSDGRR
hGPR133       147 --LY--TR--DNSMT--WEAS-FS-PPG-----PYWTHVLFTWKSKEG-LKVYVNGTL
hGPR98       1396 --LSLHYKTLGSN---ATYIAKTTVMKYL--EE-SVWLHLLII--LEDGIIEFYLDGNA
hSHBG         109 --LH--NH--WAQLT--VG---AGPRL--D--D--GRWHQVEVK--MEGDSVLLEVDGEE
bNRX1a        360 --IN--LG--SGAFE--ALVEPVNGKF--N--D--NAWHDVKVT--RNLRQVTISVDGIL
mNRX3b        150 --FN--IG--TVDIS--IKE--ERTPV--N--D--GKYHVVRFT--RNGANATLQVDNWP
mLAMa2        2997 --VD--NG--AGRFTAIYDAE--IPGHM--C--N--GQWHKVTAK--KIKNRLELVDGNQ

```

```

hCRP          134 -RV-R--KSL-K-KG--YTVGAEASIILG--Q-EQ-----SQSLV-GDIGNVNM
hSAP          133 -LV-K--KGL-R-QG--YFVEAQPKIVLG--Q-EQ-DSYGG-KFDRSQSFV-GEIGDLYM
mGPR56        110 -YL-LSSQAS-R-LL--CFQKQEQSLKQGAPLIAT-SVSSW-QIPQNTSLP-GAPSFI--
hGPR126       96 -ET-VPCD-S-T-IS--KVIPGNGKLLG--S-----NQNEIVSLK-GDIYNFRL
hGPR112       142 -IL-E---VTD--QP--HNLTPHGTLFLG--H-FLKNESE-VKSMMRSFP-GSLYYFQL
hGPR144       94 -RAGA--RGL-G-AG--HPVPSGGILVLG--Q-DQ-DSLGG-GFSVRHALS-GNLTDFHL
hGPR133       189 STSDP--SGK-V-SR--DYGESNVNLVIG--S-EQ-DQA-----KCYENGAFDEFII
hGPR98       1445 -MP-RGIKSL-K-GE--AITDGPGILRIG--A-----GINGNDRFT-GLMQDVRS
hSHBG         150 -VL-R--LRQVS-GPLTSKRHPIMRIALG--G-LLFPASNL-RLPLVPALD-GLLRRDSW
bNRX1a        404 -TT-T--GYTQE-DY--TMLGSDDFFYVG--G-SPSTADLP-GSPSVNNFM-GLKEVVY
mNRX3b        192 -VN-E--HYPTGRQL--TIFNTQAQIAIG--G-KDKG-RLFQQLSGLYYD-GLKVLNMA
mLAMa2        3042 -VD-A--QSPNS-AS--TSADTNDPVFVG--G-FPGGLNQF-GLTTNIRFR-GCIRSLKL

```

```

hCRP          180 WDFVLSPDEINTIYLGGPFSP-NVL-NWRALKYEVQGE--VFTKPQLWP
hSAP          179 WSVLPPENILSAYQTPLPA-NIL-DWQALNYEIRGY--VIIKPLVWV
mGPR56        159 FSFHNAPHKVSHNASV-----D
hGPR126       313 WNFTMAKIL-SNLSC-NVKGNVV--DWQNDFWNIPNL--ALKAESNLS
hGPR112       188 WDHILENEEFMKCLD-----G-NIV-SWEEDVWLVNKI-----I
hGPR144       278 WARALSPAQLHRARACAPPSE-GLLFRWDPGALD-----V
hGPR133       231 WERALTPDE-----I
hGPR98       1486 YERKLTLEE-----I
hSHBG         200 LDKQAEIS---ASAP-----TSLRS-----C
bNRX1a        137 KNNDVRLELSRLAKQGDP-----KMKIHG-----V
mNRX3b        241 AENNPN-----IKINGS-----VRLV
mLAMa2        3090 TKGTGKPLEVNFAKA-----LELRGVQPVS---CPTT

```

hCRP: human C-reactive_protein (P02741)
hSAPC: human Serum amyloid p-component (P02743)
mGPR56: mouse GPR56 (Q8K209)
hGPR126: human GPR126 (Q86SQ4)
hGPR112: human GPR112(Q8IZF6)
hGPR144: human GPR144(Q7Z7M1)
hGPR133: human GPR133 (Q6QNK2)
hGPR98: human GPR98 (Q9WYG9)
hSHBG: human sex hormone-binding globulin (P04278)
bNRX1a: bovine neurexin-1-alpha, isoform 9a (Q28146-9)
mNRX3b: mouse neurexin-3-beta, without splice insert 4 (Q8C985)
mLAMA2: mouse laminin subunit alpha-2 (Q60675)

Alignment S2 (p. 2 of 2)

Alignment S3. Multiple sequence alignment of PTX domains. Canonical pentraxin (PTX) domains were identified including all the well-defined PTX domains present in all aGPCR ECRs and aligned to the PLL domain from GPR56/ADGRG1. Conservation score is calculated for positions with at least some detectable conservation (9 is most conserved). PTX domain features are illustrated (including an intradomain disulfide bond, cysteine residues colored red, and the conserved PTX motif, highlighted yellow). The conserved PTX motif is defined as H Φ C*xxWxxxxG, where Φ is a hydrophobic residue and the C* participates in the intradomain disulfide bond.

