THE UNIVERSITY OF CHICAGO

STRUCTURAL BASIS FOR ADHESION G PROTEIN-COUPLED RECEPTOR GPR126 FUNCTION

A DISSERTATION SUBMITTED TO

THE FACULTY OF THE DIVISION OF THE BIOLOGICAL SCIENCES

AND THE PRITZKER SCHOOL OF MEDICINE

IN CANDIDACY FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

GRADUATE PROGRAM IN BIOCHEMISTRY AND MOLECULAR BIOPHYSICS

BY

KATHERINE LEON

CHICAGO, ILLINOIS

JUNE 2019

Table of Contents

Abbreviations

7TM:	**Seven-Transmembrane Domain**
aGPCR:	**Adhesion G Protein-Coupled Receptor**
β2AR:	**Beta-2 Adrenergic Receptor**
cAMP:	**Cyclic AMP**
CAM:	**Cell Adhesion Molecule**
CTF:	**C-Terminal Fragment**
CUB:	**Complement C1r/C1s, Uegf, Bmp1**
ECR:	**Extracellular Region**
EM:	**Electron Microscopy**
EV:	**Empty Vector**
EGFR:	**Epidermal Growth Factor Receptor**
GAIN:	**GPCR Autoproteolyis Inducing**
GPCR:	**G Protein-Coupled Receptor**
HormR:	**Hormone-Binding**
Lphn:	**Latrophilin**
NTF:	**N-Terminal Fragment**
PTX:	**Pentraxin**
SC:	**Schwann Cell**
SEA:	**Sperm protein, Enterokinase, and Agrin**
TM:	**Transmembrane**

List of Figures

List of Tables

Abstract

G protein-coupled receptors (GPCRs) are the largest superfamily of cell surface receptors and are responsible for mediating numerous physiological responses, which make them attractive drug targets. The second largest but least well-understood family of GPCRs, adhesion GPCRs (aGPCRs), have key functions in diverse biological processes including myelination, angiogenesis, neutrophil activation, synaptogenesis, and more. aGPCRs have large alternatively-spliced extracellular regions (ECRs) that mediate cell communication and may be drug targets to modulate aGPCR function. However, their structures and mechanisms of action remain unclear. The aGPCR Gpr126/ADGRG6 regulates Schwann cell myelination, heart development, and ear canal formation; and *GPR126* mutations cause myelination defects in human. We determined the structure of the complete Gpr126 ECR and revealed five domains including a previously-unknown proteolytic 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 newly-characterized calcium-binding site abolishes Schwann cell myelination in transgenic zebrafish. In addition to Gpr126, we also investigated the G protein signaling of other aGPCRs, GPR56 and Lphn, which led to the development of G protein signaling assays to probe aGPCR function. Altogether, these results demonstrate that Gpr126, and likely other aGPCRs, utilize a multi-faceted dynamic approach to regulate function and provide novel insights into ECR-targeted drug design.

Chapter 1: Introduction

Cellular communication and cell surface receptors

Plasma membranes are essential for cells to survive and function properly in their environments. Cells rely on their membranes to isolate their contents from the extracellular matrix. Without this barrier, cellular components would not be able to work together in a cohesive manner. On the other hand, plasma membranes also mediate contact with other cells and the extracellular environment in a regulated manner. Importantly, multicellular organisms rely on this intricate system of cellular communication in order to carry out essential biological processes, such as cortex development, myelination, immune response, synapse formation, and much more (Janeway et al., 1985; Scheiffele, 2003; Yuzwa and Miller, 2017).

The plasma membrane is made up of a selectively permeable phospholipid bilayer which helps regulate what comes in and out of cells (Robertson, 2018). Small hydrophobic and/or uncharged molecules can readily pass through the membrane. However, for other molecules such as ions and large polar molecules, cells must utilize membrane-bound channels and transporters in order to facilitate transport across the plasma membrane (Yang and Hinner, 2015).

In addition to molecules that physically pass the membrane, extracellular stimuli such as light, hormones, and proteins can transduce signals across the membrane via cell-surface receptors (Uings and Farrow, 2000). Because of the diversity of extracellular cues, there are numerous cell surface receptors that recognize and respond to each stimulus. An example of such cell signaling is the epidermal growth

factor receptor (EGFR), a family member of the receptor tyrosine kinases. EGFR has an extracellular ligand-binding domain, a transmembrane domain, and an intracellular kinase domain. EGFR binds to its ligand, EGF, through its extracellular domain, which leads to dimerization of EGFR and subsequent phosphorylation of each kinase domain (Wee and Wang, 2017). The recruitment of signaling proteins to these phosphorylated groups then initiates intracellular signaling pathways which, importantly, leads to regulation of gene expression to drive cell growth, differentiation, and proliferation (Scaltriti and Baselga, 2006).

G protein-coupled receptors (GPCRs) are another type of cell surface receptor that are responsible for recognizing extracellular stimuli and transmitting this information across the plasma membrane (Katritch et al., 2013). The stimuli are quite diverse, including photons, small molecules, peptide hormones, and proteins. These ligands bind to the extracellular side of GPCRs, typically an extracellular domain or the extracellular-facing seven-pass transmembrane (7TM) domain. This binding event triggers changes to the conformational landscape of the 7TM domain and allows for recruitment of G proteins to the cytoplasmic surface of the GPCR and activation of downstream signaling pathways (Weis and Kobilka, 2014). One of the more well-studied GPCRs, beta-2 adrenergic receptor (β2AR), recognizes the hormone epinephrine and activates G protein signaling in order to mediate a smooth muscle relaxation response (Barisione et al., 2010; Zhang et al., 1990).

The ability for receptor tyrosine kinases, GPCRs, and other cell surface receptors to detect and respond to changes in their environment is critical for cells to carry out biological functions in a regulated, collective manner. Studying the molecular

mechanisms of specific cell surface receptors is an important step in understanding the mechanisms of biological functions themselves and also in developing therapeutics that can modulate receptor activities. For example, propranolol is an antagonist drug which targets β1AR and β2AR to treat high blood pressure in patients (Admani et al., 2014). The rigorous research that has gone into studying cell surface receptors has led to many successful drugs currently on the market. However, there are still many receptors whose biological functions and mechanisms of action remain unknown, and this thesis will focus on a family of receptors that is relatively understudied but are involved in numerous essential cellular functions.

Cell adhesion molecules

Another important component of cellular communication is cellular adhesion, which is critical for many biological processes, including cell migration, tissue organization during development, cell differentiation, and more (Basson, 2012; De Pascalis and Etienne-Manneville, 2017; Mattes and Scholpp, 2018). Cell adhesion is the process by which cells interact with each other or with the extracellular matrix through the proteins present on their plasma membranes, termed cell adhesion molecules (CAMs) (Gumbiner, 1996). CAMs were originally thought to primarily act as a glue between cells and between cells and their extracellular matrix. However, it has become clear that CAMs also serve to transmit signals across membranes in order to regulate biological functions (Buckley et al., 1998; Cavallaro and Dejana, 2011). In general, CAMs have common domain organizations: extracellular regions with one or more ligand-binding sites, a small transmembrane region, and a cytoplasmic region that

contacts cellular signaling proteins. The main types of CAMS are integrins, cadherins, immunoglobulin superfamily cell adhesion molecules (IgCAMs), and selectins (Freemont and Hoyland, 1996) (Figure 1.1).

Integrins are heterodimeric CAMS consisting of one α chain and one β chain, with different combinations of α and β chains generating 24 distinct integrins that can bind to different extracellular matrix proteins, such as fibronectin, collagen and laminin. Both α and β chains have large, multidomain extracellular regions, single-pass transmembrane domains, and short cytoplastic regions. The extracellular regions are in in bent V-like shapes until a ligand-binding event stabilizes an extended conformation and leads to separation of the transmembrane and cytoplasmic regions between the two subunits (Calderwood, 2004; Campbell and Humphries, 2011). The cytoplasmic regions interact with proteins such as talin, α-actinin, and filamin, which link integrins to the actin cytoskeleton in order to regulate changes in cytoskeletal-driven functions such as cellular movement (Huttenlocher and Horwitz, 2011).

Similar to integrins, cadherins are made up of small cytoplasmic regions, single-pass transmembrane domains, and a large multidomain extracellular region. Cadherins are calcium dependent and form cis-dimers on the same cell surface and also form trans-dimers with other cadherin cis-dimers on neighboring cells, providing a tight adhesion between cells. The cytoplasmic regions of cadherins interact with catenins, which links cadherins to the actin cytoskeleton and can also lead to regulation of gene expression and Rho GTPase activity (Cavallaro and Dejana, 2011).

IgCAMs, like their name suggests, are cell adhesion molecules composed of extracellular immunoglobulin repeats. These repeats are typically followed by

extracellular fibronectin domains, a single-pass transmembrane domain, and a cytoplasmic region. Other IgCAMs can present themselves on cell surfaces with glycosylphosphatidylinositol (GPI) anchors. IgCAMS bind to proteins, usually integrins or other IgCAMs, on neighboring cells through their Ig and fibronectin domains. The cytoplasmic domains of IgCAMs interact with actin, ankyrins and spectrin, linking IgCAMs to the cytoskeleton, similar to integrins and cadherins. In addition, IgCAMs can also regulate receptor tyrosine kinases (Gibson, 2011). For example, neural cell adhesion molecule (NCAM) inhibits binding of the fibroblast growth factor receptor (FGFR) to FGF. This blocks FGF-dependent cell proliferation and instead activates neurite outgrowth (Francavilla et al., 2007).

Selectins have lectin and EGF-like domains in their extracellular regions. Depending on the type of selectin, this region is also followed by a series of sushi domains, as well as a transmembrane domain and cytoplasmic region. Selectins bind to carbohydrate groups on glycoproteins through their lectin domains. These molecules are known to regulate leukocyte extravasation through cell adhesion between glycoproteins on the leukocyte membrane and P-selectin on endothelial cells of blood vessels. Because this interaction has low affinity, leukocytes can “roll” along the endothelial cells and then eventually move between cells and out of the blood vessel (Crockett-Torabi, 1998; McEver, 2015).

Through coordination of their extracellular, transmembrane, and intracellular regions, this diverse set of CAMs allows cells to survive in their environments by both interacting with neighboring cells to form tissues and responding to environmental cues through cell signaling. Decades of research on these molecules have provided great

insight into their mechanisms of action and relevance for human health. The focus of this thesis, adhesion G protein-coupled receptors, have extracellular adhesion domains similar to the CAMS described above. However, because they also have seven-pass transmembrane regions common to traditional GPCRs, they are an intriguing and complex family of receptors that bridge both cell adhesion and G protein signaling.

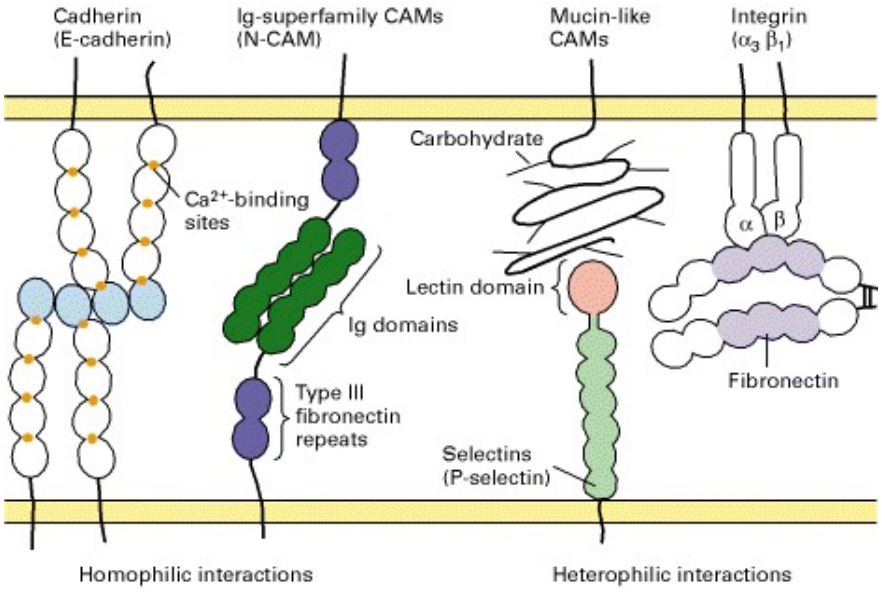


Figure 1.1 Major families of cell adhesion molecules

The major cell adhesion molecule families consist of large extracellular regions, small transmembrane domains, and cytoplasmic regions. Calcium-dependent cadherins form cis- and trans-homodimers. IgCAMs also form homodimers through their Ig domains. Selectins bind to carbohydrate groups on other proteins through their lectin domains. Integrins, composed of α and β chains bind to extracellular matrix proteins such as fibronectin. Figure adapted from Lodish H, Berk A, Zipursky SL, et al. (2000) *Molecular Cell Biology. 4th edition.* New York: W. H. Freeman.

G protein-coupled receptors

G protein-coupled receptors (GPCRs) represent the largest superfamily of cell surface signaling molecules in eukaryotes (Venkatakrishnan et al., 2013) and are split into five families based on phylogenetic analysis: rhodopsin, secretin, glutamate, frizzled, and adhesion. In humans, over 800 genes encode for GPCRs and they are responsible for regulation of almost every physiological function including taste, smell, inflammation, hormone response, and more (Fredriksson et al., 2003b). Because of these important roles, GPCRs are very attractive pharmacological targets, and currently about half of the drugs on the market target GPCR activity (Lagerstrom and Schioth, 2008).

GPCRs consist of an extracellular N-terminal region, an intracellular C-terminal region, and a canonical signaling seven-transmembrane (7TM) domain. GPCRs are responsible for detecting extracellular signals, transmitting the signal via conformation changes of the 7TM domain, and activating intracellular heterotrimeric (Gα, Gβ, Gγ) GDP-bound G proteins (Venkatakrishnan et al., 2013). Following activation of the GDP-bound G-proteins, GDP is exchanged for GTP, which results in dissociation of Gα and Gβ•Gγ from the receptor. The released G-proteins can then induce downstream signaling pathways (Cabrera-Vera et al., 2003). The signal cascade that occurs depends on the Gα subtype: $G_s\alpha$, $G_q\alpha$, $G_i\alpha$, or $G_{12/13}\alpha$. $G_s\alpha$ family members stimulate adenylyl cyclase to increase production of second messenger cyclic AMP (cAMP), $G_q\alpha$ family members activate phospholipase C to produce second messenger diacyl glycerol, $G_i\alpha$ family members inhibit adenylyl cyclase and production of cAMP, and $G_{12/13}\alpha$ family members activate RhoGEF (Cabrera-Vera et al., 2003; Knall and

Johnson, 1998). Signaling is arrested through phosphorylation of the receptor by GPCR kinases, which recruits beta-arrestins to desensitize the receptor (Luttrell and Lefkowitz, 2002). Recent studies have shown that GPCR signaling is a very complex process (Figure 1.2). For example, GPCRs are known to couple to more than one Gα protein (Baker and Hill, 2007) and can even activate alternative signaling pathways through beta-arrestins (Ma and Pei, 2007). Nevertheless, G protein signaling assays and other cell-based assays have proven to be powerful methods for studying the signaling properties of numerous GPCRs and have contributed to our better understanding of GPCR biology (Zhang and Xie, 2012).

