

ARTICLE

<https://doi.org/10.1038/s41467-019-14040-1>

OPEN

Structural basis for adhesion G protein-coupled receptor Gpr126 function

Katherine Leon^{1,2}, Rebecca L. Cunningham³, Joshua A. Riback^{1,4}, Ezra Feldman¹, Jingxian Li^{1,2}, Tobin R. Sosnick^{1,5}, Minglei Zhao¹, Kelly R. Monk^{3,6} & Demet Araċ^{1,2*}

Many drugs target the extracellular regions (ECRs) of cell-surface receptors. The large and alternatively-spliced ECRs of adhesion G protein-coupled receptors (aGPCRs) have key functions in diverse biological processes including neurodevelopment, embryogenesis, and tumorigenesis. However, their structures and mechanisms of action remain unclear, hampering drug development. The aGPCR Gpr126/Adgrg6 regulates Schwann cell myelination, ear canal formation, and heart development; and *GPR126* mutations cause myelination defects in human. Here, we determine the structure of the complete zebrafish Gpr126 ECR and reveal five domains including a previously unknown domain. Strikingly, the Gpr126 ECR adopts a closed conformation that is stabilized by an alternatively spliced linker and a conserved calcium-binding site. Alternative splicing regulates ECR conformation and receptor signaling, while mutagenesis of the calcium-binding site abolishes Gpr126 function in vivo. These results demonstrate that Gpr126 ECR utilizes a multi-faceted dynamic approach to regulate receptor function and provide relevant insights for ECR-targeted drug design.

¹Department of Biochemistry and Molecular Biology, The University of Chicago, Chicago, IL 60637, USA. ²Grossman Institute for Neuroscience, Quantitative Biology and Human Behavior, The University of Chicago, Chicago, IL 60637, USA. ³Department of Developmental Biology, Washington University School of Medicine, St. Louis, MO 63110, USA. ⁴Graduate Program in Biophysical Sciences Program, The University of Chicago, Chicago, IL 60637, USA. ⁵Institute for Biophysical Dynamics, The University of Chicago, Chicago, IL 60637, USA. ⁶Vollum Institute, Oregon Health & Science University, Portland, OR 97239, USA. *email: arac@uchicago.edu

Multicellular organisms rely on cellular communication to carry out critical biological processes, and numerous cell-surface receptors utilize their extracellular regions (ECRs) to modulate these cellular-adhesion and signaling events. For example, the ECRs of integrins, epidermal growth factor receptor (EGFR), and several G protein-coupled receptors (GPCRs) change conformation upon ligand binding, which propagates signals across the membrane^{1–9}. Targeting the essential ECRs of receptors with antibody-like drugs to trap the ECRs in distinct conformations, or to modulate ECR-ligand interactions has been an effective way to treat diseases caused by defective proteins. Currently, the anti-cancer drug cetuximab targets EGFR to prevent an activating extended ECR conformation¹⁰, and the drug etrolizumab blocks ligand binding to the ECRs of integrins in order to treat inflammatory bowel diseases¹¹. Remarkably, earlier this year, the migraine preventive drug erenumab, which blocks ligand binding to the ECR of calcitonin receptor-like receptor, became the first antibody drug against a GPCR to be approved by the Food and Drug Administration^{12,13}. Despite these and other breakthroughs, there are many essential receptors in the human genome that are not currently drugged, including the 32 adhesion GPCRs (aGPCRs), a diverse and understudied family of GPCRs with critical roles in synapse formation, angiogenesis, neutrophil activation, embryogenesis, and more^{14–16}.

Like all GPCRs, aGPCRs have canonical signaling seven-transmembrane (7TM) domains^{17,18}. However, unlike most other GPCRs, aGPCRs have large ECRs, which can extend up to almost 6000 amino acids (aa) and consist of various adhesion domains that mediate cell–cell and cell–matrix interactions¹⁹. In addition, during biosynthesis, aGPCRs are uniquely autoprolysed within a conserved GPCR Autoprolysis INDucing (GAIN) domain of the ECR that is juxtaposed to the 7TM²⁰, resulting in a fractured receptor that nevertheless remains tightly associated at the cell surface^{21,22}.

Although their protein architectures remain largely unknown, functional studies have shown that aGPCR ECRs can regulate receptor function and that antibody-like synthetic proteins that target the ECRs can modulate downstream signaling^{9,22–25}. A current model for aGPCR regulation suggests that transient interactions between the ECR and 7TM directly regulate receptor signaling^{9,22–25}. There are also numerous reports that aGPCRs use their ECRs to mediate functions in a 7TM-independent manner^{26–30}. Another non-mutually exclusive model for aGPCR activation posits that ligand binding to the ECR can exert force and cause dissociation at the autoprolysis site, revealing a tethered peptide agonist, which then activates the receptor^{31–34}. Clearly, the ECRs of aGPCRs have significant and diverse roles but remain poorly understood at a molecular level due to the scarcity of structural information, such as interdomain interactions, protein architecture, and identities of extracellular domains, which would provide insight into their mechanisms of action.

Gpr126/Adgrg6 is one of the better studied aGPCRs and is essential for Schwann cell (SC) myelination and other functions^{35–37}. In vertebrate peripheral nervous system (PNS) development, the myelin sheath surrounding axons is formed by SCs and functions to facilitate rapid propagation of action potentials³⁸. Disruption of myelination is associated with disorders such as Charcot-Marie-Tooth disease, which is characterized by muscle weakness^{39,40}. In *gpr126*-mutant zebrafish, SCs fail to express genes critical for myelination during development and are not able to myelinate axons due to deficient G-protein signaling. Additional studies have shown that this regulatory function of Gpr126 is conserved in mammals^{41,42} and that Gpr126 also plays a role in myelin maintenance through communication with the cellular prion protein⁴³. In humans, GPR126 mutations are linked to several cancers and other diseases^{44–48}, including adolescent

idiopathic scoliosis⁴⁹ and arthrogryposis multiplex congenita, a disorder characterized by multiple joint contractures⁵⁰. Furthermore, Gpr126 is required for inner ear development in zebrafish³⁵ and GPR126 is required for heart development in mouse³⁷, and it has been shown that the latter function is ECR-dependent and does not require the 7TM³⁰. While the biological significance of Gpr126 has become indisputable over recent years, the molecular mechanisms underlying Gpr126 functions remain unclear.

Gpr126 has a large ECR consisting of 839 aa. Prior to the current study, four domains in the ECR of Gpr126 had been identified through sequence-based bioinformatics: Complement C1r/C1s, Uegf, Bmp1 (CUB), Pentraxin (PTX), Hormone Receptor (HormR), and GAIN^{20,51,52}. However, a 150 aa region between PTX and HormR, could not be assigned to a known structural fold. Furin, a Golgi-localized protease, is reported to cleave human and mouse GPR126 in this region⁵¹, although any effect on protein architecture is unclear because of the unspecified structure. In addition, alternative splicing occurs in *Gpr126/GPR126*, resulting in Gpr126/GPR126 isoforms that vary in their ECRs^{51,53}. Alternative splicing of exon 6 was observed in human and zebrafish^{30,51}, producing isoforms that either include (S1 isoform, henceforth referred to as +ss) or exclude (S2 isoform, henceforth referred to as –ss) a 23 aa segment found within the unknown region between PTX and HormR. A genetic variant in *GPR126* leading to decreased inclusion of exon 6 was recently found to be associated with adolescent idiopathic scoliosis⁵⁴. Thus, determining the ECR structure, conformation, and other possible unexplored features will be instrumental in understanding Gpr126 function.

In this study, we determine the high-resolution crystal structure of the full-length ECR of zebrafish Gpr126, which reveals five domains, including a newly identified Sperm protein, Enterokinase and Agrin (SEA) domain, in which furin-mediated cleavage would occur in the human and mouse homologs. Intriguingly, the ECR is in an unexpected closed conformation that is reminiscent of the inactive closed conformation of the ECRs from EGFR and integrin families. This closed conformation is sustained by an alternatively spliced linker, while insertion of the alternatively spliced site gives rise to dynamic open-like ECR conformations and increases downstream signaling. A second feature that also mediates the closed conformation is a newly identified calcium-binding site at the tip of the ECR. Strikingly, zebrafish carrying point mutations at this site have both myelination defects and malformed ears, demonstrating the critical role of the ECR in Gpr126 function in vivo. These results altogether show that the ECR of Gpr126 has multifaceted roles in regulating receptor function, a feature that is likely true for other aGPCRs, and that will form the basis for further investigations in the efforts to drug aGPCRs.

Results

Structure of the full-length ECR of Gpr126. To determine the structure of the ECR of Gpr126, the full-length ECR (–ss) from zebrafish Gpr126 (T39–S837) was expressed and purified from insect cells using the baculovirus expression system. Zebrafish Gpr126 (Fig. 1a) has high sequence identity (47%) to its human homolog but its ECR has a fewer number of N-linked glycosylation sites (15 predicted in zebrafish, 26 in human) and no furin-cleavage site (Supplementary Fig. 1A), and thus yields a more homogeneous sample (Supplementary Fig. 1B, C). Crystals of both native and selenomethionine (SeMet)-labeled zebrafish Gpr126 ECR (–ss) were obtained and diffracted to 2.4 Å (Supplementary Fig. 1D), and the structure was determined by SeMet single-wavelength anomalous diffraction (SAD) phasing (Supplementary Table 1).