```

Conservation: 9 7 77 9          4          777          44 4          4
hCRP        47 P-LKAFTVCLHFYTELSSTRGYSIFSyatKRQDNEILI--FWSKDIGYSFTV--G-GSE-
hSAPC       48 P-LQNFTLCFRAYSdLS--RAYSLFSYNTQGRDNELLV--YKERVGEYSlyI--G-RHK-
mGPR56      28 P-REDFRCGQR-----NQTQQSTLHYDQSSEPHIF----VWNTTEETLTIRAPFLAAPD-
hGPR126    178 PELSAFTLCFEATKVGHEDSDWTAFSYSNASFTQLL-S--FGKAKSGYFLSI---SDSK-
hGPR112    50 PELSRFTCIDLVMDDNSRYWMAFSYITN---NAL----LGREDIDLGLAG--D-H-Q-
hGPR144    138 PELAALTACTHVQWDCASPDPAAALFSVAAPALPNALQLRAFAEPGGVVRAALVVR-GQHA
consensus    1 Pel aftvC h e r ysifsy t n l i fg d ysl v g k

```

```

Conservation: 4 4          4 9 9 9 4 4 3 4 4
hCRP       100 IL-F-EVP----EV-----TV-APVHICTSWESASGIVEFWVDGK-PRVRKSLKKGy
hSAPC      99 VT-S-KVI----EK-----FP-APVHICVSWESSSGIAEFWINGT-PLVKKGLRQGY
mGPR56     77 IP-----RFFPE-----P-RGLYHFCLYWSRHTGRLHLRYGKH-----DY
hGPR126   231 CL-L-NNALPVKEKEDI-----FAESFEQLCLVWNNLSLSIGVNFKRNYETVPCDSTISK
hGPR112   98 QLILYRLG---KTFSIRHHLASF-QWHTICLIWdGVKGLFLNKE--RILEVTDQPH
hGPR144   197 PF-L-AAF----RA-----DG-RWHHVCATWEQRGRWALFSDGRRRAGARGLGAGH
consensus    61 il l v e af hiClSWe sGrvelwv gk v k l gy

```