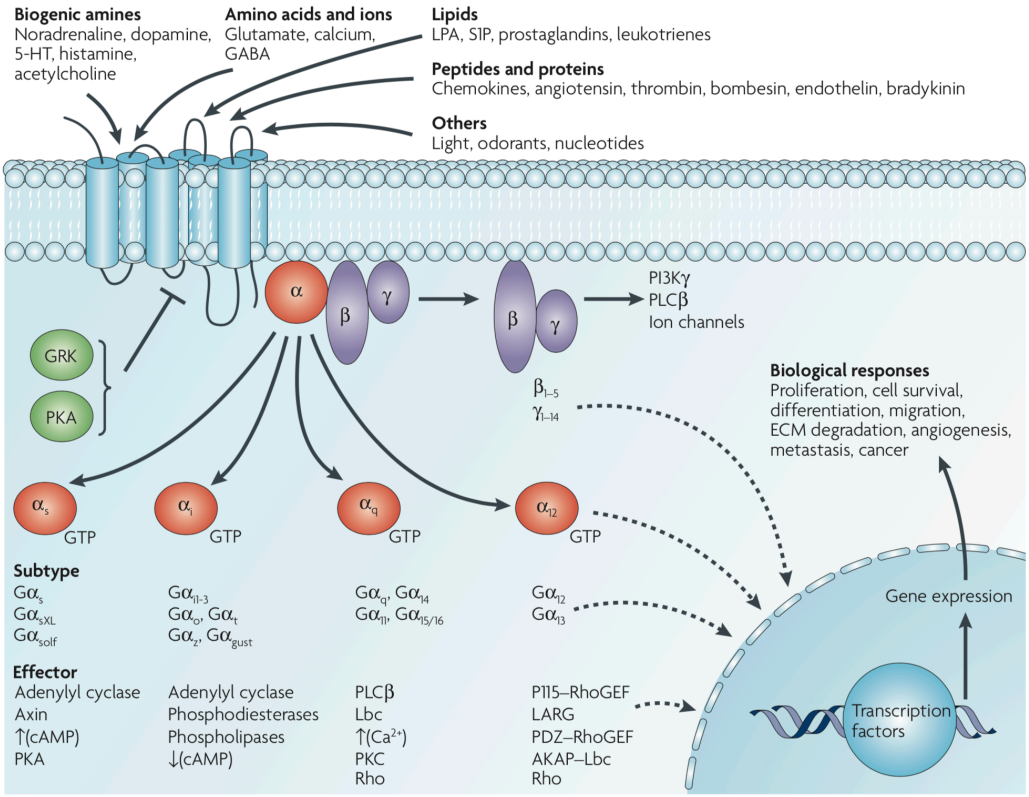


Figure 1.2 Signaling pathways of GPCRs

A diverse set of extracellular stimuli activate GPCRs, leading to activation of G proteins and downstream signaling pathways. The signaling pathway depends on the subtype of G proteins, which activate different effector proteins leading to regulation of gene expression and biological responses. After a signaling event, GPCR kinases (GRK) and protein kinase A (PKA) phosphorylate the receptor in order to arrest signaling. Figure adapted from (Dorsam and Gutkind, 2007).

An exciting advancement in the GPCR field has been the recent surge in crystallographic and cryo-EM structures of the 7TM domains of GPCRs which have provided great insight into their molecular mechanisms, including conformational changes between inactive and active structures (Gurevich and Gurevich, 2017; Thal et al., 2018; Weis and Kobilka, 2014), multi-domain regulatory interactions (Zhang et al., 2017a; Zhang et al., 2017b), activation of G proteins (Kang et al., 2018; Koehl et al., 2018; Liang et al., 2017; Rasmussen et al., 2011b), and more. There are currently high-

resolution structures of 7TM domains for all GPCR families except for the adhesion GPCR (aGPCR) family, a more recently discovered GPCR family of receptors that have uniquely large and variable extracellular regions (ECRs). Because of the lack of structural information for aGPCRs, this family is not as well understood and there remain many unanswered questions about their mechanisms of action.

Adhesion G protein-coupled receptors[1,2]

With 33 members in humans, adhesion G-Protein Coupled Receptors (aGPCRs) make up the second largest GPCR family but are the least studied and least understood (Fredriksson et al., 2003b). They are found in animals, protozoa, and alveolates, predating many other important signaling molecules (Krishnan, 2012). In contrast to the rhodopsin, secretin, glutamate, and frizzled GPCRs, members of the aGPCRs were not discovered through their ligand molecules. Instead, they were identified in the late 1990s as a result of genetic approaches that became available through the development of cDNA cloning techniques in the 1980s.

Identification of the first three aGPCRs, ADGRL1-3/Latrophilins/CIRL (calcium-independent receptor of α-latrotoxin) (Krasnoperov et al., 1997; Lelianova et al., 1997), ADGRE5/CD97 (Hamann et al., 1995) and ADGRE1/EMR1 (EGF-like module-containing, mucin-like hormone receptor 1) (Baud et al., 1995; McKnight and Gordon, 1996) led to a surprise because they are chimeric proteins comprising seven-

[1] Text from this section was taken verbatim (with minor changes) from: Araç D and Leon K (2019). Structural Biology of Adhesion GPCRs. In P. Park and B. Jastrzebska (Eds.), *GPCRs – Structure, Function, and Drug Discovery*. Amsterdam: Elsevier. My contributions to this book chapter included writing and editing the text and figures.
[2] For more detailed description of aGPCRs, see Chapter 2.

transmembrane (7TM) domains similar to the conventional GPCRs, but also have extended N-terminal extracellular regions (ECRs) possessing several adhesion domains that are different in each receptor (Figure 1.3). The difficulty of working on these proteins are due to their large sizes and their unconventional properties left them unstudied for decades. Even genome-wide association studies were not pursued, as full-length genes were unfeasible to study. Their biological functions, structures, activation mechanisms, disease mutations and whether they in fact couple to G proteins remained unknown and the field lingered dormant.

Starting in the mid-2000s, several independent genetic studies suggested essential functions (such as brain development (Piao et al., 2004), Schwann cell myelination (Monk et al., 2009), central nervous system angiogenesis (Kuhnert et al., 2010) and tumorigenesis (O'Hayre et al., 2013)) for members of the aGPCR family. Although the lack of biochemical information about these proteins left researchers still unclear about their mechanisms of function, such findings revealed the urgent need for better understanding of this forgotten family of GPCRs. Soon in 2012, structural studies led to the discovery of a novel conserved extracellular domain of aGPCRs that was later revealed as a major player for receptor activation and opened the door for the mechanistic understanding of aGPCR activation (Arac et al., 2012). However, this discovery is only the start of the major goal of understanding the molecular mechanisms of aGPCRs, and further structural studies on the ECRs and 7TM domains are needed.

This thesis will describe the advances in our knowledge made through structural and functional characterizations of various aGPCRs. For more detailed descriptions of aGPCRs, see Chapter 2.

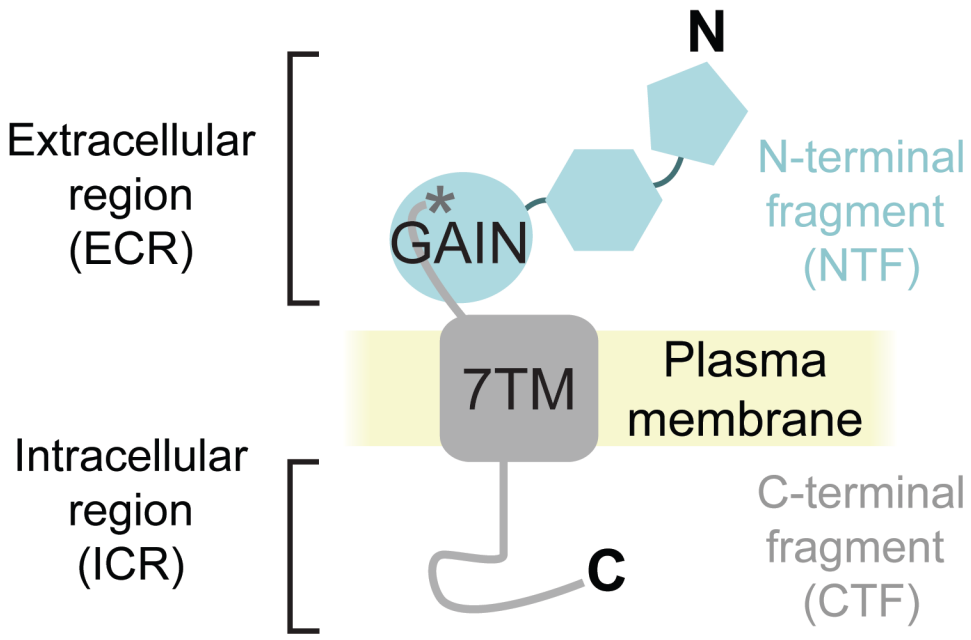


Figure 1.3 Domain structure of adhesion GPCRs

Cartoon representation of an aGPCR, which has an intracellular region (ICR), a seven-transmembrane region (7TM), and an extracellular region (ECR) consisting of variable adhesion domains including a GAIN domain. GAIN domain autoproteolysis results in a two noncovalently associated fragments: the N-terminal fragment (NTF, light blue) which includes the extracellular domains minus the *Stachel* peptide and the C-terminal fragment (CTF, gray) which includes the *Stachel* peptide, 7TM, and ICR.

Table 1.1 History of important discoveries in the aGPCR field

- The first aGPCR (EMR1) is described in 1981. (Austyn and Gordon, 1981)
- In 1995, CD97 (Hamann et al., 1995) and EMR1 (Baud et al., 1995) are the first aGPCRs to be cloned.
- In 1996, CD55 is the first ligand to be identified for an aGPCR (CD97). (Hamann et al., 1996)
- CD97 (Gray et al., 1996) and Lphn (Krasnoperov et al., 1997) are shown to be cleaved in their ECRs.
- The aGPCR family is identified as a member of the GPCR superfamily in 2003 (Fredriksson et al., 2003b).
- The mechanism of autoproteolytic cleavage of aGPCRs is described in 2004 (Lin et al., 2004).
- In 2012, the crystal structure of the GAIN domain is solved (Arac et al., 2012) leading to new models of adhesion GPCR signaling.

Latrophilins/ADGRL1-3[3]

Latrophilins (ADGRL1-3/Lphn1-3) were first identified in 1997 as calcium independent receptors for α-latrotoxin, a black widow spider toxin that triggers massive neurotransmitter release from nerve terminals (Deak et al., 2009; Krasnoperov et al., 1997; Lelianova et al., 1997; Sudhof, 2001; Sugita et al., 1999). Mutations of Lphns are linked to attention deficit hyperactivity disorder (ADHD) as well as numerous cancers in humans (Arcos-Burgos et al., 2010; Kan et al., 2010; O'Hayre et al., 2013). They are the

[3] Text from this section was taken verbatim (with minor changes) from: Araç D and Leon K (2019). Structural Biology of Adhesion GPCRs. In P. Park and B. Jastrzebska (Eds.), *GPCRs – Structure, Function, and Drug Discovery*. Amsterdam: Elsevier. My contributions to this book chapter included writing and editing the text and figures.

only aGPCRs besides flamingo-like CELSR proteins that are conserved between vertebrates and invertebrates. Lphn is required for tissue polarity during development in *C. elegans* (Langenhan et al., 2009), and for perceiving mechanical signals in *D. melanogaster* (Scholz et al., 2015). In vertebrates, Lphns mediate excitatory synapse formation through interactions involving their ECRs, which contain a lectin domain, an olfactomedin domain, a serine/threonine-rich region, a HormR domain and a GAIN domain (Arac et al., 2012; Krasnoperov et al., 1996; Lelianova et al., 1997; Sugita et al., 1998). The structures of all ECR domains of Lphns are currently available including the GAIN and HormR domains (Arac et al., 2012), the lectin domain (Vakonakis et al. 2008) and the lectin and olfactomedin domains (Jackson et al., 2015). The multi-domain ECR is a protein-protein interaction surface for numerous neuronal proteins, including FLRT proteins (O'Sullivan et al., 2012), teneurin/ODZ family proteins (Boucard et al., 2014; Levine et al., 1994; Silva et al., 2011), neurexins (Boucard et al., 2012) and α-Latrotoxin (Lelianova et al., 1997).

Despite these insights into Lphn biological functions and cell adhesion via ECR interactions, there is still a gap in our knowledge with respect to the signaling 7TM domains of Lphns. For example, we do not know which 7TM residues are important for receptor activation, how cancer mutations in the 7TM affect the ability for the receptor to signal properly, whether there are regulatory interactions between the ECR and 7TM, how the Lphn (and other aGPCR) 7TM domains differ from other GPCR 7TM domains, and more. Part of this thesis aims to answer these questions and focuses on the recent advances we have made on the structure and function of Lphns, which include

structural studies of teneurin (Chapter 5) and the development of G protein signaling assays to study Lphn function (Chapter 6).

GPR126/ADGRG6[4]

In 2003, GPR126/ADGRG6 was identified as an aGPCR due to its phylogenetic relationships with previously-identified aGPCRs (Fredriksson et al., 2003a). Little was known about this receptor except that it has a large serine/threonine-rich N-terminal region and that it contains a GPCR proteolytic site, similar to other aGPCRs. However, studies that followed shortly after showed that GPR126 contains C1r-C1s, Uegf, and Bmp1 (CUB) and pentraxin (PTX) domains in its N-terminal ECR (Moriguchi et al., 2004; Stehlik et al., 2004). In addition, GPR126 was found to be expressed in human umbilical vein endothelial cell culture as well as in the developing mouse heart and adult mouse lung. Interestingly, GPR126 was also found to proteolytically cleaved by furin in its ECR at a second site distinct from the already-known GPCR proteolytic site (Moriguchi et al., 2004). These new pieces of information suggested that the ECR of GPR126 likely plays an important role in receptor function and that GPR126 may be involved in a diverse set of biological processes.

In humans, GPR126 mutations are linked to several cancers including bladder cancer, breast cancer, and lung cancer (Garinet et al., 2018; Nik-Zainal et al., 2016; Piraino and Furney, 2017; Qin et al., 2017; Terzikhan et al., 2018), and other diseases

[4] Nomenclature is as follows:
Protein: GPR126 (human), Gpr126, (zebrafish)
Gene: *GPR126* (human), *Gpr126* (zebrafish)

including adolescent idiopathic scoliosis (Kou et al., 2013) and arthrogryposis multiplex congenita, a disorder characterized by multiple joint contractures (Ravenscroft et al., 2015).

The first major biological relevance of Gpr126 was discovered in a forward genetic screen of zebrafish with abnormally myelinated axons (Monk et al., 2009). Schwann cells in *Gpr126* mutant zebrafish were not able to express *oct6* and *krox20*, genes essential for myelination, and were arrested at the promyelinating stage. Treatment with forskolin, which mimics Gpr126 signaling, rescued the mutant phenotype, which suggests that Gpr126 G-protein signaling regulates Schwann cell myelination. Further studies showed that this regulatory function is conserved in mammals and that association with endogenous ligands such as type IV collagen, laminin-211, and the prion protein are important for Gpr126 function (Kuffer et al., 2016; Monk et al., 2011; Paavola et al., 2014; Petersen et al., 2015). In addition to myelination, Gpr126 has also been found to be important for development of the heart and ears, and interestingly, heart development has been shown to only require the ECR of Gpr126 and not the 7TM domain (Geng et al., 2013; Patra et al., 2013; Waller-Evans et al., 2010). This suggests that Gpr126 regulates several functions, perhaps through distinct molecular mechanisms involving different parts of the receptor.

To better understand the molecular mechanisms underlying these biological functions, our approach is to study the structure and function of Gpr126. In particular, we focused on the ECR since there is now much evidence suggesting that ECRs play important regulatory roles in aGPCR functions. Prior to our studies, the domains that make up the ECR were not fully characterized, the structural context of the furin

proteolysis site was unknown, and the ECR domain architecture was unclear. Chapter 3 focuses on structural and functional studies of Gpr126, which includes the high-resolution crystal structure of the full ECR, leading to the development of a model for Gpr126 function.

Chapter 2: Structural Biology of Adhesion GPCRs[1]

Biological functions of adhesion GPCRs

Cells in multicellular organisms have the extraordinary ability of adhering to each other and exchanging information. The interplay of cellular adhesion and cellular signaling is essential for the development of all organs. Extracellular adhesion coupled to intracellular signaling is a key phenomenon that is disrupted in many human diseases (Asherson and Gurling, 2012; Sudhof, 2008). However, the underlying mechanisms of such complex phenomena are unclear. aGPCRs are suggested to mediate intercellular communication via cell-cell and cell-matrix interactions.