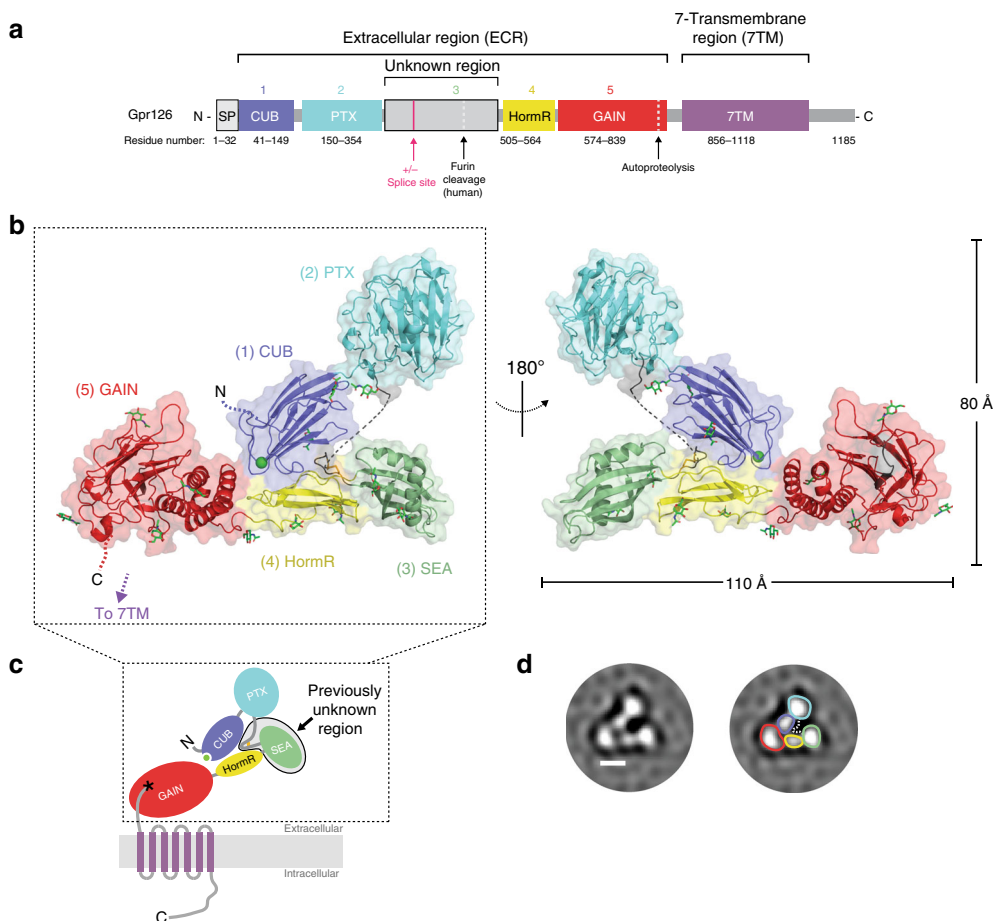


Fig. 1 Crystal structure of the full extracellular region of Gpr126. **a** Domain organization of Gpr126, indicating the ECR and 7TM regions. The unknown region includes a splice site. Cleavage sites (furin cleavage, autoproteolysis) are indicated by dashed lines. Domains are colored dark blue (CUB), cyan (PTX), gray (unknown region), yellow (HormR), red (GAIN), and purple (7TM). Domain boundaries are indicated below. SP indicates signal peptide. **b** Structure of the full ECR of (–ss) Gpr126. Domains are colored as in **a** except for the newly identified SEA domain (green). Domains are numbered (1–5) from N to C-terminus. Calcium ion in CUB domain is indicated as a green sphere. Dashed lines represent disordered residues. N-linked glycans are shown as green sticks. **c** Schematic of full-length Gpr126. The previously unknown region (SEA domain and linker region) is labeled. Autoproteolysis in GAIN domain is indicated by an asterisk and the last beta-strand of the GAIN domain is colored gray. **d** Representative negative-stain EM 2D class average of Gpr126 (–ss) ECR. Scale bar (white) represents 50 Å. Domains are assigned and colored according to color scheme noted above. The dashed line represents the linker region.

The structure, with overall dimensions of $110 \times 80 \times 35$ Å, revealed the presence of five domains (Fig. 1b, c), of which only four were identified previously. The N-terminal region of the protein is composed of the CUB domain followed very closely by the PTX domain. The 150 aa unknown region after the PTX domain was revealed to be a 22 aa linker that is partially disordered, the 23 aa alternatively spliced region (not present in crystal structure construct), and a structured domain which spans 105 aa and was identified as a SEA domain through the Dali server⁵⁵. The Gpr126 SEA domain adopts a ferredoxin-like alpha/beta sandwich fold, common to SEA domains from other proteins. Interestingly, analysis of the structure as well as sequence alignments between zebrafish and human showed that furin cleavage in humans would occur in the SEA domain (Supplementary Fig. 1A). Finally, the SEA domain is followed by the HormR and GAIN domains, the latter of which is autoproteolyzed as expected (Supplementary Fig. 1E). The HormR and GAIN domain structures are similar to previously-solved HormR + GAIN domain structures from other aGPCRs^{9,20}, with the exception of the relative orientation between HormR and GAIN. There is a 90° rotation of the

HormR domain with respect to the GAIN domain (Supplementary Fig. 1F) in Gpr126 compared to previously-solved HormR + GAIN structures from rLphn1 and hBAI3²⁰. In addition, Gpr126 was observed to have at least ten sites of glycosylation throughout all domains of the ECR except the PTX domain (Fig. 1b).

Gpr126 (–ss) ECR adopts a closed conformation. Unexpectedly, the structure revealed a compact, closed conformation where the most N-terminal CUB domain interacts with the more C-terminal HormR and GAIN domains (Fig. 1b). To ensure that this conformation is not a crystallization artifact, we utilized both negative-stain electron microscopy (EM) and small-angle X-ray scattering (SAXS) to confirm that the closed conformation is observed for Gpr126 in solution. Negative-stain 2D class averages of Gpr126 ECR showed a V-shaped protein architecture (Fig. 1d). The individual domains in the 2D class averages were assigned according to size and are consistent with the closed architecture of the crystal structure. In addition, we measured the radius of gyration (R_g) of the ECR using SAXS to confirm that the closed conformation exists in solution. The observed R_g (41.1 ± 0.1 Å) is

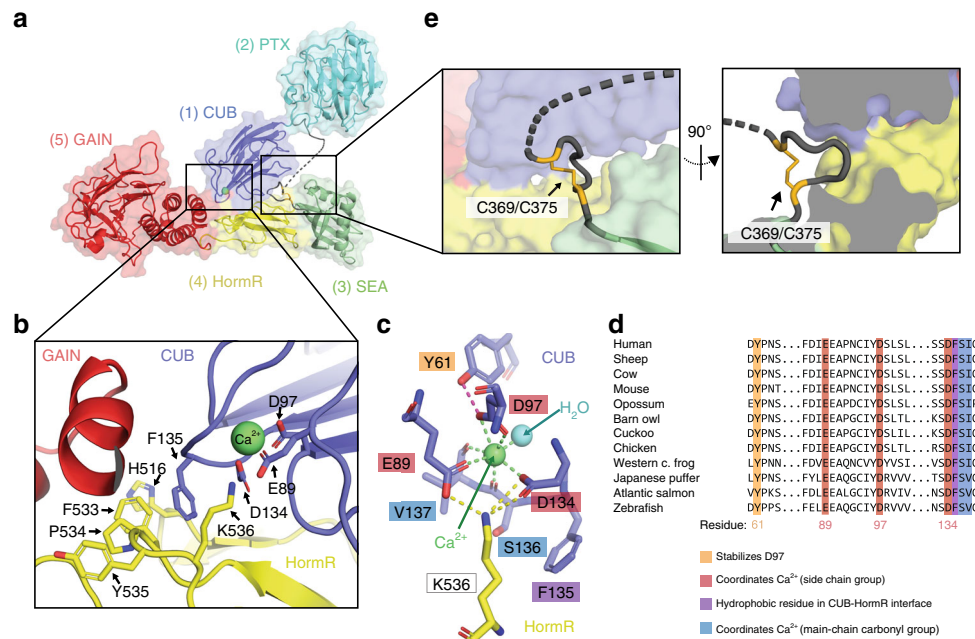


Fig. 2 Closed conformation of Gpr126 is mediated by CUB-HormR-linker interactions. **a** Structure of the full ECR of (–ss) Gpr126. **b** Close-up view of the CUB-HormR interface. Residues at the interface are shown as sticks. The calcium ion is shown as a bright green sphere. **c** Close-up view of the calcium-coordination site within CUB domain. The water molecule is shown as a blue sphere. The residues are shown as sticks. CUB residues are colored dark blue and HormR residue is colored yellow. Residue labels are colored according to their roles in CUB-HormR interaction: red (E89, D97, D134) represents calcium coordination by side-chain residue, blue (S136, V137) represents calcium coordination by main-chain carbonyl group, purple (F135) represents a hydrophobic residue in CUB-HormR interface, and orange (Y61) represents a residue that stabilizes calcium-coordinating residue D97. Calcium coordination is shown as bright green dashed lines. CUB-HormR interaction is shown as yellow dashed lines. The interaction between Y61 and D97 is shown as a magenta dashed line. **d** Sequence alignment of partial Gpr126 CUB domain from various species, highlighting important conserved residues: calcium-coordinating residues by side-chain group (red), calcium-coordinating residues by main-chain carbonyl (blue), a tyrosine residue that stabilizes a calcium-coordinating residue (orange), and a hydrophobic phenylalanine residue in the CUB-HormR interface (purple). **e** Close-up view of the disulfide-stabilized loop inserted between CUB and HormR domains. The disulfide bond is colored bright orange and is indicated by an arrow. The dashed line represents disordered residues in the linker region.

consistent with the calculated R_g of the closed-conformation crystal structure model (42.6 Å) and inconsistent with that of an extended model of Gpr126 ECR in which the CUB domain points away, rather than toward, the center of the molecule ($R_g = 52.2$ Å) (Supplementary Fig. 1G). Taken together, these results: show that Gpr126 ECR is in a closed conformation in solution, demonstrate that this conformation is not an artifact of crystal-packing contacts, and suggest that this closed conformation may play an important role in Gpr126 function.

As the closed conformation of Gpr126 (–ss) ECR was shown to exist both in solution and in the crystal lattice, we next wanted to explore the interactions that contribute to this protein architecture. Close examination of the crystal structure revealed two interaction sites that mediate the closed conformation, the first of which is a direct interaction between domains that are at opposite ends of the ECR and the second is an indirect interaction formed between two domains through a loop that holds them together (Fig. 2a).

First, a direct interaction exists at the tip of the CUB domain (close to the N-terminus), which points inward towards the center of the molecule and lies in the interface between GAIN and HormR. Residues in the HormR domain (H516, F533, P534, Y535) interact with each other through pi-pi stacking (sandwich), promoting interaction with F135 on the CUB domain through additional (T-shaped) pi-pi stacking to stabilize the CUB-HormR interaction (Fig. 2b).