```

Conservation: 4 4 4 97 4 4 6 44 7 43 4 49 4 4 4 4 6
hCRP      144 TVGAEASIIlGQEQ-----S $\Phi$ SLVGDIGNVMWDFVLSPEINTIYLGgPFSP
hSAPC     143 FVEAQPKIVLGQEQ---DSYGGKFDRS $\Phi$ SFVGEIGDLYMWDsvLPPENILSAYQGTPLPA
mGPR56    111 LLSSQASRLLCFQKQ--EQ-----SLKQGAPlATSvSSWQIPQn-----TSLPGAP
hGPR126   284 VIPNGKLLlGSNq---NE-----IVSLKGDINfRLWNfTMNAKILSN--LSCNVKG
hGPR112   151 NLTPhGTLfLGHfLKNESSEV $\Phi$ K--SMR $\Phi$ SfPGSLYfQLWdHILENEEFMK----C-LDG
hGPR144   242 PVPSSGILVLGQDQ---DSLGGGfSVRHAlSGNLTDfHLWARAlSPAQLHRARACAPPSE
consensus    121 v aggsliLgqeq ds sl qsl gdi f lWd vl pd i pl a

```

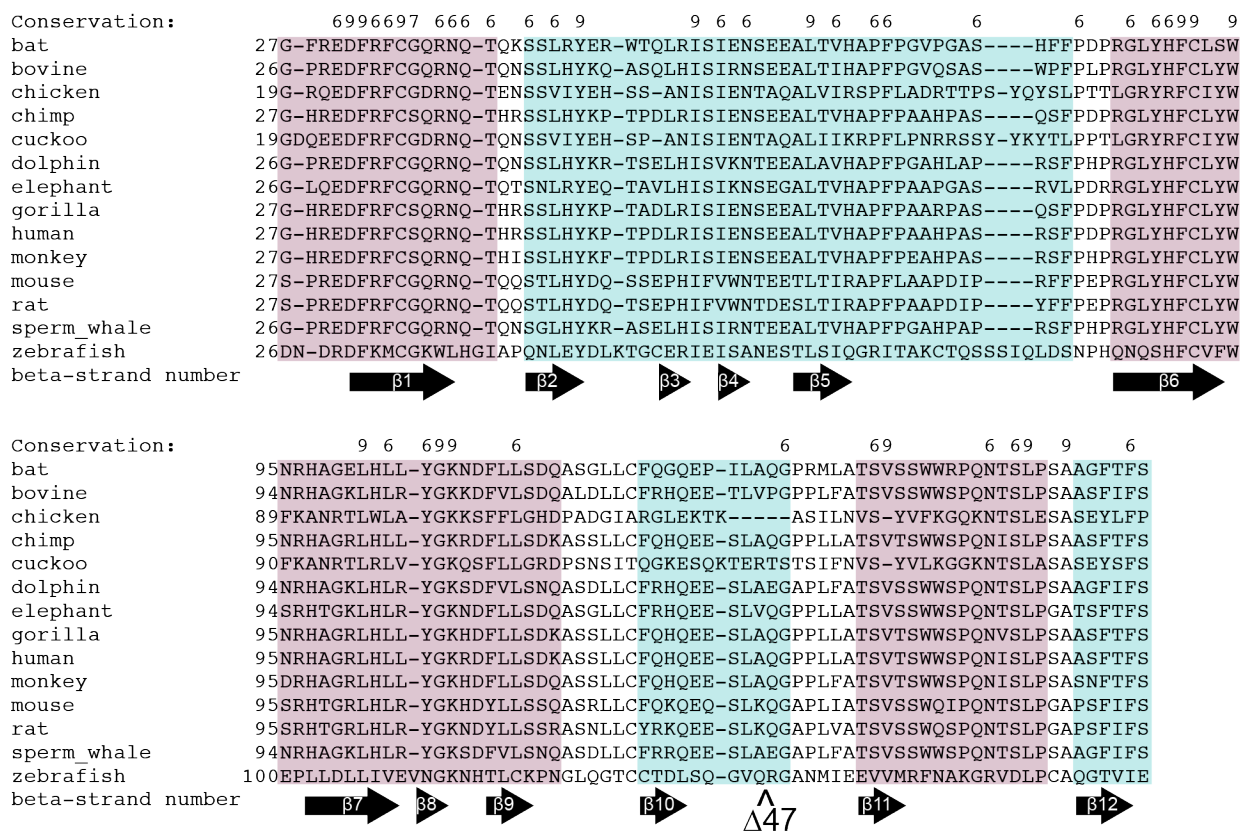
```