Overwhelming data show the essential roles of aGPCRs in numerous cellular functions. For example, ADGRG1/GPR56 has been shown to have critical roles in brain development (Bae et al., 2014; Piao et al., 2004). Mutations in GPR56 cause bilateral frontoparietal polymicrogyria (Chiang et al., 2011), which is characterized by abnormal cortical folding and layering. The latrophilins (Lphns), through interaction with protein partners such as Fibronectin leucine-rich repeat transmembrane (FLRT) (O'Sullivan et al., 2012) and teneurin, have also been shown to be important for synapse development and maintenance (Boucard et al., 2014; O'Sullivan et al., 2012). Lphn variants have been linked to susceptibility of attention-deficit/hyperactivity disorder (Arcos-Burgos et

[1] Text from this chapter was taken verbatim from: Araç D and Leon K (2019). Structural Biology of Adhesion GPCRs. In P. Park and B. Jastrzebska (Eds.), *GPCRs – Structure, Function, and Drug Discovery*. Amsterdam: Elsevier. My contributions to this book chapter included writing and editing the text and figures.

al., 2010). ADGRG6/GPR126 has been shown to be an essential regulator of myelination in the peripheral nervous system (Monk et al., 2009). GPR126 mutations have been shown to cause arthrogryposis multiplex congenita (Ravenscroft et al., 2015), characterized by multiple joint contractures. ADGRA2/GPR124 regulates central nervous system angiogenesis (Kuhnert et al., 2010) and has recently been shown to form higher order complexes with Reck, Wnt7, and Frizzled to activate Wnt signaling (Eubelen et al., 2018; Zhou and Nathans, 2014). Other receptors have additional roles in regulating synapses (Bolliger et al., 2011) and neural tube development (Chae et al., 1999; Langenhan et al., 2009; Shima et al., 2004; Usui et al., 1999).

Additionally, many aGPCRs are found to be overexpressed or underexpressed in various cancers (Kan et al., 2010; Shashidhar et al., 2005; Xu et al., 2006). For example, ADGRG2/GPR64 is overexpressed in parathyroid tumors (Balenga et al., 2017), ADGRE2/EMR2 is highly expressed in breast cancer cells (Safaee et al., 2014), ADGRF3/GPR113 is upregulated in neuroendocrine tumors (Carr et al., 2012), and ADGRG5/GPR114 is overexpressed in acute myeloid leukemia (Maiga et al., 2016). Moreover, a recent study reports aGPCRs as some of the most mutated genes in cancers (O'Hayre et al., 2013). For example, ADGRD2/GPR144 is highly mutated in lung cancer (Yu et al., 2015). A cancer-associated point mutation in Lphn3 was reported to dramatically increase constitutive signaling (Nazarko et al., 2018). Considering that many drugs target the 7TM domain of GPCRs to regulate receptor activity and have excellent therapeutic benefits, aGPCRs are promising targets for drugs to treat numerous diseases.

Table 2.1 List of important biological functions and associated diseases of adhesion GPCRs

(continued on following page)

Adhesion GPCR	Known function or related disease
Latrophilin 1-3	ADHD (Arcos-Burgos et al., 2010), synaptic adhesion (Boucard et al., 2014; O'Sullivan et al., 2012), tissue polarity (Langenhan et al., 2009), receptor for a-latrotoxin (Krasnoperov et al., 1997; Lelianova et al., 1997), cancer (Kan et al., 2010; O'Hayre et al., 2013)
GPR56	Brain development (Piao et al., 2004), cortical patterning (Bae et al., 2014), brain cancer (Xu et al., 2006; Yang et al., 2011), BFPP (Chiang et al., 2011)
BAI 1-3	Synapse maturation/elimination (Bolliger et al., 2011; Kakegawa et al., 2015), schizophrenia (DeRosse et al., 2008), cancer (Cork and Van Meir, 2011)
GPR126	Myelination of Schwann cells (Monk et al., 2009), heart development (Patra et al., 2013), arthrogryposis multiplex congenita (Ravenscroft et al., 2015), adolescent idiopathic scoliosis (Kou et al., 2013; Liu et al., 2018)
GPR124	Angiogenesis in the CNS (Cullen et al., 2011; Kuhnert et al., 2010), Wnt signaling (Eubelen et al., 2018; Zhou and Nathans, 2014), colon cancer (Oliveira et al., 2018), hypertension (Calderon-Zamora et al., 2017)
ETL	Tumor angiogenesis (Masiero et al., 2013), cardiovascular development (Lu et al., 2017)
GPR116	Lung surfactant homeostasis (Fukuzawa et al., 2013; Yang et al., 2013), cardiovascular development (Lu et al., 2017), breast cancer (Tang et al., 2013)
GPR64	Male fertility (Zhang et al., 2018a), parathyroid hormone release and cancer (Balenga et al., 2017)
CELSR 1-3/ Flamingo	Neural tube defects (Curtin et al., 2003), planar cell polarity (Chae et al., 1999; Usui et al., 1999)
EMR1-3	Neutrophil activation (Yona et al., 2008b), cancer (Safaee et al., 2014)

Table 2.1, continued

CD97	Neutrophil migration (Leemans et al., 2004), cancer (Safaee et al., 2013)
GPR123	Expressed in CNS (Lagerstrom et al., 2007)
GPR125	Planar cell polarity (Li et al., 2013), brain injury response (Pickering et al., 2008)
GPR133	Glioblastoma growth (Bayin et al., 2016)
GPR144	Lung cancer (Yu et al., 2015)
GPR115	Lung cancer (Ozer and Sezerman, 2017; Zhang et al., 2018b)
GPR111	Unknown function
GPR110	Brain development (Lee et al., 2016), lung and prostate cancer (Lum et al., 2010), breast cancer (Bhat et al., 2018)
GPR113	Neuroendocrine cancer (Carr et al., 2012)
GPR128	Deletion leads to weight loss and increased intestinal contraction frequency in mice (Ni et al., 2014)
GPR114	Overexpressed in acute myeloid leukemia (Maiga et al., 2016)
GPR97	B-cell development (Wang et al., 2013b), bone formation (Prashar et al., 2014), lymphatic remodeling (Valtcheva et al., 2013), autoimmune CNS disease (Wang et al., 2018)
GPR112	Marker for neuroendocrine carcinoma (Leja et al., 2009)
VLGR1	Usher syndrome (Weston et al., 2004), hair cell development (McGee et al., 2006), epilepsy (Myers et al., 2018; Shin et al., 2013)

7TM domain

Of the five GPCR families, aGPCRs are the second largest but the least well-understood. Despite the current lack of structural information about the 7TM domains of aGPCRs, phylogenetic analysis reveals that aGPCRs share the highest sequence identity with the secretin family because the secretin family descended from the

adhesion family (Nordstrom et al., 2009). Recent structural studies on secretin GPCRs reveal that unlike the rhodopsin GPCRs, their 7TM domains are more open on the extracellular surface, which allows for binding of peptide hormone ligands (de Graaf et al., 2017). aGPCRs are predicted to also have a more open 7TM domain, which would accommodate the binding of their proposed peptide or peptide-like endogenous ligands. Conserved motifs that are important for receptor activation have been identified for several GPCR families. For example, NPxxY in TM7, DRY in TM3, and CWxP in TM6 are motifs essential for ligand-induced activation of rhodopsin GPCRs (Audet and Bouvier, 2012). Although aGPCRs lack the conserved motifs found in rhodopsin GPCRs, the GWGxP motif in TM4 (Bortolato et al., 2014), which plays important structural roles in secretin GPCRs, is common to some aGPCRs as well.

Recent studies have aimed to determine whether aGPCRs share common activation mechanisms with rhodopsin and secretin GPCRs. In yeast solid growth assays and Serum Response Element (SRE) luciferase reporter assays, the 7TM domain of ADGRG4/GPR112 was found to retain functionally important motifs analogous to those found in rhodopsin and secretin receptors as well as distinct residues that are essential in aGPCR-specific activation (Peeters et al., 2016). For example, $His^{3.33/3.37b}$, which is conserved among all aGPCRs, leads to decreased signaling when mutated and likely plays a role in stabilizing an extracellular loop.

Another recent study aimed to identify important residues/motifs involved in the signaling activity of the aGPCR Lphn1 (Nazarko et al., 2018) through a comprehensive mutagenesis analysis using a cAMP assay and a SRE luciferase reporter assay. Despite the lack of conservation with the rhodopsin GPCRs, the Lphn1 residues that

reside in the analogous positions are important for signaling. For example, in Lphn1, the NPxxY motif is replaced with IFVFH, which when mutated, alters basal activity. Moreover, mutation of an HLY motif which replaces the rhodopsin DRY motif, also affects basal activity of the receptor. The results suggest that these residues likely play important roles in stabilizing Lphn1 in a certain conformation and, when mutated, the conformational landscape is changed and results in altered signaling. Despite the divergence in sequence between rhodopsin and aGPCR 7TM domains, aGPCRs likely have a similar “skeleton” and function similarly to rhodopsin GPCRs with some inherent differences, consistent with the conclusions from the above GPR112 study.

Adhesion GPCRs have large extracellular domains

Unlike other GPCRs, aGPCRs have long N-terminal ECRs that can be up to 5000 residues and comprise various adhesion domains (Yona et al., 2008a). These domains are specific to each aGPCR (except for the conserved GAIN domain discussed below) and are likely involved in adhering to various ligands (Arac et al., 2012). Among the domains are cadherin; leucine-rich repeat (LRR); epidermal growth factor (EGF); thrombospondin (TSP); pentraxin/laminin/neurexin/sex-hormone-binding-globuilin-like (PLL); sperm protein, enterokinase, and agrin (SEA); pentraxin (PTX); complement C1r/C1s, Uegf, Bmp1 (CUB); hormone-binding (HormR); lectin and others. Often, there are multiple copies of the same domain or motif in a receptor’s ECR, further increasing the size of the receptor. For instance, the longest isoform of Very Large GPCR (ADGRV1/VLGR1/GPR98) has 35 repeats of the Calx-ß motif, 7 repeats of epitempin (EPTP) or epilepsy associated repeat (EAR), a PTX domain, and a GAIN domain in its

ECR (McMillan and White, 2010). The known adhesion domains for each aGPCR is depicted in Figure 2.1.

Most aGPCRs are orphan receptors with no known endogenous ligands. However, several aGPCRs are known to bind to other proteins, but whether these binding partners are true ligands that activate downstream signaling pathways or simply bind to the adhesion domains is unclear (Paavola and Hall, 2012), and the role of ligand/ECR interaction in aGPCR function is currently under intense investigation. Most aGPCRs also have large cytoplasmic sequences that may be involved in intracellular interactions in addition to trimeric G-proteins. For example, PDZ-binding motifs have been reported at the C-termini of the cytoplasmic tails of ADGRD1/GPR133, ADGRA2/GPR124, ADGRA3/GPR125, ADGRB1-3/BAI1-3, and VLGR1. (Simundza and Cowin, 2013)

A hallmark of aGPCRs that was discovered in the early days of the field is autoproteolytic cleavage within the ECR at a conserved GPCR Proteolysis Site (GPS) (Chang et al., 2003; Ichtchenko et al., 1999; Krasnoperov et al., 1997; Lin et al., 2004). The GPS is well described and recognized as a vital unit for receptor function (Gray et al., 1996; Shashidhar et al., 2005; Stacey et al., 2002). Autoproteolysis cleaves the receptor into two parts: a large N-terminal fragment (NTF) and a membrane-anchored C-terminal fragment (CTF) comprising the 7TM domain and the cytoplasmic tail (Krasnoperov et al., 1997; Lelianova et al., 1997). Autoproteolysis occurs during receptor maturation in the endoplasmic reticulum and is catalyzed by a reaction similar to N-terminal nucleophilic hydrolases (Lin et al., 2004). Interestingly, unlike other autoproteolytic proteins, the cleaved fragments remain non-covalently but tightly

associated with each other as a heterodimer raising the question of why the receptor cleaves itself if the cleavage products do not separate.

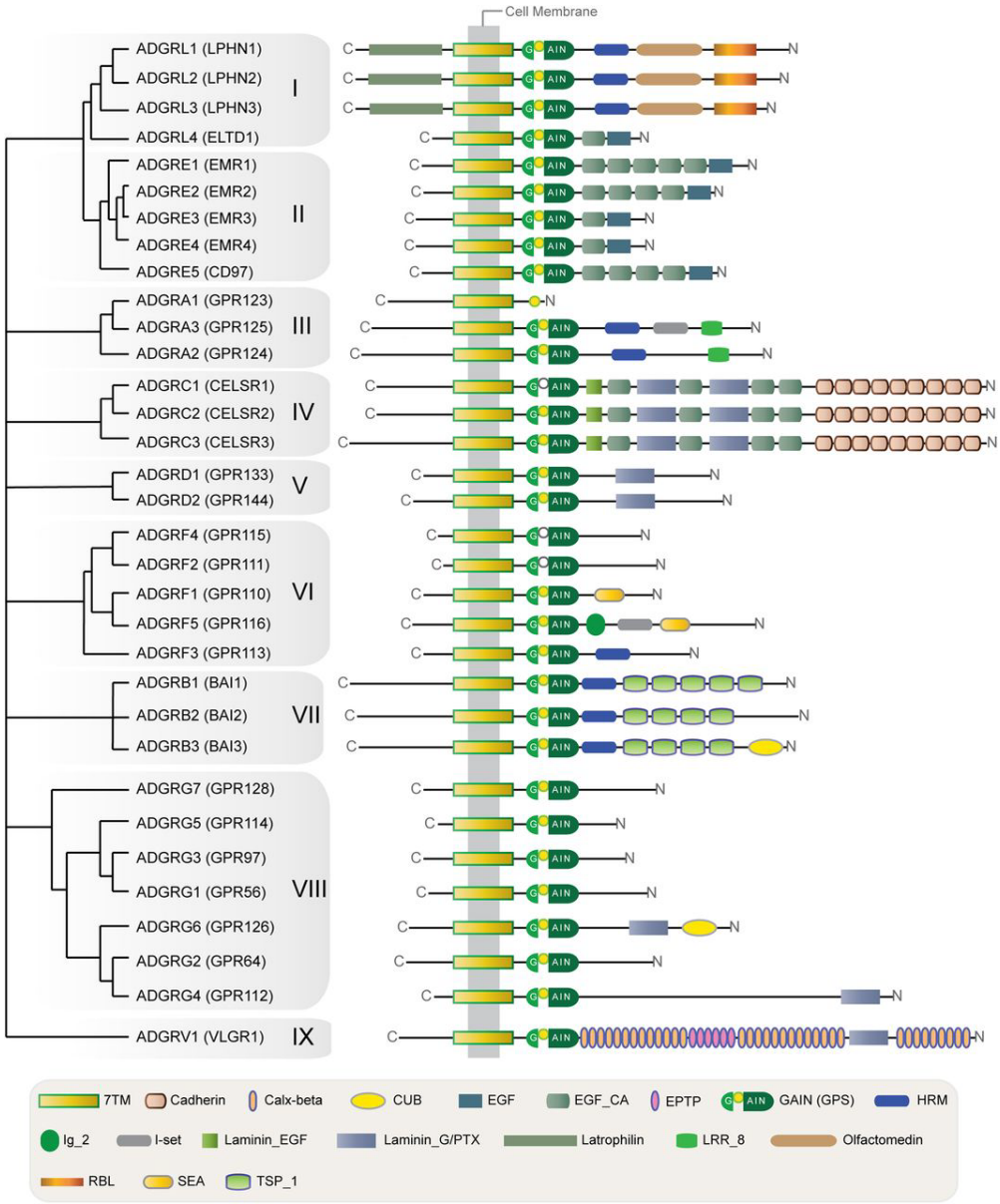


Figure 2.1 Extracellular domain organization of adhesion GPCRs

The 33 aGPCRs found in humans organized by subfamily. The ECRs consist of unique combinations of adhesion domains, indicated by the representative domain symbols. Figure adapted from (Hamann et al., 2015).