Surprisingly, examination of the 2Fo-Fc electron density map showed that there is density within the CUB domain at this

interface that does not belong to any amino acid residue (Supplementary Fig. 2A). This density is coordinated by the side-chain groups of E89, D97 (bidentate) and D134, main-chain carbonyl groups of S136 and V137, as well as a water molecule for a complex with coordination number 7 in a pentagonal bipyramid geometry (Fig. 2c). The geometry and distances between the density and the coordinating residues in Gpr126 are consistent with calcium coordination⁵⁶. Several CUB domains from extracellular proteins are reported to coordinate calcium, including Gpr126⁵⁷, and some have been discovered to use this coordination to mediate ligand binding^{57–61} (Supplementary Fig. 2B). For example, the C1s protein uses its CUB calcium-binding site to bind to ligand C1q and initiate the classical pathway of complement activation⁶¹, and the Lujo virus recognizes a calcium-binding site on the CUB domain of the neurophilin-2 receptor in order to gain cell entry⁵⁹. The calcium-coordinating residues are all conserved in the Gpr126 CUB domain (among GPR126 proteins from various species (Fig. 2d) as well as among calcium-binding CUB domains from other proteins (Supplementary Fig. 2C)), suggesting that the density is indeed calcium. Importantly, the calcium coordination aligns the coordinating residues E89 and D134 on the surface of the CUB domain such that they can interact with K536 on the HormR domain (Fig. 2c), contributing to the closed conformation.

In addition to the direct CUB-HormR interaction, a second interaction site is formed by a disulfide-stabilized loop, which provides a bridge between the CUB and HormR domains. Although 13 (C355–A367) of the 22 aa (C355–P376) in the linker

region are disordered in the structure, the rest were able to be resolved and they form a small loop stabilized by a disulfide bond between C369 and C375 (Fig. 2e). This loop is located directly N-terminal to the SEA domain and is inserted between the CUB and HormR domains, effectively bridging the two domains and likely contributing to the stabilization of the closed conformation. The cysteines that form the disulfide bond are conserved among all except four of the 94 species analyzed in this study (Supplementary Fig. 2D and Supplementary Data 1), suggesting that this disulfide bond plays an important role in Gpr126 function. The five residues (ASGLG) flanked by the cysteines are small and flexible, accommodating the formation of the disulfide loop as well as insertion into the small pocket between CUB and HormR.

Alternative splicing modulates Gpr126 ECR conformation.

Gpr126 is alternatively spliced, producing several isoforms that may modulate protein function. Skipping of exon 6 results in deletion of 23 aa in zebrafish (28 aa in human) and is of particular interest because these amino acids reside in the previously unknown region of Gpr126 ECR. The 23 aa region is rich in serine/threonine residues (10 out of 23) and contains a predicted N-linked glycosylation site, which suggests that this region may be a highly O- and N-link glycosylated stalk. From analysis of the crystal structure (–ss isoform, in which the 23 aa are deleted), we determined that the splice site is directly between the regions encoding the disulfide-stabilized loop and the SEA domain (Fig. 3a). Because the disulfide-stabilized loop makes contacts that are important for the closed conformation of Gpr126 ECR (–ss) (Fig. 2e), we hypothesized that the (+ss) isoform would disrupt the closed conformation and have a different, more open conformation.

To test whether Gpr126 ECR (+ss) and (–ss) have different conformations, the two proteins were purified and analyzed using negative-stain EM. Single particles were classified into 2D class averages and the class averages were further categorized into groups to facilitate interpretation of different conformations. The class averages for the (–ss) isoform, categorized into five main orientations (Fig. 3b), were consistent with the closed conformation of the crystal structure (Fig. 1b). However, the class averages for the (+ss) isoform (Fig. 3c) showed a diverse population of ECR molecules, as they contain additional more open-like conformations (group vi, 21% of particles, Fig. 3d) as well as closed conformations that were observed in the (–ss) isoform (Fig. 3c, d). Furthermore, individual (+ss) particles showed the presence of open conformations (Fig. 3e), including a fully extended conformation, which could not be classified into a distinct class average during image processing. These results are consistent with our hypothesis that the (+ss) ECR conformation is different from that of (–ss) and suggest that the addition of 23 aa extends the linker in (+ss), likely disrupting the indirect and direct CUB-HormR interactions and preventing the stable closed conformation that is observed in (–ss) (Fig. 3f).

The negative-stain EM data are consistent with SAXS experiments showing that the R_g of zebrafish Gpr126 ECR (+ss) is larger than that of (–ss) with a more dramatic change in R_g observed between the human GPR126 isoforms (Supplementary Fig. 3A–F and Supplementary Table 2). Size-exclusion chromatography elution profiles for both zebrafish and human constructs also showed that (+ss) elutes earlier compared to (–ss), indicative of a larger size and different shape (Supplementary Fig. 3G, H).

Alternative splicing modulates Gpr126 receptor signaling. To determine whether the two isoforms also exhibit different levels of signaling, receptor activity was measured for both isoforms

using a G protein signaling assay. Human GPR126 has been shown previously to couple to and activate $G\alpha_s$, leading to production of cAMP⁴². Therefore, we used a cAMP signaling assay in which HEK293 cells were co-transfected with a full-length zebrafish Gpr126 construct and a reporter luciferase that emits light upon binding to cAMP. Cell-surface expression levels of the constructs were quantified by flow cytometry analysis of cells stained by antibodies against N-terminal FLAG-tags (Fig. 4a and Supplementary Fig. 4A, B, C), and basal signaling results (Fig. 4b) were normalized to expression level (Fig. 4c).

Cells transfected with either (–ss) or (+ss) Gpr126 had higher cAMP levels compared to cells transfected with an empty vector (EV) (Fig. 4c), demonstrating that basal activity of Gpr126 can be detected in this assay. As a positive control, a synthetic peptide agonist that targets the 7TM activated zebrafish Gpr126 and human GPR126 signaling to a level consistent with similar, previously-published experiments on human GPR126³² (Supplementary Fig. 4D) and did not activate signaling in EV-transfected cells.

However, the closed-conformation (–ss) Gpr126 signaled significantly less compared to the more dynamic (+ss) Gpr126 (Fig. 4c), and this result was consistent between both zebrafish and human constructs (Supplementary Fig. 4C and Fig. 4d–f). This suggests that the additional amino acids in the linker region of the ECR as a result of alternative splicing plays a role in modulating the activity of Gpr126 and that the ECR of Gpr126 is coupled to receptor signaling. Taken together with the negative-stain EM results, the (–ss) and (+ss) Gpr126 isoforms are distinct in terms of ECR conformation dynamics, as well as G protein signaling activity.

In addition, we mutated calcium-binding site residues D134A/F135A in the (–ss) isoform, which we predicted would disrupt the closed conformation. Using negative-stain EM, we observed open ECR conformations for this construct (Supplementary Fig. 4E–G), similar to the wild-type (+ss) isoform (Fig. 3c and Supplementary Fig. 4H). The calcium-binding site mutation did not increase or decrease the cAMP signaling for the (–ss) Gpr126 isoform, which suggests that the ECR conformation is not solely responsible for regulation of receptor signaling. However, the same mutation in the (+ss) isoform resulted in lower cAMP levels compared to wild-type (+ss) (Supplementary Fig. 4I). Cell-surface expression levels of these mutant Gpr126 constructs in HEK293 cells were similar or higher than wild-type constructs, excluding the possibility that lower signaling was due to improper protein folding or trafficking (Supplementary Fig. 4I). Altogether, these results might be explained by a complex, rather than a simple and straightforward, model of regulation for receptor signaling and suggest a possible functional role for the calcium-binding site.

Calcium-binding site is critical for PNS myelination in vivo.

Functional sites on proteins are usually highly evolutionarily conserved. We used the ConSurf server⁶² to perform surface conservation analyses on a diverse set of 94 Gpr126 protein sequences. The conservation score for each residue was mapped onto the Gpr126 ECR structure (Supplementary Fig. 5A), which revealed that the most conserved domain in the ECR is the CUB domain. Importantly, the calcium-binding site is absolutely the most highly conserved patch within the CUB domain and within the entire Gpr126 ECR (Fig. 5a). The calcium-binding site is universally conserved among all species analyzed, which suggests that the calcium-binding site has an essential role in Gpr126 function.

We next wanted to test whether the residues in the calcium-binding site are important for Gpr126 function in vivo. Gpr126 has previously been shown to regulate both PNS myelination and