Conservation: 4 4 7 6 6
hCRP      201 NVL-NWRALKYEVQGEVfTKPQLWP
hSAPC     200 NIL-DWQALNYEIRGYVlIKPLVWV
mGPR56    156 SFIFSFHNAPhKVSHNASV-----D
hGPR126   331 NVV-DWQNDfWNIPNLALKAESNLS
hGPR112   204 NIV-SWEEDVWLVNKI-----I
hGPR144   299 GLLFRWDPGALD-----V
consensus    181 nvl w yev v

```

hCRP: human C-reactive_protein (P02741)
hSAPC: human Serum amyloid p-component (P02743)
mGPR56: mouse GPR56 (Q8K209)
hGPR126: human GPR126 (Q86SQ4)
hGPR112: human GPR112(Q8IZF6)
hGPR144: human GPR144(Q7Z7M1)

Alignment S4. Multiple sequence alignment of the GPR56 PLL domain from 14 species. Conservation score is calculated for positions with at least some detectable conservation (9 is most conserved). The β -strands were identified and numbered, and the β -sheets were colored based on the crystal structure (β -sheet A, the more divergent from PTX and LNS domains is colored cyan; β -sheet B, the more conserved with PTX and LNS domains is colored maroon [see Figure 2]). A 47-residue insertion in zebrafish Gpr56 was removed from this alignment for clarity ($\Delta 47$).



Supplementary Experimental Procedures

Cloning and purification of aGPCR extracellular fragments from insect cells

Extracellular regions (ECRs) of the following aGPCRs were cloned into pAcGP67a: mouse GPR56 (ADGRG1) full ECR (UniProt: Q8K209, residues S27-S392); mouse GPR56 GAIN domain (residues M176-S392); human GPR56 full ECR (UniProt: Q9Y653, residues G27-S392); zebrafish Gpr56 full ECR (UniProt: F1QZM9, residues T25-E358); human latrophilin 3 (LPHN3/ADGRL3) HormR+GAIN domains (UniProt: Q9HAR2, residues E496-S856); rat latrophilin 1 (LPHN1/ADGRL1) HormR+GAIN domains (UniProt: O88917, residues P460-I849); human brain angiogenesis inhibitor 3 (BAI3/ADGRB3) HormR+GAIN domains (UniProt: O60242, residues E498-E868); human GPR112 (ADGRG4) GAIN domain (UniProt: Q8IZF6), residues E2450-S2731). C-terminal 6xHIS tags were added for affinity purification. C-terminal AVI-tags corresponding to the sequence GLNDIFEAQKIEWHE were added to aid biotinylation.

A baculovirus expression system was used for expression of proteins in High Five insect cells as previously described (Arac et al., 2012). The secreted, glycosylated proteins were purified using nickel-nitrilotriacetic agarose resin (Qiagen) and size-exclusion chromatography (Superdex 200 10/300 GL; GE Healthcare).

Purification of monobodies from *E. coli*

The genes encoding the identified monobodies were cloned into an expression vector, pHBT (Sha et al., 2013). Monobodies were expressed in *E. coli* via autoinduction at 37°C for 20 hours. Monobodies were purified via an N-terminal 6xHIS tag using nickel-nitrilotriacetic agarose resin (Qiagen), and refolded on the Ni-column using the β -cyclodextrin method (Oganesyan et al., 2004). Refolded proteins were gel-filtered using a Superdex 200 10/300 GL column (GE Healthcare).

X-ray crystallography data processing, phasing, and refinement

Data were processed using HKL2000. To obtain a heavy metal substructure and calculate experimental phases, the CRANK2 (ccp4) software package (Ness et al., 2004) was used, despite the weak anomalous signal. The model generated from the anomalous data was used for molecular replacement (phaser.mr, ccp4) into the native dataset. Phenix.refine (PHENIX) was used for all refinement.

Dual luciferase SRE reporter plasmid construction

Dual luciferase SRE reporter plasmid was constructed to have constitutively expressed renilla luciferase and SRE-activated firefly luciferase on the same plasmid. The pmirGLO (Accession Number FJ376737) and pGL4.33 [luc2P SRE-Hygro] (Accession Number FJ773212) vectors were obtained from Promega. The SRE-firefly luciferase region of pGL4.33 was used to replace the PGK-firefly luciferase. The resulting reporter plasmid is referred to as dualLUC-SRE.

G protein signaling assay