A conserved domain in all aGPCR extracellular regions

Previously, it was believed that aGPCRs have a non-functional and unstructured so-called “stalk” region that precedes the GPS motif. However, computational analysis suggested that the stalk region was conserved and structured. Surprisingly, the crystal structures from two unrelated aGPCRs revealed that the GPS motif itself does not constitute an autonomously folded unit. Instead, it is an integral part of a much larger novel domain that spans the GPS motif and the stalk region and is termed the GPCR-Autoproteolysis Inducing (GAIN) domain, resolving the previously-held conflicting theories (Arac et al., 2012).

Strikingly, the GAIN domain is shared by 32 out of 33 human aGPCRs (except ADGRA1/GPR123 which has only a short sequence of ~20 residues in its ECR). The GAIN domain is unique in that it is the only domain that exists in all members of the aGPCRs in humans, indicating an essential role in aGPCR function. Moreover, database searches revealed that all five members of another human protein family, the polycystic kidney disease proteins, an unrelated family of mechanosensory membrane proteins that are also autoproteolysed, contain GAIN domains. In addition, primitive organisms, such as *Dictyostelium discoideum* that arose early in evolution before animals emerged, encode GAIN domains although they lack most other autoproteolytic domains, important adhesion and signaling domains and critical signaling pathways (Arac et al., 2012). Among all domains that are found in aGPCRs, LRR, EGF and TSP domains are the only other domains that exist in such primitive organisms. These results show that the GAIN domain is widespread and conserved in higher eukaryotes as well as in ancient organisms. Intriguingly, analysis of numerous sequences also

revealed that the GAIN domain always immediately precedes the first TM helix by a short linker raising the question of whether the GAIN domain regulates receptor signaling via the nearby 7TM domain.

Table 2.2 Hallmarks of the GAIN domain

- The GAIN domain is an ancient domain that is conserved from primitive evolutionarily distant organisms such as *Dictyostelium discoideum* and *Tetrahymena thermophila* to higher eukaryotes.
- Mutations in the GAIN domain are associated with diseases such as cancer, autosomal dominant polycystic kidney disease (ADPKD), and bilateral frontoparietal polymicrogyria (BFPP).
- It directly precedes the TM domain in all polycystic kidney disease-1 (PKD1) proteins and aGPCRs except GPR123, which lacks the GAIN domain.
- The GAIN domain can vary in size between aGPCRs.
- The GAIN domain is both necessary and sufficient for autoproteolysis at the conserved GPCR proteolysis site (GPS).
- The two fragments do not immediately dissocate following autoproteolysis and likely needs to be disrupted through mechanical force.
- Dissociation of the two fragments reveals the *Stachel* peptide, which can bind to and activate the receptor.

Structure of the GAIN domain

The crystal structures of GAIN domains from two distantly related aGPCRs, ADGRL1/Lphn1 (Figure 2.2A) and brain angiogenesis inhibitor 3 (ADGRB1/BAI3) (Figure 2.2B) revealed a novel fold that was previously unidentified (Arac et al., 2012). The GAIN domain (of Lphn1) contains an N-terminal subdomain A that is composed of six alpha helices, and a C-terminal subdomain B that is composed of a twisted β-

sandwich including thirteen β-strands and two small α-helices. The last five β-strands of subdomain B constitute the GPS motif. The GPS motif is the most conserved region of the GAIN domain. The conservation of primary sequence of the GAIN domains decreases from the C-terminus to the N-terminus. In spite of the low sequence identity between Lphn1 and BAI3 GAIN domains (24%), the very high similarity of the GAIN domain structures indicates that the three-dimensional structure is conserved more strictly than primary sequence throughout evolution. The recent structure of the ADGRG1/GPR56 GAIN domain shows that the size of the GAIN domains can be smaller but still be autoproteolysed (Figure 2.2C) (Salzman et al., 2016).

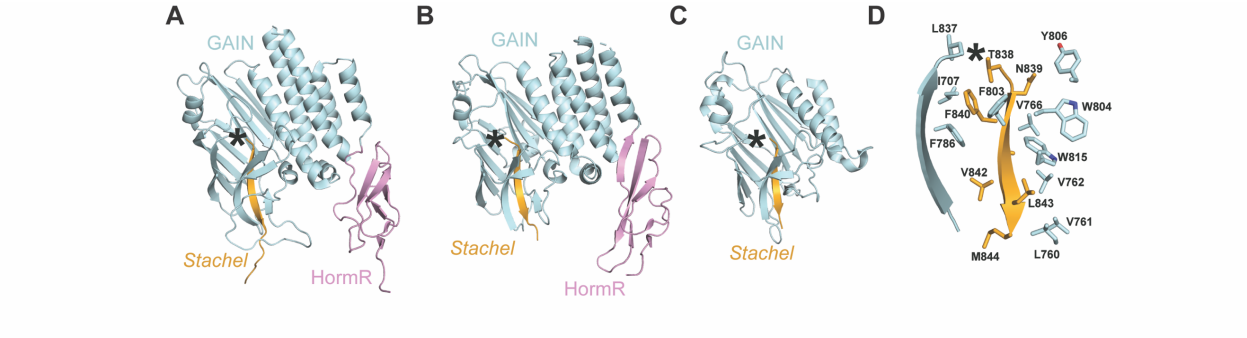


Figure 2.2 Structures of GAIN domains from various adhesion GPCRs

(A, B) Cartoon representation of the GAIN (cyan) and HormR (pink) domains from Lphnn1 (PDB: 4DLQ) (A) and BAI3 (PDB: 4DLO) (B). (C) Cartoon representation of the GAIN domain from GPR56 (PDB: 5KVM). (D) Hydrophobic interactions between the cleaved *Stachel* peptide and surrounding residues in LPHN1. The *Stachel* peptide is colored gold and autoproteolysis site is indicated by an asterisk (*).

Autoproteolysis cleaves off a short peptide but does not release it from the GAIN domain

The GAIN domain is an autoproteolytic fold that is both required and sufficient for autoproteolysis, whereas the GPS motif without the rest of the GAIN domain is not

functional. Autoproteolysis occurs in the short and kinked loop between the last two β-strands of the GAIN domain and cleaves the C-terminal β-strand from the rest of the domain. Thus, the two autoproteolysis products are: a large N-terminal fragment (NTF) comprising various adhesion domains and the GAIN domain that misses its last β-strand and a membrane-anchored C-terminal fragment (CTF) comprising the 7TM domain tethered to the missing β-stand and the cytoplasmic tail (Krasnoperov et al., 1997; Lelianova et al., 1997). However, the GAIN domain structures and biochemical studies agree that the cleaved peptide does not dissociate from the rest of the GAIN domain even after autoproteolysis (Figure 2.2D). The cleaved peptide is highly conserved and has a hydrophobic nature (TNFAVLM in Lphn1) and is involved in approximately 15 strong hydrogen bonds and numerous hydrophobic interactions that keep it tightly bound within the GAIN domain. Indeed, a critical piece of information that was revealed from the GAIN domain structures was that the cleaved peptide is “hidden” within the GAIN domain and is inaccessible to the solvent. The crystallographic B-values that represent the displacement of atoms in the cleaved peptide show that the cleaved peptide is well structured and is not flexible within the GAIN domain. The large number of hydrogen bonds and hydrophobic interactions suggest that a large force needs to be applied for the peptide to dislocate from the rest of the GAIN domain, and that once the cleaved peptide is removed from the hydrophobic core of the GAIN domain, it cannot be positioned back into the GAIN structure.

Stachel peptide mediated activation of adhesion GPCRs

The discovery of the GAIN domain was a breakthrough as it redefined the poorly studied aGPCR class, showing that members of this family share a large, unique and widespread autoproteolytic domain that may be involved in downstream signaling (Arac et al., 2012). Most importantly, the observation that the cleaved β-strand (hereafter termed *Stachel* peptide) of the GAIN domain is hidden within the GAIN domain raised questions about the consequences of exposing the *Stachel* peptide. Previous studies in 2011 had shown that truncation constructs starting from the first residue after the autoproteolysis site dramatically increased signaling and hinted that the ECR is somehow involved in receptor signaling (Paavola et al., 2011). These studies suggested that the ECR either has a direct inhibitory effect on 7TM signaling, or indirectly inhibits receptor signaling by preventing the action of an agonist. Other studies have also proposed that GPCR ECRs regulate receptor functions, likely including G-protein signaling, by binding to extracellular ligands (Booe et al., 2015; Byrne et al., 2016; Chiang et al., 2016; Coin et al., 2013; Luo et al., 2014; O'Sullivan et al., 2012; Petersen et al., 2015; Zhao et al., 2016). Currently, two complementary models are proposed for the ligand-induced activation of aGPCRs (Figure 2.3) and the validity of other possible models are still being investigated.

Independent studies from two groups showed that aGPCRs are activated by the *Stachel* peptide (also called tethered agonist or stalk) (Liebscher et al., 2014; Stoveken et al., 2015). Upon autoproteolysis, the *Stachel* peptide, which corresponds to the last β-strand of the GAIN domain, is located at the newly generated N-terminus of the 7TM domain (Figure 2.3). Addition of synthesized *Stachel* peptide on aGPCRs led to

increased signaling, suggesting a mechanism whereby the *Stachel* peptide acts as a tethered agonist that activates the receptor (Liebscher et al., 2014; Stoveken et al., 2015). Mutagenesis studies showed that the residues at the more N-terminal tail of the *Stachel* peptide (closer to the autoproteolysis site) are critical for receptor activation, reminiscent of the protease-activated receptors and secretin receptors that are also activated by short peptides. From these studies, a *Stachel*-mediated model emerged, in which the NTF serves as a protective cap for the *Stachel* and has no direct role in modulating 7TM function. Upon ligand binding to an N-terminal adhesion domain, the NTF dissociates from the CTF, termed “shedding”, exposing the *Stachel* peptide. (Demberg et al., 2015; Kishore et al., 2016; Liebscher et al., 2014; Stoveken et al., 2015). The *Stachel* peptide, which is highly hydrophobic, binds to the 7TM domain and activates the receptor. Direct interaction between the *Stachel* peptide and the 7TM was reported (Nazarko et al., 2018) and a mutagenesis screen of Lphn1 revealed residues important for the response of the receptor to the *Stachel* peptide. These results support the model of an outwardly open 7TM similar to that of secretin receptors (de Graaf et al., 2017). Key to the *Stachel*-mediated model of aGPCR activation is *Stachel* exposure. Although it has been proposed that natural ligands may induce NTF shedding upon binding to N-terminal adhesion domains, and thereby activate the receptor, direct proof of ligand-induced shedding remains elusive.

The location of the *Stachel* sequence within the GAIN domain strongly suggests that NTF shedding is autoproteolysis-dependent and irreversible. In the three-dimensional structure, the hydrophobic *Stachel* is buried within the hydrophobic core of the GAIN domain and forms extensive hydrogen-bond networks with adjacent β-strands

(Arac et al., 2012; Salzman et al., 2016). This architecture suggests that exposure of the *Stachel* requires substantial deformation of the GAIN domain. Furthermore, because the *Stachel* is a central part of the GAIN domain, the release of the *Stachel* from the GAIN domain most likely leads to a collapse of the original conformation, prohibiting re-association of the *Stachel* with the NTF, meaning that *Stachel*-dependent activation generates a receptor that cannot be reused. Moreover, because the N-terminal residues of the *Stachel* are the most deeply buried within the GAIN domain (Arac et al., 2012; Salzman et al., 2016), yet are critical for mediating receptor activation (Liebscher et al., 2014; Stoveken et al., 2015), transient exposure of the *Stachel* to activate the 7TM without causing irreversible NTF-CTF dissociation should be an extremely rare event.

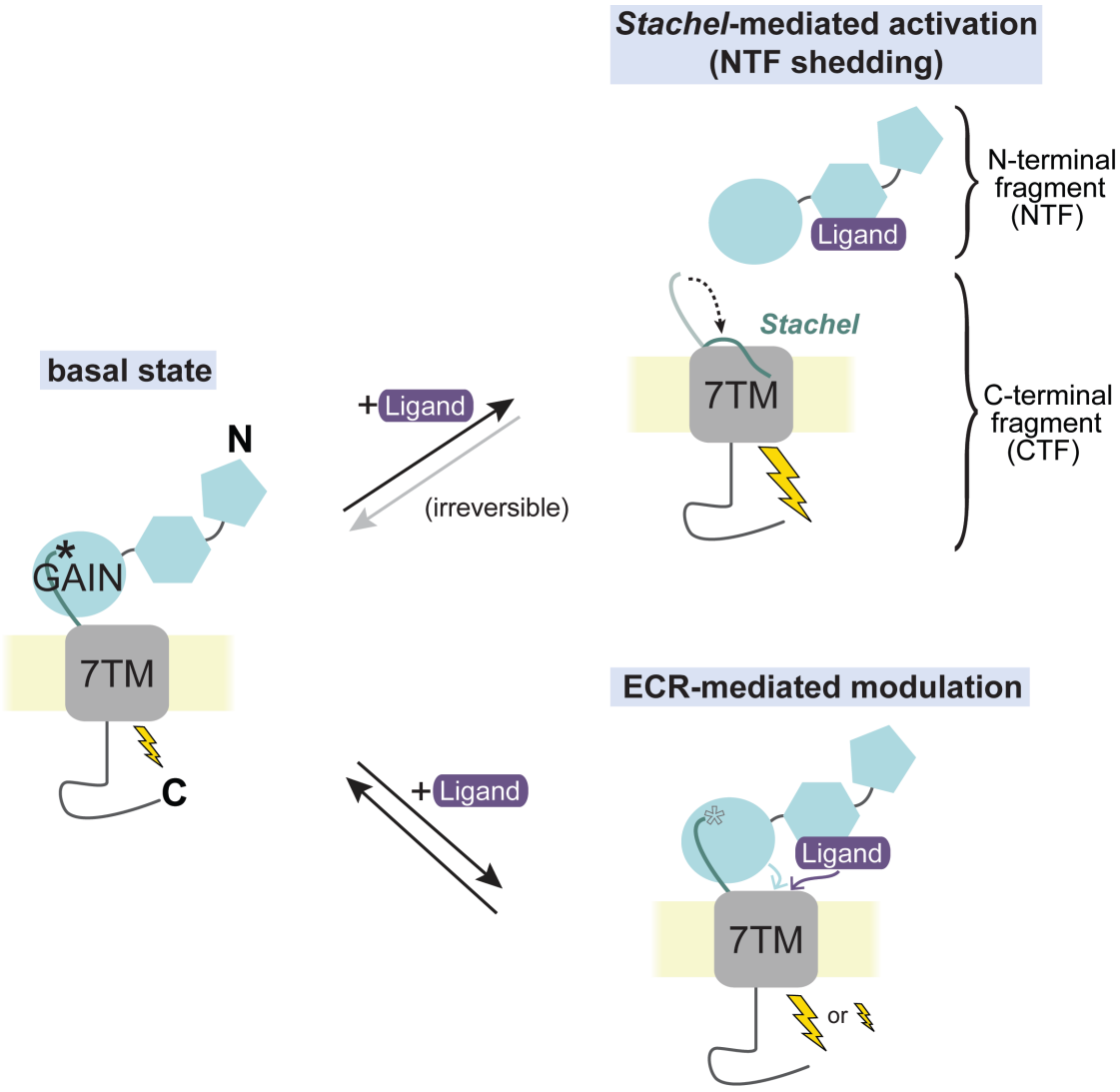


Figure 2.3 *Stachel*-dependent and *Stachel*-independent models for regulation of adhesion GPCR signaling

Stachel peptide is colored dark green and autoproteolysis site is indicated by an asterisk (*). Light blue and purple arrows indicate regulation of 7TM signaling by the ECR. Size of lightning bolt represents relative signaling level.