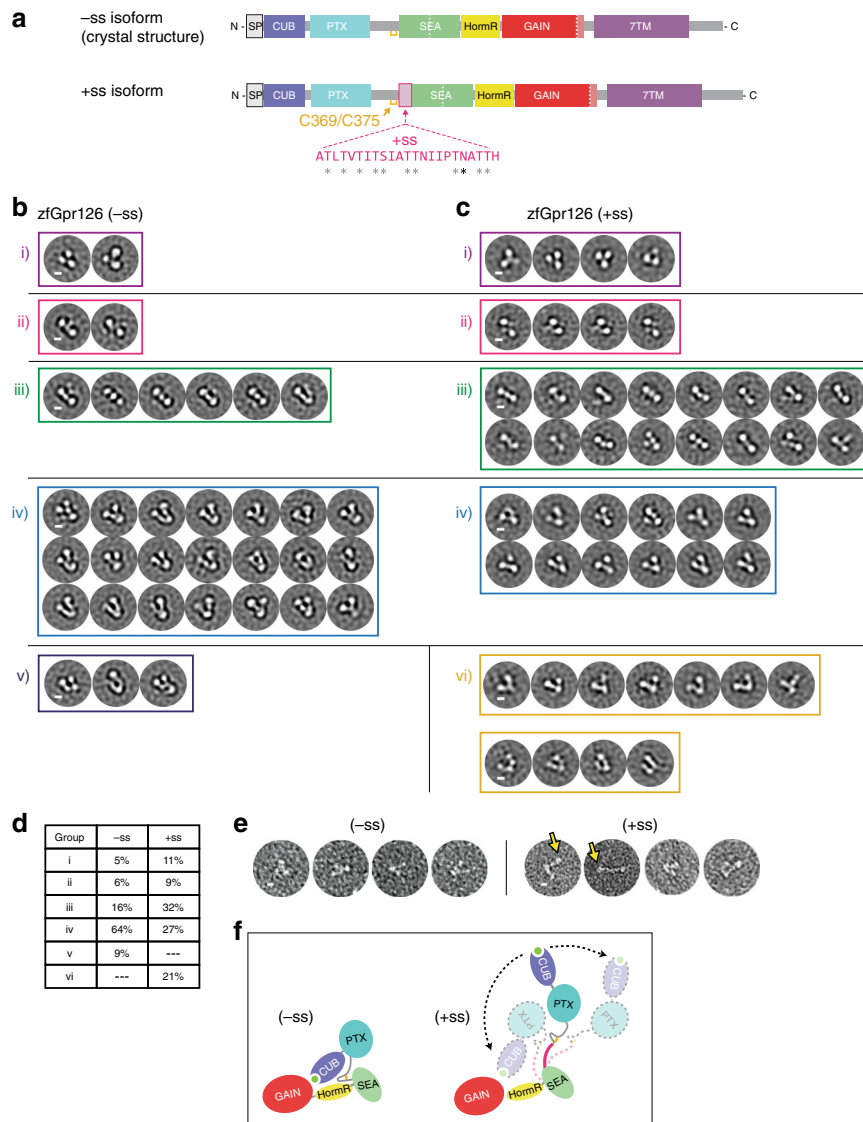


Fig. 3 Alternative splice isoforms of Gpr126 modulate ECR conformation. **a** Schematic diagram of Gpr126 splice isoforms generated by including (+ss) or excluding (-ss) exon 6. Residues encoded by exon 6 are colored magenta. Gray asterisks indicate potential O-linked glycosylation sites and the black asterisk indicates a predicted N-linked glycosylation site. The conserved disulfide bond in the linker is colored yellow. **b, c** Negative-stain EM 2D class averages for -ss (**b**) and +ss (**c**) ECR constructs. Class averages are categorized according to similar orientations: (i, ii, iii, iv, v, and vi). (i, ii, iii, iv) are observed in both -ss and +ss isoforms. (vi) represents open-like conformations ($>50^\circ$ angle) that are observed only in the +ss isoform. (v) represents unidentifiable miscellaneous views. Scale bars (white) represent 50 Å. **d** Quantification of percentage of particles per category for both isoforms. **e** Representative individual particles for both isoforms. Yellow arrows point to particles which are not in a closed conformation. **f** ECR conformations based on negative-stain EM are depicted as cartoons. The splice site is shown in magenta. Black arrows with dashed lines indicate dynamic ECR conformation.

ear development in zebrafish through elevation of cAMP^{35,36}. Zebrafish *gpr126* mutations that impair G protein signaling result in abolished myelination of the peripheral axons by SC and cause “puffy” ears^{28,32,35,36,63}. GPR126 has been shown to have a role in heart development in mouse³⁷, supported by additional studies in zebrafish^{28,30,63}. Gpr126 activity in zebrafish can be readily measured by analyzing the expression of *myelin basic protein (mbp)*, which encodes a major structural component of the myelin sheath and is essential for PNS myelination, and by assessing ear and heart morphologies of the fish. To determine whether the calcium-binding site is important for these functions, two amino acids in the site, D134 and F135, were targeted and mutated to alanines using CRISPR/Cas9-mediated homologous recombination. D134 directly coordinates the calcium ion and F135 is an adjacent hydrophobic residue which forms one arm of

the calcium-binding pocket (Figs. 2c, 5a). As a result, the mutant zebrafish, *gpr126^{stil464}*, harbor D134A and F135A mutations (Fig. 5b, Supplementary Fig. 5B). These mutations created a BstUI restriction enzyme site, which was used to genotype individual zebrafish (Fig. 5c). Expression of *gpr126* is unaffected in *gpr126^{stil464}* mutants (Supplementary Fig. 5C, D). Strikingly, compared to wild-type siblings, the *gpr126^{stil464}*-mutant zebrafish developed the puffy ears (Fig. 5d, e) that are indicative of a defect in Gpr126-mediated G protein signaling, though they do not appear to have heart defects (Supplementary Fig. 5E-T). In addition to the ear phenotype, mutant zebrafish did not express *mbp*, indicative of failed PNS myelination (Fig. 5f, g, Supplementary Fig. 5U, V). These results show that D134 and F135 in the calcium-binding pocket of Gpr126 are essential for ear and SC development in vivo.

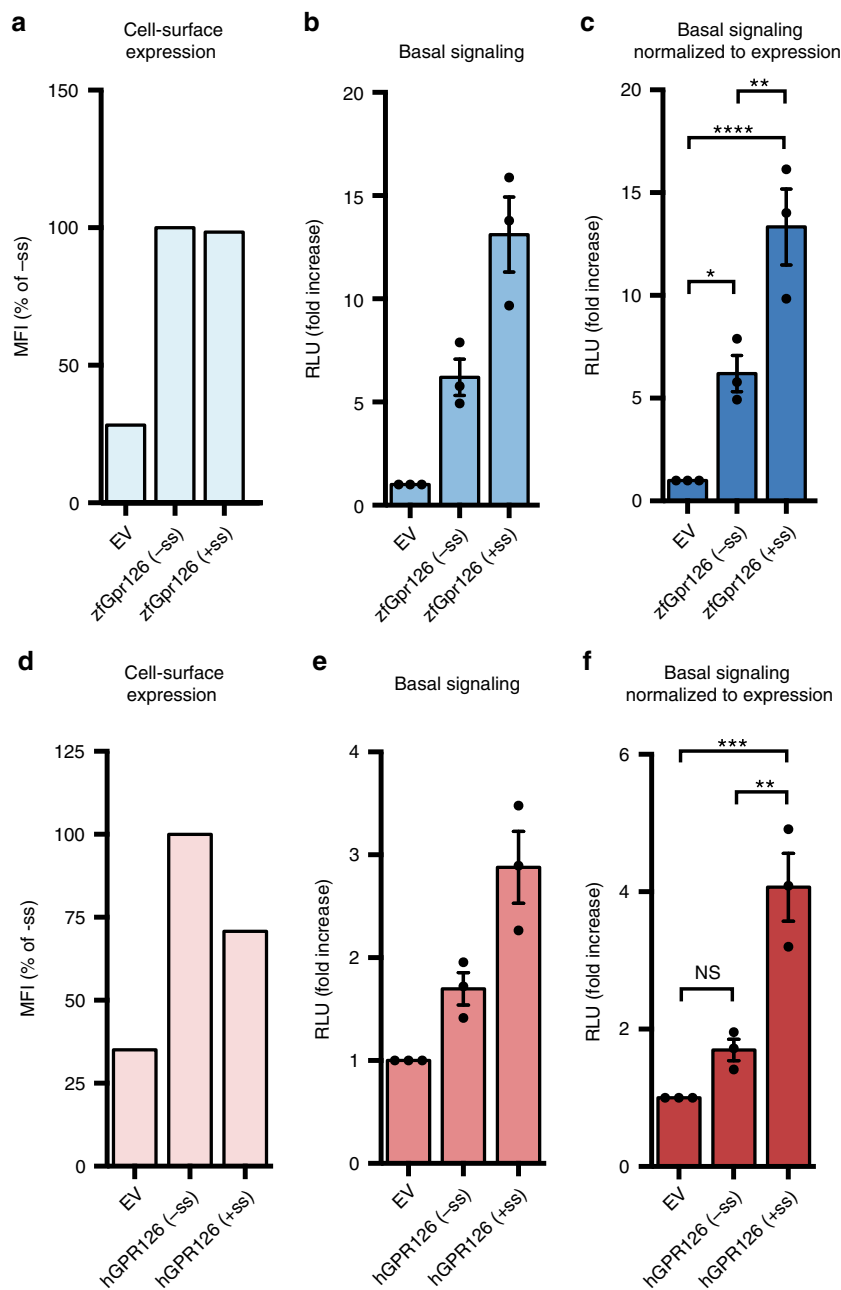


Fig. 4 Alternative splice isoforms of Gpr126 modulate receptor signaling. **a** Cell-surface expression levels for empty vector (EV), zebrafish Gpr126 splice isoforms, measured using flow cytometry to detect binding of anti-FLAG antibody to cells expressing FLAG-tagged Gpr126. The Gpr126 cell-surface expression levels are normalized to the control EV signal. Data are shown as MFI (mean fluorescence intensity). Error bars are not shown because expression levels are presented as median fluorescence intensities of 10,000 cells for each population of transfected cells, for a single flow cytometry experiment representative of at least three independent experiments. **b** Basal signaling measured by the cAMP signaling assay. Data are shown as fold increase over EV of RLU (relative luminescence units). **c** Basal cAMP signaling normalized to cell-surface expression. ns, $P > 0.05$; * $P \leq 0.05$; ** $P \leq 0.01$; *** $P \leq 0.001$; **** $P \leq 0.0001$; by one-way ANOVA and Tukey's multiple comparisons test. Data in **b** and **c** are presented as mean \pm SEM, $n = 3$, and are representative of at least three independent experiments. **d-f** Same as **a-c** but for human GPR126 splice isoforms. Source data are provided as a Source Data file.

Identification of a proteolytic SEA domain in human GPR126.