HEK293T cells were seeded in 24-well plates (45,000 cells in 0.5 mL DMEM+10% FBS/well). After 12-18 hours, cells reached 40-50% confluence and were transfected with 11.3 ng *Gpr56* (WT or mutant) + 45.0 ng dualLuc-SRE + 0.23 μ L FUGENE6 per well from a master mix. After 24 hours, media was aspirated and replaced with DMEM + 0% FBS. For monobody treatment, monobody was added to cells 6.5 hours after the start of serum starvation. After 12 hours total of serum starvation, media was aspirated. Cells were lysed using the Dual-Glo[®] Luciferase Assay System from Promega and firefly and renilla luciferase signals were measured using a Synergy[™] Neo luminescence plate reader. Signaling intensity in RLU (fold increase) is reported as:

$$\frac{(\text{Firefly}_{\text{GPR56}}/\text{Renilla}_{\text{GPR56}})}{(\text{Firefly}_{\text{EV}}/(\text{Renilla}_{\text{EV}}))}$$

Flow cytometry

HEK293T cells were transiently transfected with FL or mutant mouse *GPR56* constructs (openbiosystems clone ID: 3709247) using Fugene6. After 48 hours, cells were detached and stained. Flow cytometry was performed using Guava[®] easyCyte as previously described (Lu et al., 2015).

Monobody $\alpha 5$ staining: To measure binding affinity, cells were stained primarily with biotinylated monomeric monobody at a range of concentrations, and secondarily with labeled neutravidin. To detect binding (as in Figure S3), cells were stained with 100 nM pre-tetramerized monobody on labeled neutravidin in order to increase avidity in a single staining reaction.

FLAG staining: Cells expressing N-terminally FLAG-tagged constructs were stained primarily with 1:1000 mouse anti-FLAG M2 (Sigma) and secondarily with 1:100 anti-mouse-FITC.

Streptavidin pull-down and western blot

Streptavidin pull-down and western blot were performed as previously described (Stoveken et al., 2015). Briefly, HEK293T cells were transiently transfected with GPR56 constructs using FUGENE6. After 48 hours, cells were treated with EZ-Link™ Sulfo-NHS-LC-Biotin (ThermoFisher), quenched, lysed, subject to pull-down with Streptavidin MagneSphere® Paramagnetic Particles (Promega), and subject to western blot using an antibody against the GPR56 CTF (Millipore Cat#: ABS1028, RRID: AB_2617058; 1:1000 dilution).

Zebrafish stocks and rearing conditions

Zebrafish (*Danio rerio*) were maintained in the Washington University Zebrafish Consortium facility (<http://zebrafish.wustl.edu/>), and all experiments were performed in compliance with Washington University's institutional animal protocols. WT (AB-Tubigen) embryos were collected from harem matings and reared at 28.5°C in egg water (5 mM NaCl, 0.17 mM KCl, 0.33 mM CaCl₂, 0.33 mM MgSO₄). Embryos were staged in hours post-fertilization (hpf) as previously described (Kimmel et al., 1995). To prevent pigmentation in embryos > 1 dpf, egg water was supplemented with phenylthiourea to 0.003%.

Whole mount *in situ* hybridization

Whole-mount *in situ* hybridizations were performed as described previously (Thisse and Thisse, 2008). Briefly, embryos were fixed at 65 hpf in 4% paraformaldehyde at 4°C overnight followed by dehydration in 100% methanol. Following dehydration, embryos were washed in 0.2% PBS-Tween (PBSTw), permeabilized in proteinase K (20mg/ul diluted 1:1000 in 0.2% PBSTw), and incubated with an *mbp* Digoxigenin-labeled riboprobe (Lyons et al., 2005) overnight at 65°C in hybridization buffer (50% formamide). Following overnight incubation, embryos were washed to remove formamide, blocked in 2% blocking medium supplemented with 10% normal sheep serum and 0.2% Triton™, and incubated in primary antibody (Anti-Dig, Fab fragments (1:2000), Product # 11214667001, Roche) overnight in block. Embryos were then washed in Maleic Acid Buffer with 0.2% Triton™, and developed by alkaline phosphatase treatment. After colorimetric development was complete, embryos were post-fixed in 4%