ECR-mediated activation of adhesion GPCRs (*Stachel*-independent activation)

The biological relevance of *Stachel*-mediated aGPCR activation is clear (Chiang et al., 2017; Luo et al., 2014; Petersen et al., 2015). However, unanswered questions and recent observations have necessitated the introduction of the complementary

Stachel-independent model (Kishore et al., 2016). For example, overexpression of autoproteolysis-deficient lat-1/ADGRL1 in lat-1-knockout *C. elegans* rescues the wildtype phenotype, suggesting that some aGPCR functions do not require autoproteolysis (Promel et al., 2012a). Additionally, there are several aGPCRs that lack the conserved residues critical for autoproteolysis and therefore, remain uncleaved (Arac et al., 2012; Promel et al., 2012b). Furthermore, several aGPCRs, including ADGRG1/GPR56, are found partially uncleaved *in vivo* (Arac et al., 2012; Iguchi et al., 2008; Promel et al., 2012b). For example, GPR56 in skeletal muscle was found to be almost completely uncleaved (Iguchi et al., 2008), although it plays critical roles in skeletal muscle cells (White et al., 2014; Wu et al., 2013). Together, these observations suggest that *Stachel*-independent mechanisms may play important roles in aGPCR signaling.

The involvement of ECRs in aGPCR functions was previously proposed. In the ECR-mediated model, the ECR (i.e., associated NTF and *Stachel*) has a direct role in modulating the 7TM signaling (Kishore et al., 2016; Ohta et al., 2015; Paavola et al., 2011; Salzman et al., 2016). Regulation by this mechanism is independent of *Stachel*-mediated activation, although the *Stachel* residues are present within the core of the GAIN domain. In this model, the ECR directly communicates with the 7TM (i.e., via transient interactions), such that ligand binding events or conformational changes in the ECR may directly result in altered signaling, either inhibitory or activating (Figure 2.3). In recent studies, Kishore et al. (Kishore et al., 2016) and Kishore and Hall (Kishore and Hall, 2017) measured the basal activities of GPR56 constructs with various ECR truncations through multiple signaling pathways. Using an SRE luciferase assay, a

construct lacking the NTF (i.e., 7TM with exposed *Stachel*) had the highest activity, whereas one lacking both the NTF and the *Stachel* (i.e., just the 7TM) had the lowest activity (Kishore et al., 2016), confirming the agonistic function of *Stachel* on the 7TM. In comparison, the full-length constructs of both the wild-type and an autoproteolysis-deficient mutant exhibited a moderate level of activity (Kishore et al., 2016), suggesting that the ECR modulates 7TM signaling.

In a different study, synthetic protein ligands (called monobodies) targeted to the ECR of human GPR56 were engineered and several allosteric agonists and allosteric inverse-agonists were obtained, showing that GPR56 signaling can be modulated by targeting its ECR (Salzman et al., 2017). The activity of these synthetic ligands on an autoproteolysis-defective and thus, shedding-defective receptor, was tested and the results showed that autoproteolysis is not required for each of these functional monobodies to modulate signaling. Furthermore, to test if the monobodies functioned in an autoproteolysis-independent but *Stachel*-dependent manner, a highly conserved *Stachel* residue, previously shown to be critical for *Stachel*-mediated activation of several aGPCRs, including GPR56 (Liebscher et al., 2014; Stoveken et al., 2015) was mutated and it was shown that the *Stachel* mutation did not abolish monobody-mediated modulation of GPR56 signaling, suggesting that neither autoproteolysis nor *Stachel*-mediated activation are required for the monobody-mediated modulation of GPR56 signaling. These results strongly suggest that perturbations to the ECR are directly sensed by the 7TM, resulting in altered signaling without NTF shedding and *Stachel* exposure.

The structure of the GPR56 ECR in complex with an allosteric modulator

A recent study reported the crystal structure of the ECR of GPR56 in complex with an inhibitory monobody (termed α5) and provided a functional framework to understand the molecular mechanisms by which aGPCR ECRs govern receptor function; excitingly, the structure represents the first crystal structure of a full ECR for any aGPCR (Figure 2.4). The structure revealed the identity and boundaries of two extracellular domains: a previously unidentified domain with a β-sandwich architecture at the N-terminus and, as predicted from the sequence, a GAIN domain at the C-terminus. The previously unidentified N-terminal domain has low homology to all known folds and was termed pentraxin/laminin/neurexin/sex-hormone-binding-globuilin-like (PLL) domain due to its remote similarity to well-studied domains that are involved in adhesion functions. This remote similarity supports an adhesion-related role for the PLL domain in GPR56 as was previously suggested (Koirala et al., 2009). The structure also revealed that the monobody interacts with the N-terminal PLL domain and GAIN domain simultaneously. The entire ECR is necessary for α5 binding, indicating that any effect on GPR56 activity mediated by α5 is due to its interaction with the ECR. Therefore, α5 represents an ‘allosteric inverse-agonist’ for GPR56 (Christopoulos, 2014).

A highly conserved interdomain disulfide bond links the PLL and GAIN domains. This disulfide bond may restrict the movements of the two domains with respect to each other as observed in other proteins with interdomain disulfide bonds (Bustanji and Samori, 2002). The 15-residue linker between the two GPR56 domains is ordered in the crystal, despite its lack of defined secondary structure. An interface between the PLL and GAIN is composed of mostly conserved and hydrophobic residues and has a large

buried surface area of ~680Å^2. It is likely that the interdomain disulfide bond and the hydrophobic residues are sufficient to stabilize this conformation of the ECR. These observations raised the possibility that aGPCR ECRs might exhibit rigid conformational states within different domains and the 7TM. Binding of natural ligands or synthetic proteins such as monobodies to the ECR might alter such conformations and transduce the signal to the 7TM domain (Figure 2.4).

Notably, the entire PLL domain is deleted in an alternatively spliced variant of GPR56 (S4), but not the other variants. Surprisingly, *in vitro* signaling assays showed that deletion of this domain increased basal activity of the receptor. This result suggests that the S4 isoform that lacks the PLL domain may have increased basal G protein signaling (Figure 2.4). This observation suggests that by regulating alternative splicing, a cell may generate GPR56 with or without a PLL domain in the ECR, which could diversify functionality. Moreover, eliminating the conserved disulfide bond between the PLL and GAIN domains in full-length GPR56 also increased GPR56 basal activity, suggesting the importance of restricting the flexibility within the ECR in keeping the receptor in the basal state. These results further support the ECR-mediated regulation of aGPCR function.

Finally, a highly conserved, surface-exposed patch was identified on the PLL domain. Mutagenesis of a histidine to alanine within this conserved patch on the GPR56 surface abolished previously-reported central nervous system (CNS) myelination function of GPR56 in zebrafish, suggesting an essential role of this evolutionarily conserved residue in CNS myelination and raised the possibility that the conserved patch is a ligand binding site (Figure 2.4). Together these results elucidated the

multifaceted manner by which the ECR regulates aGPCR function and broadened our understanding of aGPCR biology.

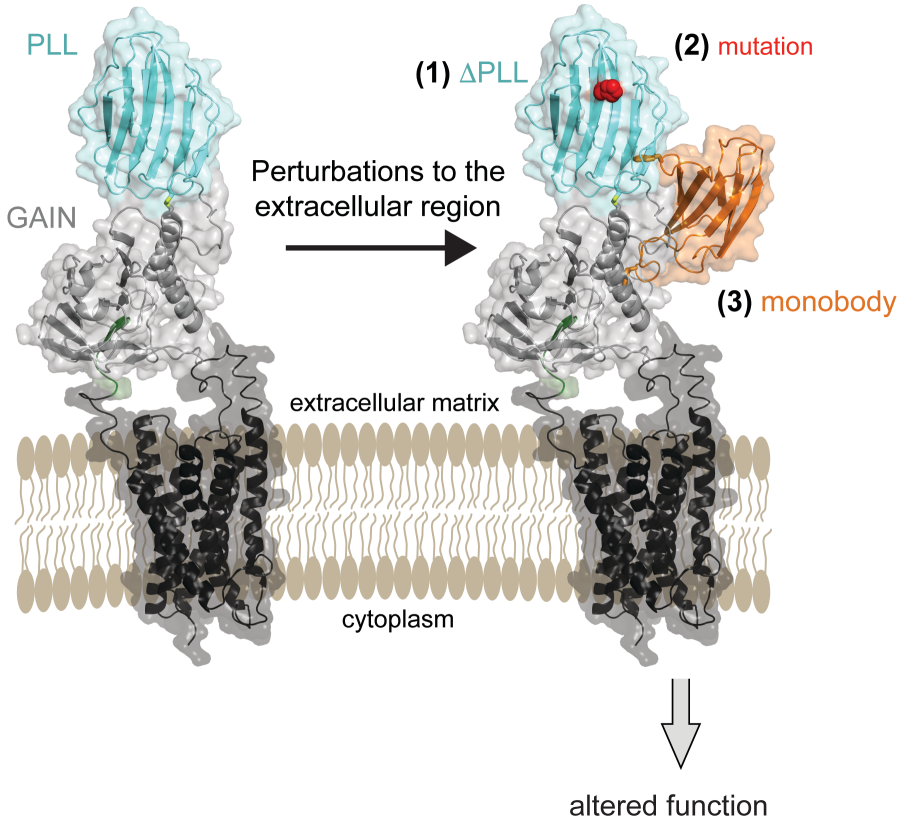


Figure 2.4 Mechanisms for perturbing GPR56 function by its extracellular region

Model of full-length GPR56 based on crystal structure of GAIN (gray) and PLL (cyan) domains (PDB: 5KVM) and 7TM model (black) from the glucagon receptor (PDB: 46LR). Pertubations to the extracellular region are noted as (1) deletion of PLL domain in S4 splice isoform which alters signaling, (2) H89A in PLL domain which alters *in vivo* function, and (3) treatment with α5 monobody which alters signaling. Figure is courtesy of Gabriel Salzman.

Adhesion GPCR N-termini as sensors of mechanical force

The mechanisms proposed for the activation of aGPCRs provided immense insight. However, critical questions still remain unanswered. Both the *Stachel*-dependent mechanism and ECR-mediated mechanism for aGPCR activation suggest that ligand binding to the ECR triggers 7TM signaling. In the *Stachel*-mediated mechanism, the *Stachel* peptide, which is deeply buried within the GAIN domain somehow escapes and confers a signal to the 7TM domain. In the ECR-mediated mechanism, ligand binding to the ECR either changes the conformation or the orientation of the ECR with respect to the 7TM domain to modulate receptor signaling. A critical question then is how is the energy derived in order to separate the *Stachel* peptide from the rest of the GAIN domain or to move the ECR? Recent studies suggest that mechanical force may be at play, which could apply enough force on the ligand in a certain direction that can then be transduced onto the ECR of the receptor. Such a mechanical force might be high enough to lead to shedding of the ECR and release of the *Stachel* peptide from the GAIN domain, enabling the *Stachel* peptide to activate the receptor in an irreversible manner. Or it may be a smaller force that is not strong enough to expose the *Stachel* but enough to pull or push the ECR to induce conformational/orientational changes of the ECR and modulate receptor signaling.

Recent work from several groups uncovered that mechanical cues trigger the activity of aGPCRs under physiological conditions, adding mechanosensation to the sensory repertoire of the GPCR superfamily. For example, the ADGRE5/CD97 NTF is released from the CTF after engagement with the ligand CD55, but only under mechanical shaking conditions that resembles the shear stress associated with

circulating blood (Karpus et al., 2013). Similarly, laminin-211, a ligand of ADGRG6/GPR126 stimulates the receptor only under shaking conditions and inhibited receptor activity under static conditions (Petersen et al., 2015). It is possible that the mechanical forces helped laminin-211 disengage the NTF from its CTF, whereas without shaking, the ligand binding likely stabilized the inhibitory NTF-CTF interaction. Furthermore, in ADGRG1/GPR56 mutant mice, muscle hypertrophy that is induced by resistance exercise is attenuated highlighting a potential role for GPR56 in sensing muscle fiber size, which could be mediated by mechanical sensitivity to stretch (White et al., 2014). A key *in vivo* study on aGPCR-mediated mechanosensation demonstrated that *D. melanogaster* larvae lacking the ADGL1/Lphn1 ortholog, CIRL, exhibited decreased sensitivity to mechanical stimuli (Scholz et al., 2015). These examples support the idea that, for at least some ligand-receptor pairs, mechanical force is a key determinant of the signaling output that results from the interaction.

Interaction of endogenous ligands with ECR of adhesion GPCRs: Latrophilins as a model

Lphn / FLRT interaction

Lphns mediate excitatory synapse formation by interacting with FLRT proteins, essential cell-adhesion molecules in cortical development and synapse formation (O'Sullivan et al., 2012). FLRTs additionally interact with Unc5/Netrin receptors to control the migration of neurons in the developing cortex (O'Sullivan et al., 2012; O'Sullivan et al., 2014). The crystal structure reveals that the N-terminal Ig domain of

Unc5 contacts the LRR domain of FLRT3 to form a high-affinity complex (Seiradake et al., 2014). The first structure of an aGPCR in complex with its ligand was determined for Lphn3 and FLRT3. (Lu et al., 2015b; Ranaivoson et al., 2015). The complex structure of the Lphn3 olfactomedin domain bound to the FLRT3 LRR domain revealed a large Lphn3-binding interface on the concave surface of FLRT3. Multiple cancer and ADHD mutations are localized on the Lphn3-binding surface of FLRT3 and on the FLRT3-binding surface of Lphn3 that specifically disrupt the interaction. The Lphn3/FLRT3 structure also showed that Lphn3 binds to FLRT3 at a site distinct from Unc5. Structural modeling in combination with biochemical studies demonstrated that Unc5 and Lphn3 can simultaneously bind to FLRT3, forming a trimeric complex, and that FLRT3 can form transsynaptic complexes with both Lphn3 and Unc5 (Lu et al., 2015b) (Figure 2.5). Later, the complex structure of the LPHN/FLRT/Unc5 corroborated these findings (Jackson et al., 2016). Structural studies based on crystal packing also proposed the formation of supercomplexes that may induce receptor clustering, however, the validity of such complexes at the synaptic junctions remains elusive (Jackson et al., 2016).

These results suggest that one distinct function of the N-terminal domains of Lphn is to tether the ECR to specific extracellular ligands through high-affinity protein-protein interactions. The Lphn/FLRT pair can interact in *trans* and cause cell-aggregation. In addition, a second function of the N-terminal domains of Lphn may be to mediate intercellular signaling. For example, FLRT may act as a ligand to modulate Lphn signaling, likely by sensing mechanical forces that are applied on FLRT.

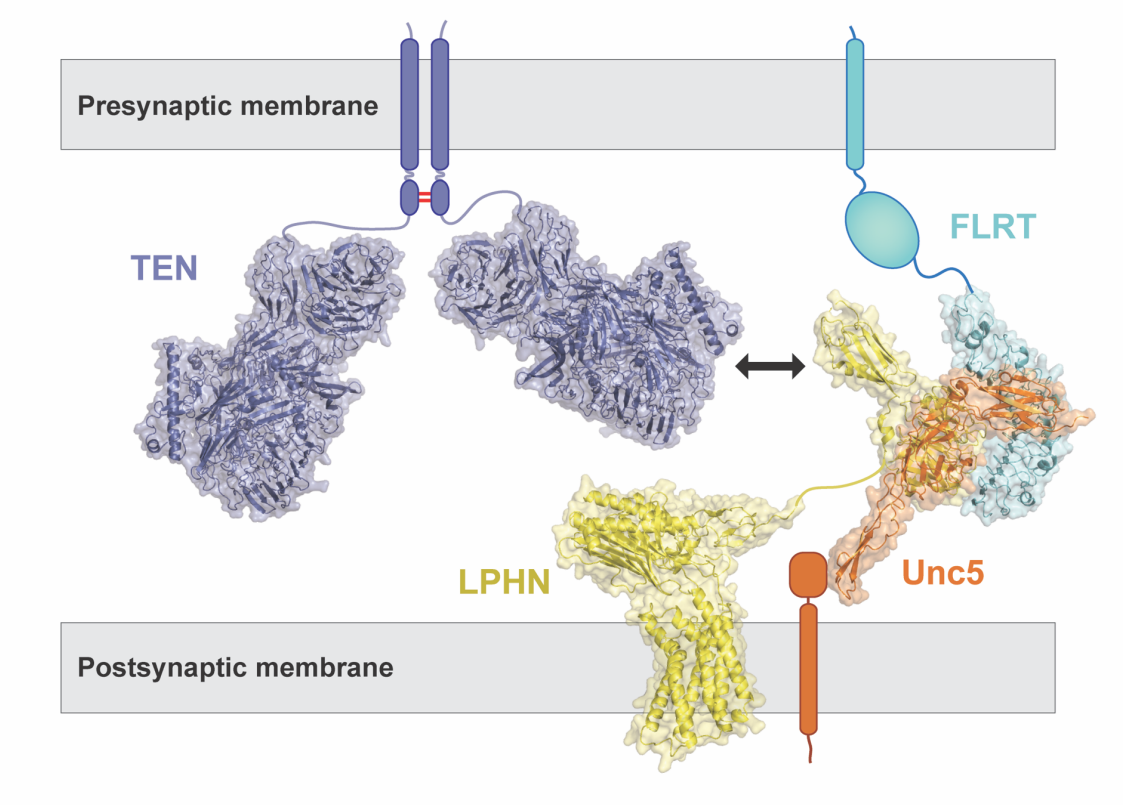


Figure 2.5 Model for TEN/Lphn/FLRT/Unc5 interaction

Representation of protein interactions at the synapse. Lphn3 (yellow), FLRT (cyan), Unc5 (orange) form a trimeric complex. Lphn1 is modeled using structure of GAIN (PDB: 4DLQ), structure of lectin-olfactomedin (PDB: 5AFB), and 7TM modeled from corticotropin-releasing factor receptor 1 (PDB: 4K5Y). Lphn1/FLRT/Unc5 complex is adapted from crystal structure of the complex (PDB: 5FTT). TEN (dark blue) is modeled from partial cryo-electron microscopy structure (PDB: 6CMX) and homology models.