As mentioned earlier, the previously unknown region in the Gpr126 ECR contains a structured domain, which we revealed to be a SEA domain (Fig. 6a). Gpr126 SEA superimposes well over known SEA domains from Mucin-1 and Notch-2^{64,65}, which are cleaved (via autoproteolysis and furin, respectively), both in the same loop between beta-strand 2 and beta-strand 3 (Fig. 6b, c). Although the GPR126 furin-cleavage site is conserved in many

mammals and birds (Supplementary Data 1, Supplementary Fig. 1A), with a consensus sequence of (R/K)-X-K-R], it is not conserved in zebrafish Gpr126. Using sequence alignments (Fig. 6d) and homology modeling, we mapped the furin-cleavage site in human GPR126 (Fig. 6e, Supplementary Fig. 6A) to the same loop that is cleaved in Mucin-1 and Notch-2, suggesting that SEA domain cleavage plays similar roles in each of these proteins. Consistent with a previous study⁵¹, R468A mutations

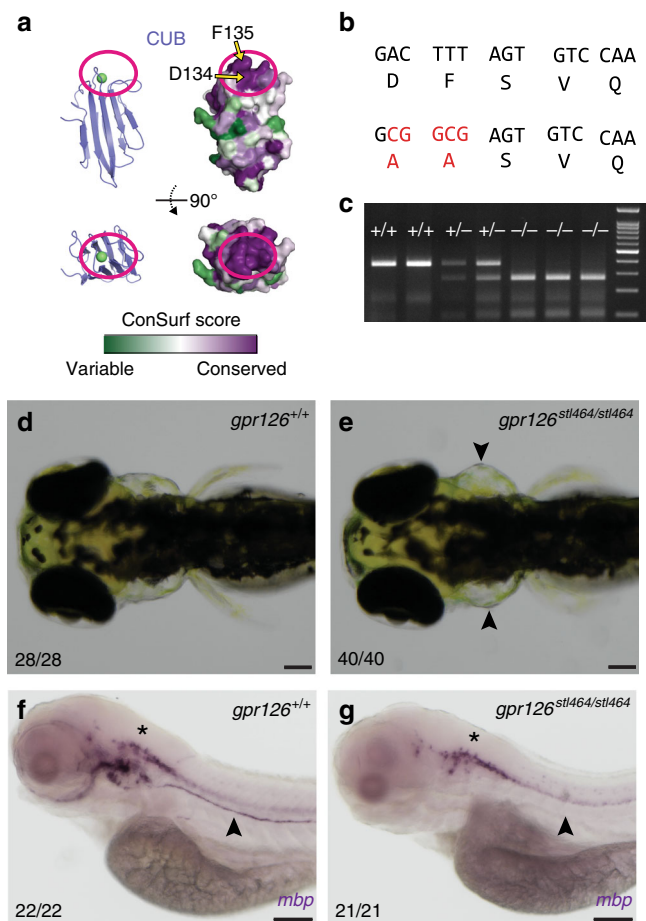


Fig. 5 The calcium-binding site is required for Gpr126 function in vivo. **a** Surface conservation analysis (green, variable; purple, conserved) of CUB domain. The calcium-binding site is circled in magenta. D134 and F135 are indicated by arrows. **b** D134 and F135 were both mutated to alanines through homologous recombination of a 150 bp ssODN containing a 5 bp mutation (red nucleotides). **c** Genotyping assay for the *gpr126^{stl464}* lesion. The 5 bp mutation introduces a BstUI restriction enzyme binding site. **d** 4 dpf wild-type larva compared to **e** 4 dpf *gpr126^{stl464/stl464}* larva with puffy ears (arrowheads). Scale bars (black) represent 100 μ m. **f** 4 dpf wild-type larvae express *mbp* throughout the posterior lateral line nerve (PLLn, arrowhead), whereas **g** 4 dpf *gpr126^{stl464/stl464}* larva lack *mbp* expression along the PLLn (arrowhead). Scale bars (black) represent 100 μ m. Asterisks indicate CNS.

abolish furin cleavage in both human GPR126 (–ss) and (+ss) isoforms (Supplementary Fig. 6B, C). In addition, these mutant GPR126 constructs were transfected into HEK293 cells and were detected on the cell surface (Supplementary Fig. 6D), and therefore, the importance of furin cleavage is likely not primarily important for proper expression and trafficking.

To our knowledge, SEA and GAIN are the only known protein domains that are proteolyzed and remain associated even after proteolysis. In proteins like Mucins and Notch, the cleaved SEA domain remains intact^{65,66} and shear forces likely unfold the domain and separate the protein into two fragments^{67,68}. The Gpr126 SEA domain shows several noncovalent interdomain interactions, particularly between all four of the beta-strands that form a beta-sheet (Fig. 6f). The separation of the human GPR126 furin-cleaved SEA domain into two fragments does not readily occur immediately following cleavage as the cleaved protein resists separation when purified by size-exclusion

chromatography (Supplementary Fig. 6B), similar to the aforementioned SEA domains as well as to GAIN domain autoproteolysis. Instead, the two fragments likely stay associated noncovalently until a disruptive event, such as ligand binding and mechanical force, unfolds the SEA domain and leads to separation or shedding of the region N-terminal to the furin-cleavage site (CUB, PTX, linker, half of SEA) and the C-terminal region (half of SEA, HormR, GAIN, 7TM).

Discussion

aGPCRs make up the second largest family of GPCRs with 32 members in humans and are essential for numerous biological processes such as synapse formation, cortex development, neutrophil activation, angiogenesis, embryogenesis, and many more. Recent studies have shown that the ECRs of aGPCRs play important roles in these functions; however, the relative lack of information about the structures of ECRs and their mechanisms of activation hampers further studies toward drugging these receptors. Here we show that the large ECR of Gpr126, an aGPCR with critical functions in PNS myelination, ear development, and heart development, adopts an unexpected closed conformation where the most N-terminal CUB domain interacts with the more C-terminal HormR domain. The structure of the Gpr126 ECR revealed that the closed conformation is mediated through a calcium-binding site as well as a disulfide-stabilized loop. Interestingly, the residues involved in these intramolecular interactions are highly conserved among Gpr126 sequences, including that of zebrafish, raising questions about their role in Gpr126 function.

Alternative splicing is a mechanism to increase the functional diversity of metazoan genomes and has been repeatedly demonstrated to play a role in the regulation of brain function. For example, alternative splicing contributes to the functional diversification of DSCAMs, protocadherins, calcium channels, neuroligins, and neuroligins^{69–72}. It is also proposed that alternative splicing may cause a large conformational change in the ECR of the synaptic protein teneurin, since alternative splicing allows the protein to act as a switch in regulating ligand binding despite the ligand-binding site being distant from the seven aa alternatively spliced site⁷³. Since *gpr126* is alternatively spliced in the region encoding the ECR, we examined the functional differences between isoforms. Our negative-stain EM and SAXS results suggest that alternative splicing between the regions encoding the PTX and SEA domains in *gpr126* perturbs the closed conformation and generates a population of ECR conformations that range from closed to extended (Fig. 7a). Several of the inserted residues resulting from alternative splicing are predicted to be sites of glycosylation. These glycosylation sites as well as the state of the other glycosylation sites may contribute to the change in ECR conformation. Our signaling assay results also show that alternative splicing leads to changes in basal receptor activity, which suggests that the architecture and conformation of aGPCR ECRs play more important roles in their functions than previously thought. However, the signaling assay results showing that the change of Gpr126 ECR conformation is not solely responsible for changes in signaling may be confusing and contradictory. Rather, a more complex model that combines changes in ECR conformation with exposure of potential functional sites due to these changes may be key for alternative splicing-mediated regulation.

Importantly, we identified the calcium-binding site in Gpr126 as a potential functional site. Our in vivo results showed that zebrafish carrying two point-mutations in the calcium-binding site have defective SC and ear development, suggesting that the calcium-binding site is essential for the in vivo functions of Gpr126 (Fig. 7b). Since a subset of CUB domains from other

determined by SAD phasing using Crank2 in CCP4. Refinement was performed with both REFMAC5 (CCP4) and phenix.refine (PHENIX).

Negative-stain electron microscopy. Uranyl formate (0.75%) solution was freshly prepared by adding 5 mL boiling water to 37.5 mg uranyl formate (Electron Microscopy sciences, 22450). After stirring for 5 min in the dark, 10 μ L 5 M NaOH was added and stirred for an additional 5 min. The solution was syringe filtered (Millipore, SLGV033RS) and stored in the dark. Purified Gpr126 (–ss), (+ss), and (–ss) D134A/F135A ECR constructs were diluted to ~5 μ g mL⁻¹ and applied to glow-discharged EM grids (Electron Microscopy Sciences, CF400-Cu), using a conventional negative-stain protocol⁸⁶. To the grid was applied 2 μ L diluted protein for 30 sec. The protein was blotted off with filter paper (Sigma–Aldrich, WHA1001110), and then the grid was touched to a 25 μ L drop of distilled, filtered water. The water was blotted off, and the grid was touched to a second 25 μ L drop of water and blotted off. The grid was then touched to a 25 μ L drop of 0.75% uranyl formate for 30 sec and blotted off. The grid was air-dried for 30 sec. The sample was imaged on a Tecnai G2 F30 operated at 300 kV. Gpr126 –ss (6565 particles), +ss (2529 particles), and –ss D134A/F135A (3916 particles) were processed using EMAN2⁸⁷.

Small-angle X-ray scattering. SAXS measurements were performed at the Advanced Photon Source at Argonne National Laboratory (beamline 18-ID) with an in-line SEC columns (Superdex 200 or Biorad EnRich 5–650 10–300) equilibrated with 20 mM HEPES, pH 7.4, and 150 mM NaCl. Data were analyzed using autorg and datnom using the commands “autorg –sminrg 0.55 –smxrg 1.1” and “datnom ‘1.dat -r ‘2’ – skip ‘3’ -o ‘1.out.’”, respectively, where ‘1’ is the file name, ‘2’ is the R_g determined by autorg, and ‘3’ is the number of points removed at low q as determined from autorg. SAXS curves of molecular models were generated with Crysol version 2.83⁸⁸.

cAMP signaling assay. Full-length wild-type and mutant Gpr126/GPR126 constructs were cloned into pCMV5. All constructs include N-terminal FLAG-tags for measuring cell-surface expression levels. The following primers were used for amplification of zebrafish Gpr126: F: 5'-GCTGACTACAAAGACGATGACGACAAGCTTTGCAATGTGGTGTCTACCGACTCCAGGGC-3' and R: 5'-CCTGGCCAGGCCTCTGGTCCATGAGGCCCTTATTGACAGGTTACTATCTGCATTACTGTG-3'. The following primers were used for amplification of human GPR126: F: 5'-GCTGACTACAAAGACGATGACGACAAGCTTTGCGCAAACGTGATGGTGTCTCGA-3' and R: 5'-CAGGCCTCTGGTCCATGAGGCCCTCAACAGGGGCCAGTTTCCACCAG-3'. Zebrafish Gpr126 D134A/F135A mutant constructs were generated using the QuikChange method (Agilent) with primers: F: 5'-GATGGAGTTTCTTAACTCCGCGCTAGTGTCCAAAGAAAGGCTTCC-3' and R: 5'-GGAAGCCTTTCTTTGGACACTAGCGCGGAGTTAAAGAAA CCTCCATC-3'.