paraformaldehyde and stored long-term in 70% glycerol. Embryos were mounted on slides with glycerol and imaged at 10x with an AxioCam MRM on a light microscope (Zeiss AxioImager M2). Imaging was done in a blinded fashion such that the interpreter did not know whether the larvae were injected with WT or mutant mRNA. For further detail, please see previous studies which have employed similar *mbp* quantification protocols (Petersen et al., 2015, Ackerman et al., 2015).

Direct G protein-coupling assay using insect cell membranes

Protocol was implemented as previously described (Stoveken et al., 2015), the only difference being that for the present study, insect cell membranes were pre-incubated with G proteins for 5 minutes before starting the assay.

Matrix-assisted laser desorption time of flight (MALDI-TOF) mass spectrometry

Matrix was prepared by first preparing solvent: 50% acetonitrile + 50% (0.1% TFA in H₂O). 1mg sinapinic acid was dissolved in 0.1mL solvent to form saturated solution with precipitate. Samples for MALDI-TOF were prepared by spotting 0.5μL purified protein and 0.5μL matrix on the target plate. Samples were analyzed on a Bruker ultrafleXtreme MALDI-TOF/TOF instrument.

Circular Dichroism Spectroscopy

Purified mouse GPR56 ECR and GAIN domain were diluted to 200 nM and 250 nM, respectively in 100 mM NaCl, 25 mM sodium phosphate, pH 7.5. The sample was placed in a 1 cm path length cuvette and CD signal at $\lambda=220\pm 5$ nm was observed as the temperature was increased. All measurements were taken on a Jasco J-715 CD spectrometer.

Supplementary References

- Ackerman, S. D., Garcia, C., Piao, X., Gutmann, D. H. & Monk, K. R. 2015. The adhesion GPCR Gpr56 regulates oligodendrocyte development via interactions with Galpha12/13 and RhoA. *Nat Commun*, 6, 6122.
- Arac, D., Boucard, A. A., Bolliger, M. F., Nguyen, J., Soltis, S. M., Sudhof, T. C. & Brunger, A. T. 2012. A novel evolutionarily conserved domain of cell-adhesion GPCRs mediates autoproteolysis. *EMBO J*, 31, 1364-78.
- Kimmel, C. B., Ballard, W. W., Kimmel, S. R., Ullmann, B. & Schilling, T. F. 1995. Stages of embryonic development of the zebrafish. *Dev Dyn*, 203, 253-310.
- Lu, Y. C., Nazarko, O. V., Sando, R., 3rd, Salzman, G. S., Sudhof, T. C. & Arac, D. 2015. Structural Basis of Latrophilin-FLRT-UNC5 Interaction in Cell Adhesion. *Structure*, 23, 1678-91.
- Lyons, D. A., Pogoda, H. M., Voas, M. G., Woods, I. G., Diamond, B., Nix, R., Arana, N., Jacobs, J. & Talbot, W. S. 2005. *erbb3* and *erbb2* are essential for schwann cell migration and myelination in zebrafish. *Curr Biol*, 15, 513-24.
- Ness, S. R., De Graaff, R. A., Abrahams, J. P. & Pannu, N. S. 2004. CRANK: new methods for automated macromolecular crystal structure solution. *Structure*, 12, 1753-61.
- Oganesyan, N., Kim, R. & Kim, S.-H. 2004. One-column Chemical Refolding of Proteins. *PharmaGenomics*, 22-26.
- Petersen, S. C., Luo, R., Liebscher, I., Giera, S., Jeong, S. J., Mogha, A., Ghidinelli, M., Feltri, M. L., Schoneberg, T., Piao, X. & Monk, K. R. 2015. The adhesion GPCR GPR126 has distinct, domain-dependent functions in Schwann cell development mediated by interaction with laminin-211. *Neuron*, 85, 755-69.
- Sha, F., Gencer, E. B., Georgeon, S., Koide, A., Yasui, N., Koide, S. & Hantschel, O. 2013. Dissection of the BCR-ABL signaling network using highly specific antibody inhibitors to the SHP2 SH2 domains. *Proc Natl Acad Sci U S A*, 110, 14924-9.
- Stoveken, H. M., Hajduczuk, A. G., Xu, L. & Tall, G. G. 2015. Adhesion G protein-coupled receptors are activated by exposure of a cryptic tethered agonist. *Proc Natl Acad Sci U S A*, 112, 6194-9.
- Thisse, C. & Thisse, B. 2008. High-resolution in situ hybridization to whole-mount zebrafish embryos. *Nat Protoc*, 3, 59-69.