Lphn / Teneurin interaction

Lphns form high affinity trans-cellular adhesion complexes with TENs to mediate functions during embryonic development and synaptogenesis (Boucard et al., 2014; Woelfle et al., 2015). The TEN/Lphn interaction is mediated by the lectin and olfactomedin domains of the Lphn ECR, with the lectin domain contributing most of the binding affinity (Boucard et al., 2014; O'Sullivan et al., 2014). TENs are type-II transmembrane proteins with large (>2000 amino acids) C-terminal ECRs that mediate

heterophilic and homophilic trans-cellular interactions that are key for various TEN functions (Mosca, 2015). With four members in humans, TENs are evolutionarily conserved cell-adhesion molecules that mediate intercellular communication and play central roles in embryogenesis, tissue polarity, heart development, axon guidance, and synapse formation (Lossie et al., 2005; Mosca et al., 2012; Nakamura et al., 2013; Tucker and Chiquet-Ehrismann, 2006; Young and Leamey, 2009).

Lphns were reported to mediate embryogenesis in *C. elegans* by modulating intracellular cAMP levels (Winkler and Promel, 2016), however no ligand was reported to induce a change of cAMP levels in a Lphn-dependent manner. Recently, a robust signaling assay was established for Lphns that detects decrease in cAMP levels in mammalian cell culture. This assay was used as a readout to monitor receptor activity in a setup that mimics trans-cellular signaling (Li et al., 2018; Nazarko et al., 2018). Cells expressing Lphn signaled more when co-cultured with cells expressing TEN2, suggesting that TEN2 may induce trans-cellular signaling in a Lphn-dependent manner by modulating cAMP levels.

TEN/Lphn interaction mediates key biological functions, however, a molecular understanding of this interaction remained unclear partly because most of the TEN sequences did not exhibit readily identifiable domains by sequence analysis. Recent structures of TENs revealed a striking similarity to bacterial Tc-toxins (Jackson et al., 2018; Li et al., 2018) (Figure 2.5) and showed that the ECR has an unusual architecture whereby a large cylindrical β-barrel partially encapsulates a C-terminal toxin-like domain that emerges from the barrel and is tethered to its outer surface. An immunoglobulin (Ig)-like domain seals the bottom of the barrel while a β-propeller is attached in a

perpendicular orientation. Mammalian cell-based binding experiments demonstrated that the tethered toxin-like domain of TENs mediates interactions with Lphn (Li et al., 2018; Silva et al., 2011). However, molecular details of the TEN/Lphn interaction still remain unknown. Surprisingly, it was shown that an alternatively spliced region within the β-propeller acts as a switch to regulate trans-cellular adhesion of TEN2 to Lphns (Li et al., 2018). Consequently, one splice variant activates trans-cellular signaling in a Lphn-dependent manner, whereas the other induces inhibitory postsynaptic differentiation in a Lphn-independent manner. These results highlight the unique structural organization of TENs giving rise to multifarious aGPCR-dependent and -independent functions.

Alternative splicing regulates adhesion GPCR function

Alternative splicing in other GPCR families is relatively rare. Almost all human aGPCRs are alternatively spliced (Bjarnadottir et al., 2007). Alternative splicing occurs mainly in the ECRs but also in the 7TM and intracellular regions. As a result of the abovementioned findings, the role of alternative splicing on the functions of aGPCRs is becoming more intriguing. Alternative splicing of the entire PLL domain in GPR56 affects the basal activity of the receptor and also deletes critical ligand binding sites from the receptor (Salzman et al., 2016). Lphn can interact with its ligand TEN only if TEN lacks a seven-amino acid splice insert (Li et al., 2018). In addition, the strength of the TEN/Lphn interaction is regulated by a four-amino acid splice site between the lectin and olfactomedin domains of Lphn (Boucard et al., 2014). ADGRG6/GPR126 is another aGPCR that is alternatively spliced, and the ECRs of GPR126 splice isoforms likely

adopt different conformations and lead to altered basal activity. In ADGRG5/GPR114, deletion of a single residue in the loop that connects the GAIN domain and the TM domain changes the basal activity of the receptor (Wilde et al., 2016). These observations suggest that aGPCRs have multiple layers of regulation and alternative splicing is likely another mechanism for fine-tuning aGPCR function.

Adhesion GPCRs are putative hormone receptors

The Hormone Receptor (HormR) domain is the second most frequently observed domain in aGPCRs (found in 13 out of 33 human aGPCRs, specifically in ADGRL1-3/Lphn1-3, ADGRA2/GPR124, ADGRA3/GPR125, ADGRC1-3/CELSR1-3, ADGRF3/GPR113, ADGRB1-3/BAI1-3, and ADGRG6/GPR126). Crystal structures show that the HormR domains of Lphn1 and BAI3 are ~70-residue domains formed by two anti-parallel β-sheets with conserved disulfide bonds and tryptophan residues, and yielded a low RMSD (1.1 Å) in spite of the low sequence identity (24%) (Arac et al., 2012). Since no hormone ligand has yet been found for aGPCRs, it was believed that the HormR domain in aGPCRs may not be a true hormone-binding region. However, a DALI search revealed that the HormR domain of the secretin-family corticotrophin releasing factor receptor (CRFR) (PDB ID: 3EHU) is strikingly similar to the HormR domains of Lphn1 and BAI3 yielding RMSDs of 0.7 Å and 1.1 Å, respectively, in spite of the low sequence identity (29%) (Grace et al., 2007). The unusually high structural similarity of the HormR domains of Lphn1 and BAI3 to that of the CRFR raised the possibility that aGPCRs can be bona fide hormone receptors.

HormR domain-containing aGPCRs are homologous to secretin GPCRs, which all include a HormR domain (Lagerstrom and Schioth, 2008). Superposition of the Lphn1 structure with the structure of the CRFR HormR domain bound to corticotrophin releasing factor, a 41-amino acid peptide hormone, showed that a similar hormone could not bind to Lphn1 HormR because the GAIN domain is blocking the homologous hormone binding site on the HormR domain of Lphn1 (Arac et al., 2012). Clearly, hormone binding would require a conformational change in Lphn1 to expose the putative hormone binding site (such as the reduction of one of the conserved disulfide bonds to switch to an open conformation). In secretin GPCRs, the HormR domain precedes the 7TM and is juxtaposed to the membrane. According to the two-domain model of secretin GPCR activation, the interaction of the extracellular HormR domain with the hormone promotes the interaction of the hormone with the 7TM, leading to the activation of the receptor (Hoare, 2005). However, the mechanism of aGPCR activation upon hormone binding may be different because the GAIN domain lies between the HormR domain and the 7TM.

Adhesion GPCRs as drug target: potential for 7TM- and ECR-mediated modulation

Emerging data show the involvement of aGPCRs in numerous physiological functions and human diseases. aGPCRs, thus, constitute a large group of molecules that are potential drug targets. As other GPCR family members are successfully targeted by a variety of molecules binding to either orthosteric or allosteric binding sites in the 7TM domain, drugging aGPCRs is possible. Rational design of aGPCR ligands

that interact with a 7TM binding cavity relies on good receptor models. The determination of the first structure for the 7TM domain of an aGPCR will open the way to the development of aGPCR 7TM-targeted drugs.

A major challenge in GPCR-targeted drug design is the high conservation of the 7TM, which requires high specificity of drugs in order to minimize undesirable side-effects (Schlyer and Horuk, 2006). As aGPCR ECRs are much more diverse than their 7TMs, the pursuit of aGPCR ECR-targeted (*i.e.*, allosteric) synthetic ligands, such as monobodies or antibodies, will likely result in highly specific reagents. The finding that monobodies alter basal activity by binding to the ECR of aGPCRs is an encouraging proof of concept for developing highly selective modulators of aGPCRs. ECR-mediated receptor regulation is likely to be moderate. Thus, the ECR-mediated regulation should be suited for fine-tuning the signaling near the basal levels. Furthermore, as the therapeutic potential of allosteric GPCR modulators that exhibit moderate effects has been demonstrated (Christopoulos, 2014; Wootten et al., 2013; Wootten et al., 2016), the aGPCR ECRs are potential drug targets. The powerful combination of ECR-targeted synthetic ligands will be invaluable in future mechanistic and pharmacological studies of aGPCRs.

Chapter 3: Structural Basis for Gpr126 Function[1]

Introduction

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 (Chun et al., 2012; de Graaf et al., 2017; Hu and Luo, 2013; Kovacs et al., 2015; Luo et al., 2007; Rana et al., 2013; Salzman et al., 2016; Shimaoka and Springer, 2003; Sun et al., 2016). 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 conformation (Li et al., 2005), and the drug etrolizumab blocks ligand binding to the ECRs of integrins in order to treat inflammatory bowel diseases (Ley et al., 2016). Remarkably, earlier this year, the migraine preventive

[1] The text from this chapter was taken verbatim from: Leon K, Cunningham RL, Riback JA, Feldman E, Li J, Sosnick TR, Zhao M, Monk KR, Araç D (2019). Structural basis for adhesion G protein-coupled receptor GPR126 function. 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.

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 (Shi et al., 2016; Traynor, 2018). Despite these and other breakthroughs, there are many essential receptors in the human genome that are not currently drugged, including the 33 adhesion GPCRs (aGPCRs), a diverse and understudied family of GPCRs with critical roles in synapse formation, angiogenesis, neutrophil activation, embryogenesis, and more (Hamann et al., 2015; Langenhan et al., 2016; Yona et al., 2008a).

Like all GPCRs, aGPCRs have canonical signaling seven-transmembrane (7TM) domains (Manglik and Kruse, 2017; Paavola and Hall, 2012). 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 interactions (Langenhan et al., 2013). In addition, during biosynthesis, aGPCRs are uniquely autoproteolysed within a conserved GPCR Autoproteolysis INducing (GAIN) domain of the ECR that is juxtaposed to the 7TM (Arac et al., 2012), resulting in a fractured receptor that nevertheless remains tightly associated at the cell surface (Chang et al., 2003; Paavola et al., 2011).

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 (Kishore and Hall, 2017; Kishore et al., 2016; Paavola et al., 2011; Salzman et al., 2016; Salzman et al., 2017). A current model for aGPCR regulation suggests that transient interactions between the ECR and 7TM directly regulate receptor signaling (Kishore and

Hall, 2017; Kishore et al., 2016; Paavola et al., 2011; Salzman et al., 2016; Salzman et al., 2017). There are also numerous reports that aGPCRs use their ECRs to mediate functions in a 7TM-independent manner (Koh et al., 2004; Muller et al., 2015; Patra et al., 2013; Petersen et al., 2015; Promel et al., 2012a). Another non-mutually exclusive model for aGPCR activation posits that ligand binding to the ECR can exert force and cause dissociation at the autoproteolysis site, revealing a tethered peptide agonist, which then activates the receptor (Demberg et al., 2015; Liebscher et al., 2014; Stoveken et al., 2015; Wilde et al., 2016). 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.

ADGRG6/Gpr126 is one of the better studied aGPCRs and is essential for Schwann cell (SC) myelination and other functions (Geng et al., 2013; Monk et al., 2009; Waller-Evans et al., 2010). In vertebrate peripheral nervous system (PNS) development, the myelin sheath surrounding axons is formed by SCs and functions to facilitate rapid propagation of action potentials (Pereira et al., 2012). Disruption of myelination is associated with disorders such as Charcot-Marie-Tooth disease, which is characterized by muscle weakness (d'Ydewalle et al., 2012; Nave et al., 2007). 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 (Mogha et al., 2013; Monk et al., 2011) and that Gpr126 also plays a role in

myelin maintenance through communication with the cellular prion protein (Kuffer et al., 2016). In humans, GPR126 mutations are linked to several diseases (Garinet et al., 2018; Nik-Zainal et al., 2016; Piraino and Furney, 2017; Qin et al., 2017; Terzikhan et al., 2018), including arthrogryposis multiplex congenita, a disorder characterized by multiple joint contractures (Ravenscroft et al., 2015). Furthermore, Gpr126 is required for inner ear development (Geng et al., 2013) and heart development (Waller-Evans et al., 2010), and it has been shown that the latter function is ECR-dependent and does not require the 7TM (Patra et al., 2013). 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 (Figure 3.1). 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 (Arac et al., 2012; Moriguchi et al., 2004; Stehlik et al., 2004). However, a 150 aa region between PTX and HormR, could not be identified. Furin, a Golgi-localized protease, is reported to cleave Gpr126 in this region (Moriguchi et al., 2004), although any effect on protein architecture is unclear because of the unspecified structure. In addition, alternative splicing occurs in *GPR126*, resulting in Gpr126 isoforms that vary in their ECRs (Bjarnadottir et al., 2007; Moriguchi et al., 2004). Alternative splicing of exon 6 was observed in human and zebrafish (Moriguchi et al., 2004; Patra et al., 2013), 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, further

emphasizing the importance of determining the ECR structure, conformation, and other possible unexplored features in understanding Gpr126 function.