HEK293 cells (ATCC CRL-1573) were seeded in 6-well plates with Dulbecco's Modified Eagle Medium (DMEM; Gibco, 11965092) supplemented with 10% FBS (Sigma–Aldrich, F0926). At 60–70% confluency, the cells were co-transfected with 0.35 μ g Gpr126 DNA, 0.35 μ g GloSensor reporter plasmid (Promega, E2301), and 2.8 μ L transfection reagent Fugene 6 (Promega, PRE2693). After a 24-h incubation, the transfected cells were detached and seeded (50,000 cells per well) in a white 96-well assay plate. Following another 24-h incubation, the DMEM was replaced with 100 μ L Opti-MEM (Gibco, 31985079) and incubated for 30 min. To each well was then added 1 μ L GloSensor substrate and 11 μ L FBS. Basal-level luminescence measurements were taken after 30 min to allow for equilibration. For activation assays, the cells were then treated with either 1 mM p14 synthetic peptide (GenScript, N-THFGVLMDLPRASEKEK-Biotin-C) or vehicle DMSO for 15 min. Measurements were taken with a Synergy HTX BioTeck plate reader at 25 °C.

Flow cytometry to measure cell-surface expression of Gpr126. HEK293 cells were transfected as described above and incubated for 24 h. The cells were then detached and seeded in a 24-well plate. Following another 24-h incubation, the cells were detached with citric saline and washed with phosphate-buffered saline (PBS). The cells were then washed twice with PBS supplemented with 0.1% BSA (Sigma–Aldrich, A3803). The cells were incubated with mouse anti-FLAG primary antibody (1:1000 dilution in PBS + 0.1% BSA; Sigma–Aldrich, F3165) at room temperature for 30 min and washed twice with PBS + 0.1% BSA. The cells were then incubated with donkey anti-mouse Alexa Fluor 488 secondary antibody (1:500 dilution in PBS + 0.1% BSA; Invitrogen, A21202) at room temperature for 30 min and washed twice with PBS + 0.1% BSA. Stained cells were resuspended in PBS + 0.1% BSA and were analyzed with a BD Accuri C6 flow cytometer.

Zebrafish rearing. Zebrafish were maintained in the Washington University Zebrafish Consortium Facility (<http://zebrafish.wustl.edu>), and the following experiments were performed according to Washington University animal protocols. All zebrafish experiments were performed in compliance with institutional ethical regulations for animal testing and research at Washington University and Oregon Health and Science University (OHSU). Experiments were approved by the Animal Care and Use Committee of Washington University School of Medicine (St. Louis, MO) and the Institutional Care and Use Committee of OHSU (Portland,

OR). The *gpr126*^{sl464} zebrafish were generated within the wild-type AB* background. All crosses were either set up as pairs or harems and embryos were raised at 28.5 °C in egg water (5 mM NaCl, 0.17 mM KCl, 0.33 mM CaCl₂, 0.33 mM MgSO₄). Larvae were staged at days post fertilization (dpf). *gpr126*^{sl464} larvae can be identified at 4 dpf by a puffy ear phenotype.

Genotyping. To identify carriers of the *gpr126*^{sl464} allele, the following primers were used to amplify the 381 base pair (bp) locus of interest: F: 5'-GTTGTGCTCAAGACCCGGCAC-3' and R: 5'-TCCACCTCCCAGCTACAATTCC-3'. After amplification by PCR, the product was digested with either DrdI (NEB) at 37 °C or BstUI (NEB) at 60 °C, and then run on a 3% agarose gel. The mutation both disrupts a DrdI binding site and introduces a BstUI binding site. DrdI cleaves wild-type PCR product into 275 and 105 bp products, and the mutant product is 380 bp. BstUI cleaves mutant PCR product into 274 and 106 bp products, and the wild-type product is 380 bp. We recommend using BstUI for genotyping. Any larvae identified with the puffy ear phenotype were always genotyped as *gpr126*^{sl464} homozygous mutant ($n = 20/20$).

Guide RNA synthesis. Potential gRNA templates were generated by CHOPCHOP (<http://chopchop.cbu.uib.no/>). The chosen forward and reverse oligonucleotides, 20 bps upstream of the PAM sequence, were ordered with additional nucleotides added to the 5' end to permit cloning into the pDR274 vector⁸⁹. The oligonucleotide forward sequence used was: 5' - tag gAC TTT AGT GTC CAA AAG AA - 3' and oligonucleotide reverse sequence used was: 5' - aaa cTT CTT TTG GAC ACT AAA GT - 3'. 2 μ M of each oligonucleotide was mixed in annealing buffer (10 mM Tris, pH 8, 50 mM NaCl, 1 mM EDTA) and incubated at 90 °C for 5 min, then cooled to 25 °C over a 45 min time interval. The pDR274 vector was linearized with BsaI and oligonucleotides were ligated into the vector with T4 ligase (NEB) for 10 min at room temperature. The ligation reaction was transformed into competent cells and then plated on kanamycin LB plates. Selected colonies were grown, mini-prepped (Zyppy Plasmid Kits, Zymo Research), and Sanger sequenced. The gRNA DNA sequence was then PCR amplified from 50 ng μ L⁻¹ of the plasmid with Phusion (NEB) and the following primers: F: 5'-GTTGGAACCTCTTACGTGCC-3' and R: 5'-AAAAGCACCGACTCGGTG-3'. The PCR product was digested with DpnI at 37 °C for 1 h, heat inactivated at 80 °C for 20 min, and then purified with a Qiagen PCR Purification column. RNA was synthesized with a MEGAscript T7 Transcription Kit (Ambion).

Design of ssODN and microinjections. One-cell stage wild-type embryos were injected with either 2 or 3 nl of a solution containing ~132 ng μ L⁻¹ gRNA, ~148 ng μ L⁻¹ of *Cas9* mRNA (obtained from the Hope Center Transgenic Core at Washington University in St. Louis), and 60 ng μ L⁻¹ of the ssODN. The 150 bp ssODN was ordered from IDT and contained a 5 bp mutation (uppercase): 5'-atacat aacatacccttctgtaactgatggaagccttctttggacactCGCCGcgagtaagaaacctccatcac attccagtgaggtag-3'. Please note that an extra C (bolded), beginning after exon 3, is present in the ssODN that is not present in the *gpr126* reference sequence. The extra nucleotide was not integrated into the *sl464* mutants. At 1 dpf, embryos were genotyped for disruption of the wild-type DrdI binding site and screened for the characteristic *gpr126* puffy ear mutant phenotype. Mutations that were successfully transmitted to the F1 offspring were screened for by restriction enzyme digest analysis. Mutant bands were gel extracted (Qiagen Gel Extraction Kit) and Sanger sequenced to identify the incorporation of the ssODN containing the mutation of interest.

Whole mount in situ hybridization. 1 dpf larvae were treated with 0.003% phenylthiourea to inhibit pigmentation until fixation in 4% paraformaldehyde at 4 dpf. After fixation, larvae pooled in microcentrifuge tubes were dehydrated in methanol (5 by 5 min washes while nutating) and then stored at –20 °C. To begin in situ hybridization, larvae were re-hydrated with 50, 70, and 100% 0.2% PBS-Tween: methanol (5 min washes on a nutator). PBS-Tween washes were then continued (4 by 5 min washes while nutating). Larvae were then treated with 1:900 Proteinase-K (20 mg/mL) in PBS-Tween for 45 min at room temperature (not on a nutator). After Proteinase-K treatment, two quick PBS-Tween washes were performed, and then larvae were post-fixed in 4% paraformaldehyde for 20 min while nutating. Larvae were then washed with PBS-Tween (5 by 5 min washes on a nutator) and then incubated in hybridization buffer (Hyb(+)) at 65 °C for at least 1 h in a dry heat block. The riboprobe of interest in Hyb(+) was then incubated with the larvae overnight at 65 °C in a dry heat block. The following day, larvae were washed in 100% Hyb(+), 75% Hyb(+); 25% 2X SSCTween, 50% Hyb(+); 50% 2X SSCTween, 25% Hyb(+); 75% 2X SSCTween, all preheated to 65 °C, in a dry heat block. Larvae were then washed with 2X SSCTween (two washes, 30 min each) and 0.2X SSCTween (two washes, 30 min each) at 65 °C in a dry heat block. Larvae were then washed with MAB-TritonX-100 for 10 min at room temperature while nutating and then blocked in blocking solution (2% blocking reagent in MAB + 0.2% TritonX-100 + 10% sheep serum) for at least 1 h at room temperature on a nutator. Larvae were then treated with anti-DIG AP Fab fragments (1:2000 – Roche 11093274910) in blocking solution overnight at 4 °C on a nutator. The following day, larvae were washed with MAB-TritonX-100 (6 by 30 min washes on a nutator) at room temperature. After a 10 min wash in alkaline phosphatase/NTMT (AP

buffer, larvae were moved to a 24-well plate, covered in aluminum foil to prevent light exposure, and incubated with NBT ($2.2 \mu\text{L mL}^{-1}$) + BCIP ($1.6 \mu\text{L mL}^{-1}$) in AP buffer until the reaction completed. After development of the probe, larvae were washed with three quick PBS-Tween washes and then passed through 30, 50, and 70% glycerol washes. A complete protocol with detailed notes is available⁹⁰. The previously characterized riboprobes utilized in this study were *mhp* (GenBank: [AY860977.1](https://www.ncbi.nlm.nih.gov/nuclot/AY860977.1))⁹¹, and *gpr126*, originally synthesized with 5'- ggaattctgatggagctgtg-gaacatagc-3' and 5'-agtgtcgaactctcatctcaactcagcagc-3' primers³⁶. For *mhp*, larvae were scored for either presence or absence of signal expression along the PLLn.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The accession number for the coordinates and diffraction data for the Gpr126 (–ss) ECR crystal structure reported in this paper is PDB: [6V55](https://www.rcsb.org/structure/6V55). The SASBDB IDs for the SAXS experimental data are: [SASDFT9](https://www.sasbdb.org/entry/SASDFT9), [SASDFU9](https://www.sasbdb.org/entry/SASDFU9), [SDSDFV9](https://www.sasbdb.org/entry/SDSDFV9), [SASDFW9](https://www.sasbdb.org/entry/SASDFW9), [SASDFX9](https://www.sasbdb.org/entry/SASDFX9). The source data underlying Fig. 4b, c, e, f and Supplementary Figures 1B, 4C, H, and 6B, C are provided as a Source Data file. All other data are available from the corresponding author on reasonable request

Received: 2 May 2019; Accepted: 11 December 2019;

Published online: 10 January 2020

References

- Chun, L., Zhang, W. H. & Liu, J. F. Structure and ligand recognition of class C GPCRs. *Acta Pharmacologica Sin.* **33**, 312–323 (2012).
- de Graaf, C. et al. Extending the structural view of class B GPCRs. *Trends Biochem. Sci.* **42**, 946–960 (2017).
- Hu, P. & Luo, B. H. Integrin bi-directional signaling across the plasma membrane. *J. Cell. Physiol.* **228**, 306–312 (2013).
- Kovacs, E., Zorn, J. A., Huang, Y., Barros, T. & Kuriyan, J. A structural perspective on the regulation of the epidermal growth factor receptor. *Annu. Rev. Biochem.* **84**, 739–764 (2015).
- Rana, R. et al. Structural insights into the role of the Smoothed cysteine-rich domain in Hedgehog signalling. *Nat. Commun.* **4**, 2965 (2013).
- Shimaoka, M. & Springer, T. A. Therapeutic antagonists and conformational regulation of integrin function. *Nat. Rev. Drug Discov.* **2**, 703 (2003).
- Sun, Z., Guo, S. S. & Fassler, R. Integrin-mediated mechanotransduction. *J. Cell Biol.* **215**, 445–456 (2016).
- Luo, B. H., Carman, C. V. & Springer, T. A. Structural basis of integrin regulation and signaling. *Annu. Rev. Immunol.* **25**, 619–647 (2007).
- Salzman, G. S. et al. Structural basis for regulation of GPR56/ADGRG1 by its alternatively spliced extracellular domains. *Neuron* **91**, 1292–1304 (2016).
- Li, S., Schmitz, K. R., Jeffrey, P. D., Wiltzius, J. J., Kussie, P. & Ferguson, K. M. Structural basis for inhibition of the epidermal growth factor receptor by cetuximab. *Cancer cell* **7**, 301–311 (2005).
- Ley, K., Rivera-Nieves, J., Sandborn, W. J. & Shattil, S. Integrin-based therapeutics: biological basis, clinical use and new drugs. *Nat. Rev. Drug Discov.* **15**, 173–183 (2016).
- Traynor, K. FDA approves licensing of erenumab-aoe to prevent migraine. *Am. J. Health-Syst. Pharm. Off. J. Am. Soc. Health Syst. Pharmacists* **75**, 929–930 (2018).
- Shi, L. et al. Pharmacologic characterization of AMG 334, a potent and selective human monoclonal antibody against the calcitonin gene-related peptide receptor. *J. Pharmacol. Exp. Ther.* **356**, 223–231 (2016).
- Hamann, J. et al. International Union of Basic and Clinical Pharmacology. XCIV. Adhesion G protein-coupled receptors. *Pharmacol. Rev.* **67**, 338–367 (2015).
- Yona, S., Lin, H.-H., Siu, W. O., Gordon, S. & Stacey, M. Adhesion-GPCRs: emerging roles for novel receptors. *Trends Biochem. Sci.* **33**, 491–500 (2008).
- Langenhan, T., Piao, X. & Monk, K. R. Adhesion G protein-coupled receptors in nervous system development and disease. *Nat. Rev. Neurosci.* **17**, 550–561 (2016).
- Manglik, A. & Kruse, A. C. Structural basis for G protein-coupled receptor activation. *Biochemistry* **56**, 5628–5634 (2017).
- Paavola, K. J. & Hall, R. A. Adhesion G protein-coupled receptors: signaling, pharmacology, and mechanisms of activation. *Mol. Pharmacol.* **82**, 777–783 (2012).
- Langenhan, T., Aust, G. & Hamann, J. Sticky signaling—adhesion class G protein-coupled receptors take the stage. *Sci. Signal.* **6**, re3 (2013).
- Arac, D. et al. A novel evolutionarily conserved domain of cell-adhesion GPCRs mediates autoprolysis. *EMBO J.* **31**, 1364–1378 (2012).
- Chang, G. W., Stacey, M., Kwakkenbos, M. J., Hamann, J., Gordon, S. & Lin, H. H. Proteolytic cleavage of the EMR2 receptor requires both the extracellular stalk and the GPS motif. *FEBS Lett.* **547**, 145–150 (2003).
- Paavola, K. J., Stephenson, J. R., Ritter, S. L., Alter, S. P. & Hall, R. A. The N terminus of the adhesion G protein-coupled receptor GPR56 controls receptor signaling activity. *J. Biol. Chem.* **286**, 28914–28921 (2011).
- Kishore, A. & Hall, R. A. Disease-associated extracellular loop mutations in the adhesion G protein-coupled receptor G1 (ADGRG1; GPR56) differentially regulate downstream signaling. *J. Biol. Chem.* **292**, 9711–9720 (2017).
- Kishore, A., Purcell, R. H., Nassiri-Toosi, Z. & Hall, R. A. Stalk-dependent and stalk-independent signaling by the adhesion G protein-coupled receptors GPR56 (ADGRG1) and BA11 (ADGRB1). *J. Biol. Chem.* **291**, 3385–3394 (2016).
- Salzman, G. S., Zhang, S., Gupta, A., Koide, A., Koide, S. & Arac, D. Stachel-independent modulation of GPR56/ADGRG1 signaling by synthetic ligands directed to its extracellular region. *Proc. Natl Acad. Sci. USA* **114**, 10095–10100 (2017).
- Koh, J. T. et al. Extracellular fragment of brain-specific angiogenesis inhibitor 1 suppresses endothelial cell proliferation by blocking alphavbeta5 integrin. *Exp. Cell Res.* **294**, 172–184 (2004).
- Muller, A. et al. Oriented cell division in the *C. elegans* embryo is coordinated by G-protein signaling dependent on the adhesion GPCR LAT-1. *PLoS Genet.* **11**, e1005624 (2015).
- Petersen, S. C. et al. The adhesion GPCR GPR126 has distinct, domain-dependent functions in Schwann cell development mediated by interaction with laminin-211. *Neuron* **85**, 755–769 (2015).
- Promel, S. et al. The GPS motif is a molecular switch for bimodal activities of adhesion class G protein-coupled receptors. *Cell Rep.* **2**, 321–331 (2012).
- Patra, C. et al. Organ-specific function of adhesion G protein-coupled receptor GPR126 is domain-dependent. *Proc. Natl Acad. Sci. USA* **110**, 16898–16903 (2013).
- Demberg, L. M., Rothmund, S., Schoneberg, T. & Liebscher, I. Identification of the tethered peptide agonist of the adhesion G protein-coupled receptor GPR64/ADGRG2. *Biochem. Biophys. Res. Commun.* **464**, 743–747 (2015).
- Liebscher, I. et al. A tethered agonist within the ectodomain activates the adhesion G protein-coupled receptors GPR126 and GPR133. *Cell Rep.* **9**, 2018–2026 (2014).
- Stoveken, H. M., Hajduczuk, A. G., Xu, L. & Tall, G. G. Adhesion G protein-coupled receptors are activated by exposure of a cryptic tethered agonist. *Proc. Natl Acad. Sci. USA* **112**, 6194–6199 (2015).
- Wilde, C. et al. The constitutive activity of the adhesion GPCR GPR114/ADGRG5 is mediated by its tethered agonist. *FASEB J.* **30**, 666–673 (2016).
- Geng, F. S. et al. Semicircular canal morphogenesis in the zebrafish inner ear requires the function of *gpr126* (lauscher), an adhesion class G protein-coupled receptor gene. *Development* **140**, 4362–4374 (2013).
- Monk, K. R. et al. A G protein-coupled receptor is essential for Schwann cells to initiate myelination. *Science* **325**, 1402–1405 (2009).
- Waller-Evans, H. et al. The orphan adhesion-GPCR GPR126 is required for embryonic development in the mouse. *PLoS ONE* **5**, e14047 (2010).
- Pereira, J. A., Lebrun-Julien, F. & Suter, U. Molecular mechanisms regulating myelination in the peripheral nervous system. *Trends Neurosci.* **35**, 123–134 (2012).
- d'Ydewalle, C., Benoy, V. & Van Den Bosch, L. Charcot-Marie-Tooth disease: emerging mechanisms and therapies. *Int. J. Biochem. Cell Biol.* **44**, 1299–1304 (2012).
- Nave, K. A., Sereda, M. W. & Ehrenreich, H. Mechanisms of disease: inherited demyelinating neuropathies—from basic to clinical research. *Nat. Clin. Pract. Neurol.* **3**, 453–464 (2007).
- Monk, K. R., Oshima, K., Jors, S., Heller, S. & Talbot, W. S. Gpr126 is essential for peripheral nerve development and myelination in mammals. *Development* **138**, 2673–2680 (2011).
- Mogha, A. et al. Gpr126 functions in Schwann cells to control differentiation and myelination via G-protein activation. *J. Neurosci.* **33**, 17976–17985 (2013).
- Kuffer, A. et al. The prion protein is an agonistic ligand of the G protein-coupled receptor Adgrg6. *Nature* **536**, 464–468 (2016).
- Garinet, S. et al. High prevalence of a hotspot of non-coding somatic mutations in intron 6 of GPR126 in bladder cancer. *Mol. Cancer Res.* **17**, 469–475 (2018).
- Nik-Zainal, S. et al. Landscape of somatic mutations in 560 breast cancer whole-genome sequences. *Nature* **534**, 47–54 (2016).
- Piraino, S. W. & Furney, S. J. Identification of coding and non-coding mutational hotspots in cancer genomes. *BMC Genomics* **18**, 17 (2017).
- Qin, X. et al. Genetic variant of *gpr126* gene is functionally associated with adolescent idiopathic scoliosis in Chinese population. *Spine* **42**, E1098–E1103 (2017).