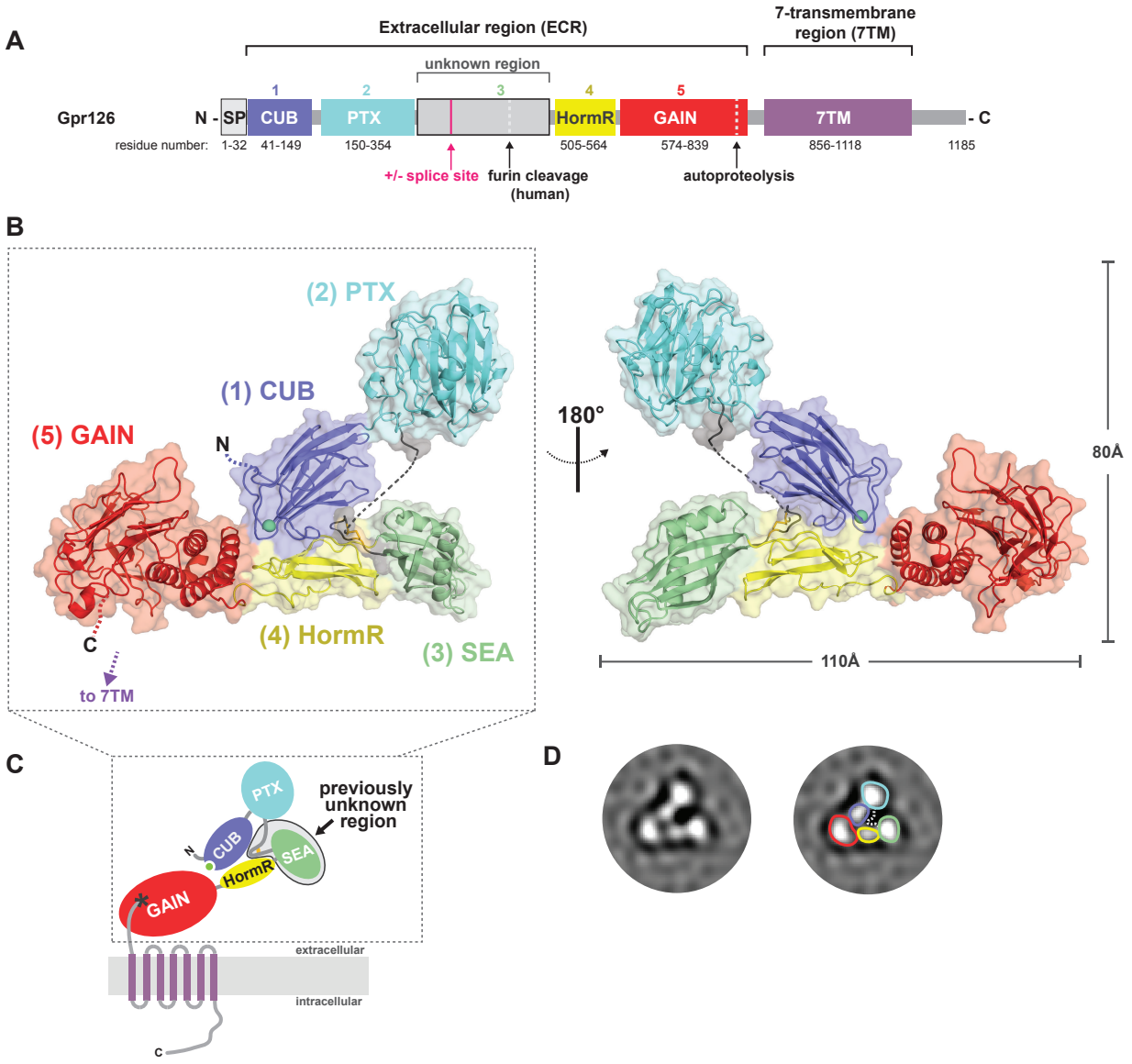


Figure 3.1 Crystal structure of the full extracellular region of Gpr126

(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), grey (unknown region), yellow (HormR), red (GAIN) and purple (7TM). Domain boundaries are indicated below. SP indicates signal peptide. (continued on next page)

(continued from previous page) (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. (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 β-strand of the GAIN domain is colored grey. (D) Representative negative stain EM 2D class average of Gpr126 (-ss) ECR. Domains are assigned and colored according to color scheme noted above. The dashed line represents the linker region.

In this study, we determined the high-resolution crystal structure of the full-length ECR of Gpr126, which reveals five domains, including a newly identified Sperm protein, Enterokinase and Agrin (SEA) domain which is the site of furin-mediated cleavage. 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

The structure of the full-length ECR of Gpr126 (-ss) reveals five domains in a closed conformation

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 has high sequence identity (47%) to its human homolog but has a fewer number of N-linked glycosylation sites (16 predicted in zebrafish, 26 in human) and no furin cleavage site (Figure 3.2A), and thus yields a more homogeneous sample (Figures 3.2B and 3.2C). Crystals of both native and selenomethionine (SeMet)-labeled zebrafish Gpr126 ECR (-ss) were obtained and diffracted to 2.4 Å (Figure 3.2D), and the structure was determined by SeMet single-wavelength anomalous diffraction (SAD) phasing (Table 3.1).

Table 3.1 Data Collection and Refinement Statistics

Protein	Gpr126 ECR Native Dataset	Gpr126 ECR SeMet SAD
Data Collection		
Integration Package	HKL2000	HKL2000
Wavelength	1.033 Å	0.979 Å
Space group	C2	C2
Cell dimensions		
a, *b*, *c*	144.96 Å, 59.45 Å, 168.39 Å	145.20 Å, 59.33 Å, 168.72 Å
α, β, γ	90°, 107.82°, 90°	90°, 108.16°, 90°
Resolution	50–2.40 Å (2.46–2.40 Å)	50–2.34 Å (2.38–2.34 Å)
R_{sym} or R_{merge}	0.064 (0.664)	0.086 (0.492)
$CC_{1/2}$, highest res. bin	0.557	0.376
$I/\sigma I$	15.0 (0.9)	13.0 (1.4)
Completeness	77.9% (25.4%)	71.6% (9.7%)
Redundancy	2.9 (1.7)	3.8 (1.3)
Number of measured reflections	125687	157602
Number of unique reflections	43044	41385
Refinement Statistics		
R_{work}/R_{free}	0.216/0.271	
Number of atoms		
Protein	5844	
Water	146	
Other	169	
Average B-factors		
Protein	28.8 $Å^2$	
Water	26.1 $Å^2$	
Other	57.5 $Å^2$	
Rmsds		
Bond lengths	0.009 Å	
Bond angles	1.118°	
Ramachandran plot statistics		
Most favorable	90.98%	
Allowed	8.22%	
Disallowed	0.8%	

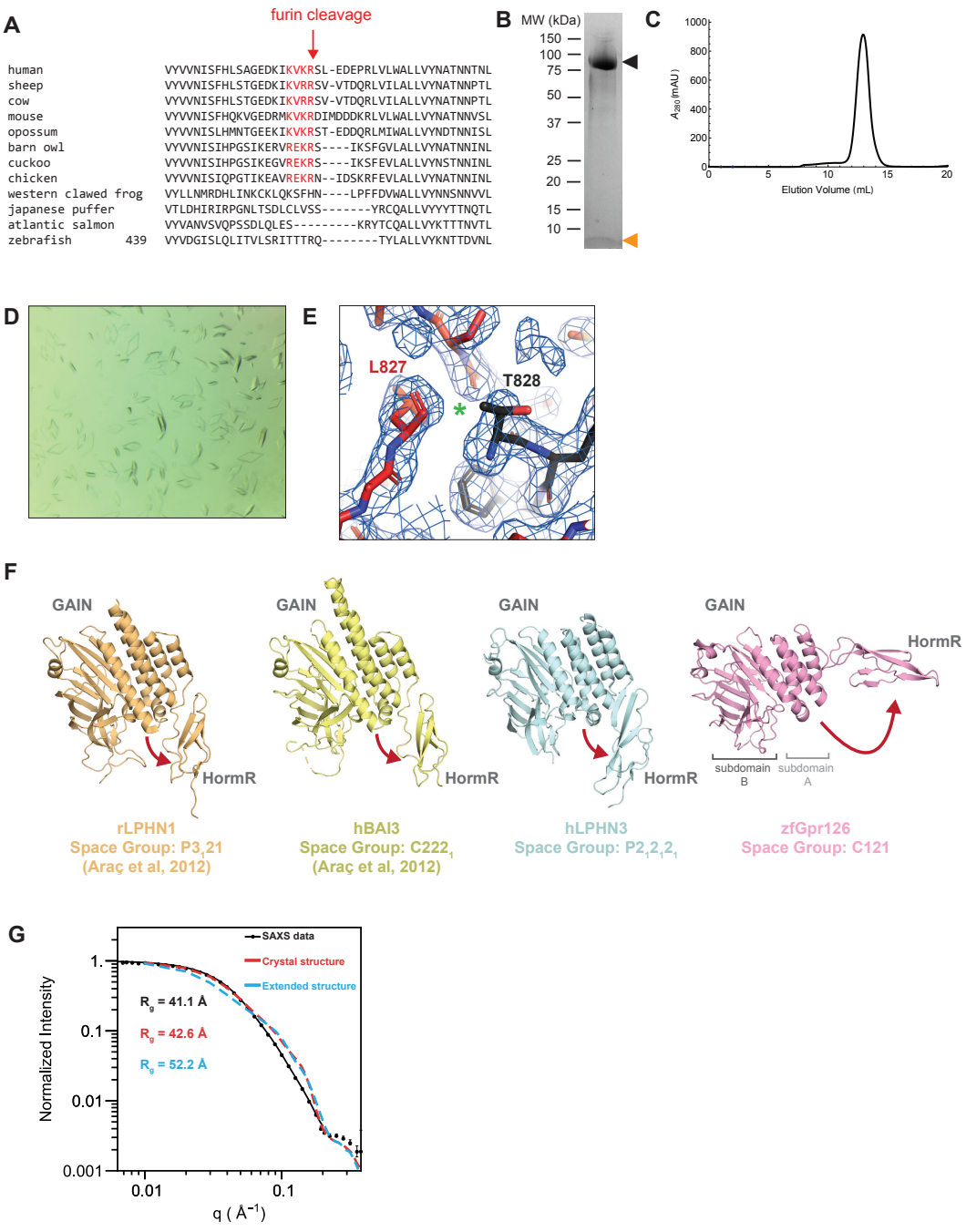


Figure 3.2 Purification and crystallization of Gpr126 ECR

(A) Sequence alignment of Gpr126 furin-cleavage site from various species. Consensus cleavage site residues are colored red. (B) SDS-PAGE of purified Gpr126 ECR showing bands for N-terminal fragment (black arrowhead) and tethered peptide (orange arrowhead). (C) Size exclusion column profile of purified Gpr126 ECR. (D) Crystals of Gpr126 ECR. (continued on next page)

(continued from previous page) (E) Electron density for Gpr126 ECR showing lack of electron density at autoproteolysis site (asterisk) generating an N-terminal fragment (red) and C-terminal fragment (black). (F) HormR+GAIN domain structures and space groups of rLPHN1 (orange), hBAI3 (yellow), hLPHN3 (light blue), and zfGpr126 (pink). Red arrows indicate angles between GAIN and HormR domains, highlighting difference seen in zfGpr126 compared to all other structures. (G) Experimental SAXS curve for Gpr126 ECR (black) fit to a smooth regularized scattering curve (i.e. the Fourier transform of p(r), solid black line, see methods), simulated SAXS curve for Gpr126 ECR based on crystal structure (red), and simulated SAXS curve for extended model of Gpr126 ECR (blue).

The structure, with overall dimensions of 110Å x 80Å x 35Å, revealed the presence of five domains (Figures 3.1B and 3.1C), 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 server (Holm and Laakso, 2016). The Gpr126 SEA domain adopts a ferredoxin-like alpha/beta sandwich fold, and superimposition over known SEA domain structures from Mucin-1 and Notch-2 (RMSD: 2.8Å and 4.8Å, respectively) (Gordon et al., 2007; Macao et al., 2006) showed that the Gpr126 SEA domain indeed adopts the same fold. 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. Finally, the SEA domain is followed by the HormR and GAIN domains, the latter of which is autoproteolyzed as expected (Figure 3.2E). The HormR and GAIN domain structures are similar to previously-solved HormR+GAIN domain structures from other aGPCRs (Arac et al., 2012; Salzman et al., 2016), with the exception of the

relative orientation between HormR and GAIN. There is a 90º rotation of the HormR domain with respect to the GAIN domain (Figure 3.2F) in Gpr126 compared to previously-solved HormR+GAIN structures from rLphn1, hBAI3, and hLphn3 (Arac et al., 2012).

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 (Figure 3.1B). 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 confirmation is observed for Gpr126 in solution. Negative-stain 2D class averages of Gpr126 ECR showed a V-shaped protein architecture (Figure 3.1D). 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 $\pm$ 0.1 Å) is consistent with the calculated R_g of the crystal structure (42.6 Å) and inconsistent with that of an extended model of GPR126 ECR (R_g = 52.2Å) (Figure 3.2G). Taken together, these results: show that the closed, rather than extended, conformation of Gpr126 ECR exists 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.

A calcium-binding site and a disulfide-stabilized linker mediate the closed conformation of Gpr126 ECR

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 (Figure 3.3A).

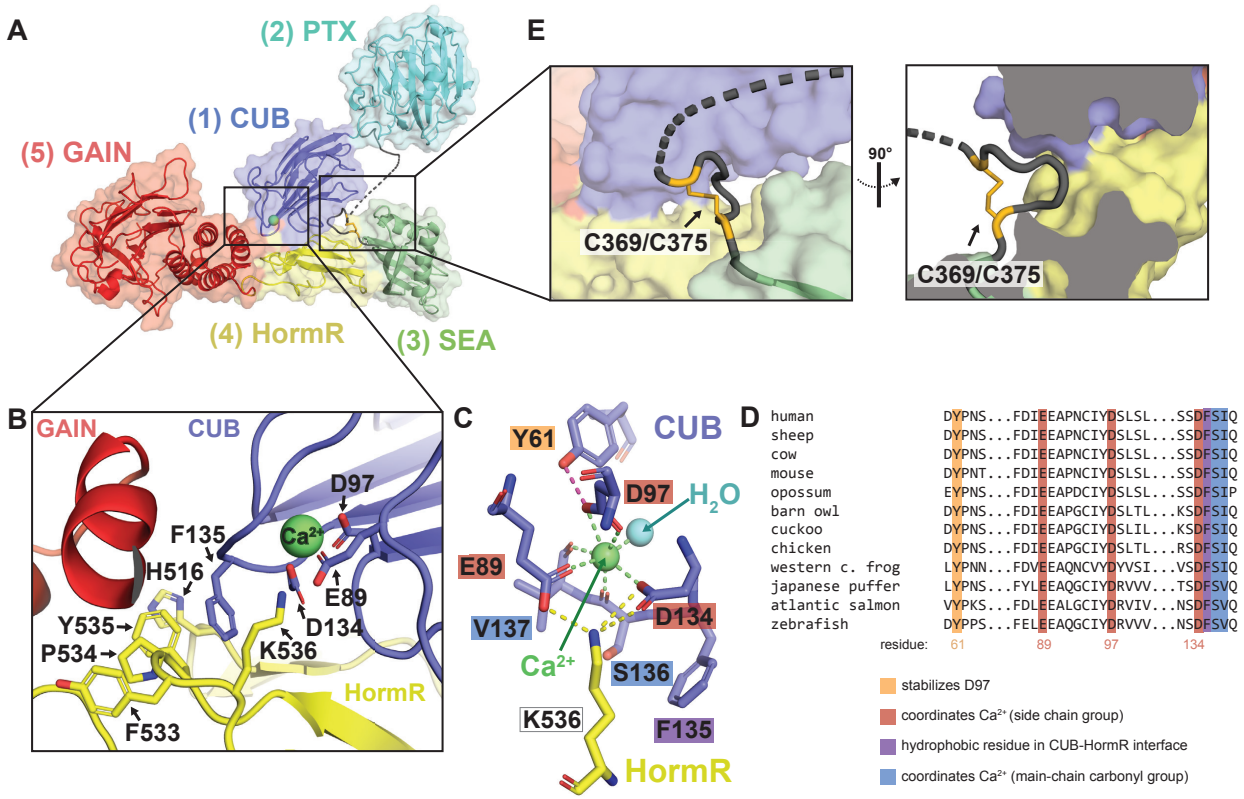
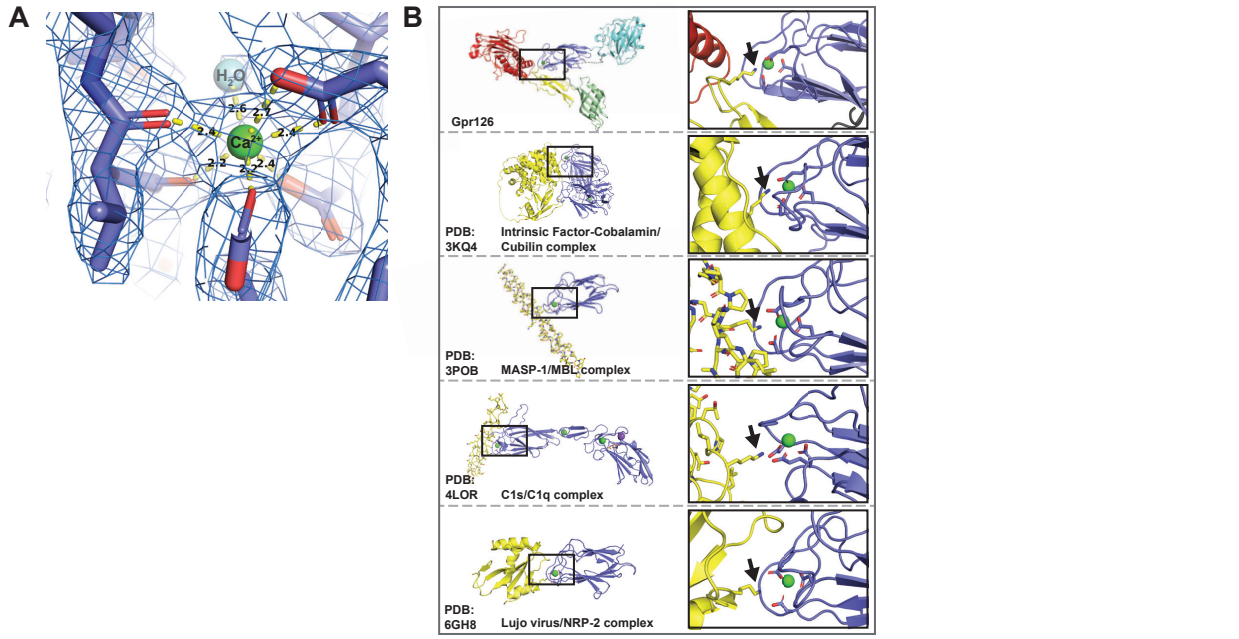


Figure 3.3 The closed conformation of Gpr126 is mediated by CUB-HormR-linker interactions

(A) Structure of the full ECR of (-ss) Gpr126. (B) Close-up view of the CUB-HormR interface. Resides at the interface are shown as sticks. The calcium ion is shown as a bright green sphere. (continued on next page)

(continued from previous page) (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.