48. Terzikhan, N. et al. Heritability and genome-wide association study of diffusing capacity of the lung. *Eur. Respir. J.* **52**, pii: 1800647 (2018).
49. Kou, I. et al. Genetic variants in GPR126 are associated with adolescent idiopathic scoliosis. *Nat. Genet.* **45**, 676–679 (2013).
50. Ravenscroft, G. et al. Mutations of GPR126 are responsible for severe arthrogyposis multiplex congenita. *Am. J. Hum. Genet.* **96**, 955–961 (2015).
51. Moriguchi, T., Haraguchi, K., Ueda, N., Okada, M., Furuya, T. & Akiyama, T. DREG, a developmentally regulated G protein-coupled receptor containing two conserved proteolytic cleavage sites. *Genes Cells.* **9**, 549–560 (2004).
52. Stehlik, C., Kroismayr, R., Dorfleutner, A., Binder, B. R. & Lipp, J. VIGR—a novel inducible adhesion family G-protein coupled receptor in endothelial cells. *FEBS Lett.* **569**, 149–155 (2004).
53. Bjarnadottir, T. K., Geirardsdottir, K., Ingemansson, M., Mirza, M. A., Fredriksson, R. & Schiöth, H. B. Identification of novel splice variants of Adhesion G protein-coupled receptors. *Gene* **387**, 38–48 (2007).
54. Xu, E., Shao, W., Jiang, H., Lin, T., Gao, R. & Zhou, X. A genetic variant in GPR126 causing a decreased inclusion of exon 6 is associated with cartilage development in adolescent idiopathic scoliosis population. *BioMed. Res. Int.* **2019**, 4678969 (2019).
55. Holm, L. & Laakso, L. M. Dali server update. *Nucleic Acids Res.* **44**, W351–355 (2016).
56. Harding, M. M. The geometry of metal-ligand interactions relevant to proteins. II. Angles at the metal atom, additional weak metal-donor interactions. *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **56**, 857–867 (2000).
57. Gaboriaud, C., Gregory-Pauron, L., Teillet, F., Thielens, N. M., Bally, I. & Arlaud, G. J. Structure and properties of the Ca(2+)-binding CUB domain, a widespread ligand-recognition unit involved in major biological functions. *Biochem. J.* **439**, 185–193 (2011).
58. Andersen, C. B., Madsen, M., Storm, T., Moestrup, S. K. & Andersen, G. R. Structural basis for receptor recognition of vitamin-B(12)-intrinsic factor complexes. *Nature* **464**, 445–448 (2010).
59. Cohen-Dvashi, H., Kilimnik, I. & Diskin, R. Structural basis for receptor recognition by Lujo virus. *Nat. Microbiol.* **3**, 1153–1160 (2018).
60. Gingras, A. R. et al. Structural basis of mannan-binding lectin recognition by its associated serine protease MASP-1: implications for complement activation. *Structure* **19**, 1635–1643 (2011).
61. Venkatraman Girija, U. et al. Structural basis of the C1q/C1s interaction and its central role in assembly of the C1 complex of complement activation. *Proc. Natl Acad. Sci. USA* **110**, 13916–13920 (2013).
62. Ashkenazy, H. et al. ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res.* **44**, W344–350 (2016).
63. Paavola, K. J., Sidik, H., Zuchero, J. B., Eckart, M. & Talbot, W. S. Type IV collagen is an activating ligand for the adhesion G protein-coupled receptor GPR126. *Sci. Signal.* **7**, ra76 (2014).
64. Gordon, W. R., Vardar-Ulu, D., Histen, G., Sanchez-Irizarry, C., Aster, J. C. & Blacklow, S. C. Structural basis for autoinhibition of Notch. *Nat. Struct. Mol. Biol.* **14**, 295–300 (2007).
65. Macao, B., Johansson, D. G., Hansson, G. C. & Hard, T. Autoproteolysis coupled to protein folding in the SEA domain of the membrane-bound MUC1 mucin. *Nat. Struct. Mol. Biol.* **13**, 71–76 (2006).
66. Gordon, W. R. et al. Effects of S1 cleavage on the structure, surface export, and signaling activity of human Notch1 and Notch2. *PLoS ONE* **4**, e6613 (2009).
67. Pelaseyed, T., Zäch, M., Petersson, Å. C., Svensson, F., Johansson, D.G.A., Hansson, G.C. Unfolding dynamics of the mucin SEA domain probed by force spectroscopy suggest that it acts as a cell protective device. *FEBS J.* **280**, 1491–1501 (2013).
68. van Putten, J. P. M. & Strijbis, K. Transmembrane mucins: signaling receptors at the intersection of inflammation and cancer. *J. Innate Immun.* **9**, 281–299 (2017).
69. Aoto, J., Martinelli, D. C., Malenka, R. C., Tabuchi, K. & Südhof, T. C. Presynaptic neuexin-3 alternative splicing trans-synaptically controls postsynaptic AMPA receptor trafficking. *Cell* **154**, 75–88 (2013).
70. Irimia, M. et al. A highly conserved program of neuronal microexons is misregulated in autistic brains. *Cell* **159**, 1511–1523 (2014).
71. Thalhammer, A. et al. Alternative splicing of P/Q-type Ca(2+) channels shapes presynaptic plasticity. *Cell Rep.* **20**, 333–343 (2017).
72. Fuccillo, M. V. et al. Single-cell mRNA profiling reveals cell-type-specific expression of neuexin isoforms. *Neuron* **87**, 326–340 (2015).
73. Li, J. et al. Structural basis for teneurin function in circuit-wiring: a toxin motif at the synapse. *Cell* **173**, 735–748.e715 (2018).
74. Sandberg, A., Johansson, D. G., Macao, B. & Hard, T. SEA domain autoproteolysis accelerated by conformational strain: energetic aspects. *J. Mol. Biol.* **377**, 1117–1129 (2008).
75. Lum, A. M., Wang, B. B., Beck-Engeser, G. B., Li, L., Channa, N. & Wabl, M. Orphan receptor GPR110, an oncogene overexpressed in lung and prostate cancer. *BMC Cancer* **10**, 40 (2010).
76. Cho, H. S. & Leahy, D. J. Structure of the extracellular region of HER3 reveals an interdomain tether. *Science* **297**, 1330–1333 (2002).
77. Eubelen, M. et al. A molecular mechanism for Wnt ligand-specific signaling. *Science* **361**, pii: eaat1178 (2018).
78. Vallon, M. et al. A RECK-WNT7 receptor-ligand interaction enables isoform-specific regulation of Wnt bioavailability. *Cell Rep.* **25**, 339–349.e339 (2018).
79. Zhou, Y. & Nathans, J. Gpr124 controls CNS angiogenesis and blood-brain barrier integrity by promoting ligand-specific canonical wnt signaling. *Dev. Cell* **31**, 248–256 (2014).
80. Cho, C., Smallwood, P. M. & Nathans, J. Reck and Gpr124 are essential receptor cofactors for Wnt7a/Wnt7b-specific signaling in mammalian CNS angiogenesis and blood-brain barrier regulation. *Neuron* **95**, 1056–1073.e1055 (2017).
81. Jackson, V. A. et al. Structural basis of latrophilin-FLRT interaction. *Structure* **23**, 774–781 (2015).
82. Jackson, V. A. et al. Super-complexes of adhesion GPCRs and neural guidance receptors. *Nat. Commun.* **7**, 11184 (2016).
83. Lu, Y. C. et al. Structural basis of latrophilin-FLRT-UNC5 interaction in cell adhesion. *Structure* **23**, 1678–1691 (2015).
84. Tu, Y. K., Duman, J. G. & Tolias, K. F. The adhesion-GPCR BAI1 promotes excitatory synaptogenesis by coordinating bidirectional trans-synaptic signaling. *J. Neurosci.* **38**, 8388–8406 (2018).
85. Dong, G., Wearsch, P. A., Peaper, D. R., Cresswell, P. & Reinisch, K. M. Insights into MHC class I peptide loading from the structure of the tapasin-ERp57 thiol oxidoreductase heterodimer. *Immunity* **30**, 21–32 (2009).
86. Booth, D. S., Avila-Sakar, A. & Cheng, Y. Visualizing proteins and macromolecular complexes by negative stain EM: from grid preparation to image acquisition. *J. Vis. Exp.* **22**(58), pii: 3227 (2011).
87. Tang, G. et al. EMAN2: an extensible image processing suite for electron microscopy. *J. Struct. Biol.* **157**, 38–46 (2007).
88. Petoukhov, M. V. et al. New developments in the ATSAS program package for small-angle scattering data analysis. *J. Appl. Crystallogr.* **45**, 342–350 (2012).
89. Hwang, W. Y. et al. Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat. Biotechnol.* **31**, 227–229 (2013).
90. Cunningham, R. L. & Monk, K. R. Whole mount in situ hybridization and immunohistochemistry for zebrafish larvae. *Methods Mol. Biol.* **1739**, 371–384 (2018).
91. Lyons, D. A. et al. *erbb3* and *erbb2* are essential for schwann cell migration and myelination in zebrafish. *Curr. Biol.* **15**, 513–524 (2005).

Acknowledgements

We thank Engin Özkan for assistance with crystal structure determination and for insightful discussions, the He lab for the use of their luminescence plate reader, and other members of the Araç lab for helpful discussions. We also thank the staff at the University of Chicago Advance Electron Microscopy Facility, the Advanced Photon Source (APS) at Argonne National Labs (ANL), the Washington University Zebrafish Consortium Facility, and Fernanda Coelho for assistance with in situ hybridization. Crystal diffraction and SAXS data were collected at the APS, GM/CA 23-ID-D (supported by NCI ACB-12002 and the NIGMS AGM-12006 grants) and BioCAT 18-ID (supported by NIH grant 9 P41 GM103622-18), respectively. This research used resources of the APS, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by ANL under Contract No. DE-AC02-06CH11357. The Eiger 16 M detector was funded by an NIH–Office of Research Infrastructure Programs, High-End Instrumentation Grant (1S10OD012289-01A1). Use of the Pilatus 3 1 M detector was provided by grant 1S10OD018090-01 from NIGMS. This work was supported by NIH grants R01-GM120322 (D.A.), R01-NS079445 (K.R.M.), and T32GM007183 (K.L.). R.L.C. was supported by the National Science Foundation Graduate Research Fellowship (DGE-1745038). The content is solely the responsibility of the authors and does not necessarily reflect the official views of the National Institute of General Medical Sciences or the National Institutes of Health.

Author contributions

K.L. cloned, expressed, purified proteins (with assistance from E.F.), carried out bioinformatic and biochemical characterizations, performed crystallography experiments (with assistance from J.L.) and structure determination, performed cAMP-based signaling assays, and collected and analyzed negative-stain EM data (with assistance from M.Z.). R.L.C. and K.R.M. designed and analyzed and R.L.C. performed zebrafish experiments. J.A.R., T.R.S., and K.L. designed and performed SAXS experiments. K.L. and D.A. designed all experiments, interpreted results, and wrote the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41467-019-14040-1>.

Correspondence and requests for materials should be addressed to D.A.

Peer review information *Nature Communications* thanks Haydyn Mertens, and the other, anonymous, reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020