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 (Figure 3.3B).



C

```
NRP2_CUB1          --CGGRLNSKDAGYITSPGYPQDYPSHQNCEWIVYAPEPNQKIVLNFNPHFEIE-----KHDCKYDFIEIRDGDSESADLLGKHCGNIAP----PTIISSGSMLYIKFTSDYARQG--AGFSLRYEIF
MASP1_CUB2         --CSGNLFTQRTGTITSPDYPNPYPKSSECSYTIDLEE-GFMVTLQFEDIFDIED--HPEVPCPYDYIKIKAGS----KVWGPFCGEKS----PEPISTQSHSIQILFRSDNSGEN--RGWRLSYRAA
C1s_CUB1           -------EPTMYGEILSPNYPQAYPSEVEKSWDIEVPE-GYGIHLYFT-HLDIEL----SENCAYDSVQIISGDTEEGRLCGQRSSNNPHSPIVEEFQVPYNKLQVIFKSDFSNEERFTGFAAYYVAT
MASP2_CUB1         TPLGPKWPEPVFGRLASPGFPGEYANDQERRWTLTAPP-GYRLRLYFT-HFDLEL----SHLCEYDFVKLSSGAKVLATLCGQESTDTERAPGKDTFYSLGSSLDITFRSDYSNEKPFTGFEAFYAAE
hGPR126_CUB        --CRVVLS-NPSGTFTSPCYPNDYPNSQACMWTLRAPT-GYIIQITFN-DFDIEE----APNCIYDSLSLDNGE-----SQTKFCGATAK---GLSFNSSANEMHVSFSSDFSIQK--KGFNASYIRV
zfGpr126_CUB    41 --CNVVLT-DSQGSFTSPCYPNDYPPSQSCNWTIQAPA-GFIVQITFL-DFELEE----AQGCIYDRVVVKTGT-----SDAKFCGLTAN---GLTLNSTGNVMEVFFNSDFSVQK--KGFHISYKQV
```

D

```
human                   ASCADLGTLCQ
sheep                   ASCADLGTLCQ
cow                     ASCADLGTLCQ
mouse                   TNCSELGTLCQ
opossum                 TSCADLGTLCQ
barn owl                TSCANLGSLCQ
cuckoo                  TSCANLGSLCQ
chicken                 TSCANLGSLCQ
western clawed frog     TTCKNLGGVCQ
japanese puffer         TSCDPPGVGCP
atlantic salmon         TSCSSPGLDCP
zebrafish         367   AGCAS-GLGCP
                          └ S-S ┘
                        disulfide bond
```

Figure 3.4 Analysis of the Gpr126 CUB domain

(A) Electron density for Gpr126 ECR showing density for calcium ion (green sphere). Coordination interactions are shown as yellow dashed lines and distances (Å) are noted. (B) (Left panels) Structures of proteins that use calcium-binding CUB domains (blue) to bind ligands (yellow). (Right panels) Close-up views of the interfaces with arrows pointing to the lysine residues on the ligands which interact with calcium-binding residues on the CUB domains. (C) Sequence alignment of CUB domains from various proteins, highlighting conserved calcium-coordinating residues (red). (D) Alignment of the disulfide-bond loop region from various species showing the cysteines (bright orange) are universally conserved.

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 (Figure 3.4A). 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 (Figure 3.3C). The geometry and distances between the density and the coordinating residues in Gpr126 are consistent with calcium coordination (Harding, 2000). Several CUB domains from extracellular proteins are reported to coordinate calcium, including Gpr126 (Gaboriaud et al., 2011), and some have been discovered to use this coordination to mediate ligand binding (Andersen et al., 2010; Cohen-Dvashi et al., 2018; Gaboriaud et al., 2011; Gingras et al., 2011; Venkatraman Girija et al., 2013) (Figure 3.4B). For example, the C1s protein uses its CUB calcium-binding site to bind to ligand C1q and initiate the classical pathway of complement activation (Venkatraman Girija et al., 2013), and the Lujo virus recognizes a calcium-binding site on the CUB domain of the neurophilin-2 receptor in order to gain cell entry (Cohen-Dvashi et al., 2018). The calcium-coordinating residues are all conserved in the Gpr126 CUB domain (among GPR126 in various species (Figure 3.3D) as well as among calcium-binding CUB domains in other proteins (Figure 3.4C)), 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 (Figure 3.3C), 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 (Figure 3.3E). 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 one of the 90 species analyzed in this study (Figure 3.4D and Table 3.2), 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.

Table 3.2 Information for various Gpr126 species

(continued on following pages)

class	common name	furin cleavage site?	evidence of splice site between PTX/SEA?	conserved C369/C375 cysteines?
mammalia	giant panda	yes	yes	yes
reptilia	american alligator	yes	yes	yes
aves	mallard	yes	yes	yes
reptilia	green anole [lizard]	yes	yes	yes
aves	chuck-wills-widow	yes	no	yes
aves	emperor penguin	yes	no	yes
actinopterygii	mexican tetra	no	yes	yes
mammalia	domestic yak	yes	yes	yes
mammalia	cow	yes	yes	yes
aves	rhinoceros hornbill	no	no	yes
mammalia	marmoset	yes	yes	yes
aves	anna's hummingbird	yes	no	yes
mammalia	wild bactrian camel	yes	yes	yes
mammalia	dog	yes	yes	yes
aves	killdeer	yes	no	yes
aves	macqueen’s bustard	yes	no	yes

Table 3.2, continued

mammalia	green monkey	yes	yes	yes
aves	speckled mouse-bird	yes	no	yes
aves	american crow	yes	yes	yes
mammalia	chinese hamster	yes	yes	yes
reptilia	australian saltwater crocodile	yes	yes	yes
aves	cuckoo	yes	no	yes
actinopterygii	zebrafish	no	yes	yes
mammalia	beluga whale	yes	yes	yes
mammalia	ord's kangaroo rat	yes	yes	yes
aves	little egret	n/a	no	yes
mammalia	horse	yes	yes	yes
mammalia	european hedgehog	yes	yes	yes
aves	collared flycatcher	yes	yes	yes
actinopterygii	mummichog	no	yes	yes
aves	chicken	yes	yes	yes
mammalia	gorilla	yes	yes	yes
aves	white-tailed eagle	yes	no	yes
mammalia	naked mole-rat	yes	yes	yes
mammalia	human	yes	yes	yes
actinopterygii	channel catfish	no	yes	yes
actinopterygii	large yellow croaker	no	yes	yes
actinopterygii	baramundi perch	no	yes	yes
actinopterygii	spotted gar	yes	yes	yes
mammalia	african bush elephant	yes	yes	yes
mammalia	crab-eating macaque	yes	yes	yes
mammalia	rhesus macaque	yes	yes	yes
aves	golden-collared manakin	no	no	yes
aves	turkey	yes	no	yes
aves	northen carmine bee-eater	n/a	no	yes
mammalia	golden hamster	yes	yes	yes
mammalia	opossum	yes	yes	yes
mammalia	mouse	yes	yes	yes
mammalia	ferret	yes	yes	yes
mammalia	vesper bat	yes	no	yes
mammalia	little brown bat	yes	no	yes
aves	crested ibis	no	no	yes
mammalia	northern white-cheeked gibbon	yes	yes	yes
actinopterygii	turquoise killifish	no	yes	no
actinopterygii	coho salmon	no	yes	yes
actinopterygii	rainbow trout	no	yes	yes
actinopterygii	chinook salmon	no	yes	yes
aves	hoatzin bird	yes	no	yes

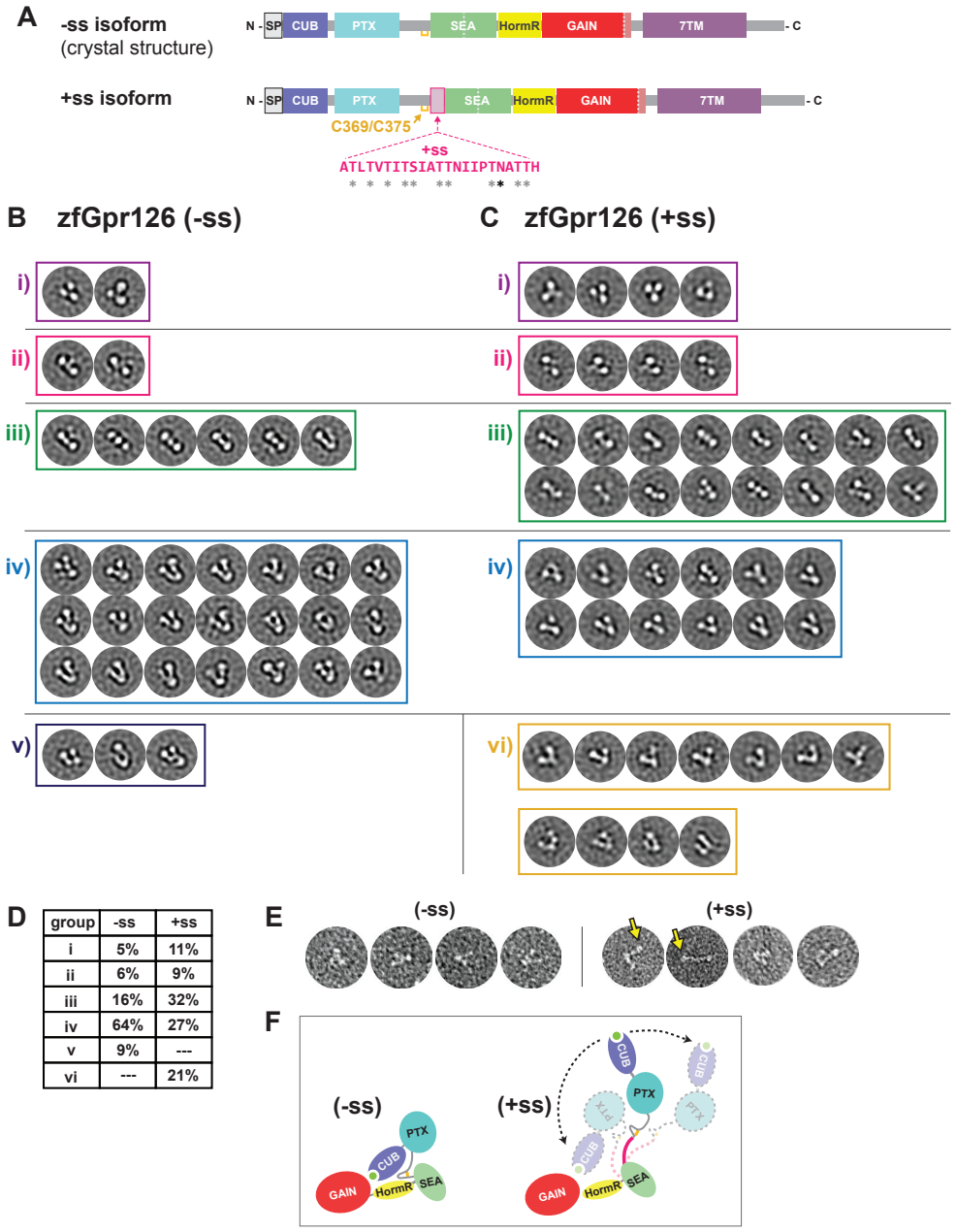
Table 3.2, continued

actinopterygii	nile tilapia	no	yes	yes
mammalia	platypus	yes	no	yes
mammalia	european rabbit	yes	yes	yes
mammalia	northern great galago	yes	yes	yes
mammalia	sheep	yes	yes	yes
mammalia	chimpanzee	yes	yes	yes
mammalia	leopard	yes	yes	yes
mammalia	olive baboon	yes	yes	yes
aves	dalmatian pelican	yes	no	yes
reptilia	chinese soft-shell turtle	no	yes	yes
aves	white-tailed tropicbird	yes	no	yes
aves	great cormorant	no	no	yes
aves	downy woodpecker	no	no	yes
aves	great crested grebe	yes	no	yes
actinopterygii	amazon molly	no	yes	yes
mammalia	sumatran orangutan	yes	yes	yes
aves	yellow-throated sandgrouse	yes	no	yes
mammalia	black flying fox	yes	no	yes
aves	adelie penguin	no	no	yes
mammalia	brown rat	yes	yes	yes
actinopterygii	atlantic salmon	no	yes	yes
mammalia	tasmanian devil	no	yes	yes
mammalia	pig	yes	yes	yes
aves	zebra finch	yes	no	yes
actinopterygii	japanese puffer	no	yes	yes
aves	red-crested turaco	no	no	yes
reptilia	three-toed box turtle	yes	yes	yes
actinopterygii	green pufferfish	no	yes	yes
aves	white-throated tinamou	no	no	yes
aves	barn owl	yes	no	yes
amphibia	western clawed frog	n/a	yes	yes
actinopterygii	southern platyfish	n/a	yes	yes

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 (Figure 3.5A). Because the disulfide-stabilized loop makes contacts that are important for the closed conformation of Gpr126 ECR (-ss) (Figure 3.3E), we hypothesized that the (+ss) isoform would disrupt the closed conformation and have a different, more open conformation.



group	-ss	+ss
i	5%	11%
ii	6%	9%
iii	16%	32%
iv	64%	27%
v	9%	---
vi	---	21%

Figure 3.5 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. Grey 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. (continued on next page)

(continued from previous page) (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 that are observed only in the +ss isoform. (v) represents unidentifiable miscellaneous views. (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.

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 (Figure 3.5B), were consistent with the closed conformation of the crystal structure (Figure 3.1B). However, the class averages for the (+ss) isoform showed a diverse population of ECR molecules, as they contain additional more open-like conformations (group vi, 21% of particles, Figure 3.5D) as well as closed conformations that were observed in the -ss isoform (Figures 3.5C and 3.5D). Furthermore, individual (+ss) particles showed the presence of open conformations (Figure 3.5E, yellow arrows), 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) (Figure 3.5F).

The negative stain EM data is consistent with SAXS experiments showing that the R_g of zebrafish Gpr126 ECR (+ss) is larger than that of (-ss) (Figures 3.6A and 3.6B), with a more dramatic change in R_g observed between the human GPR126 isoforms. 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 (Figures 3.6C and 3.6D).

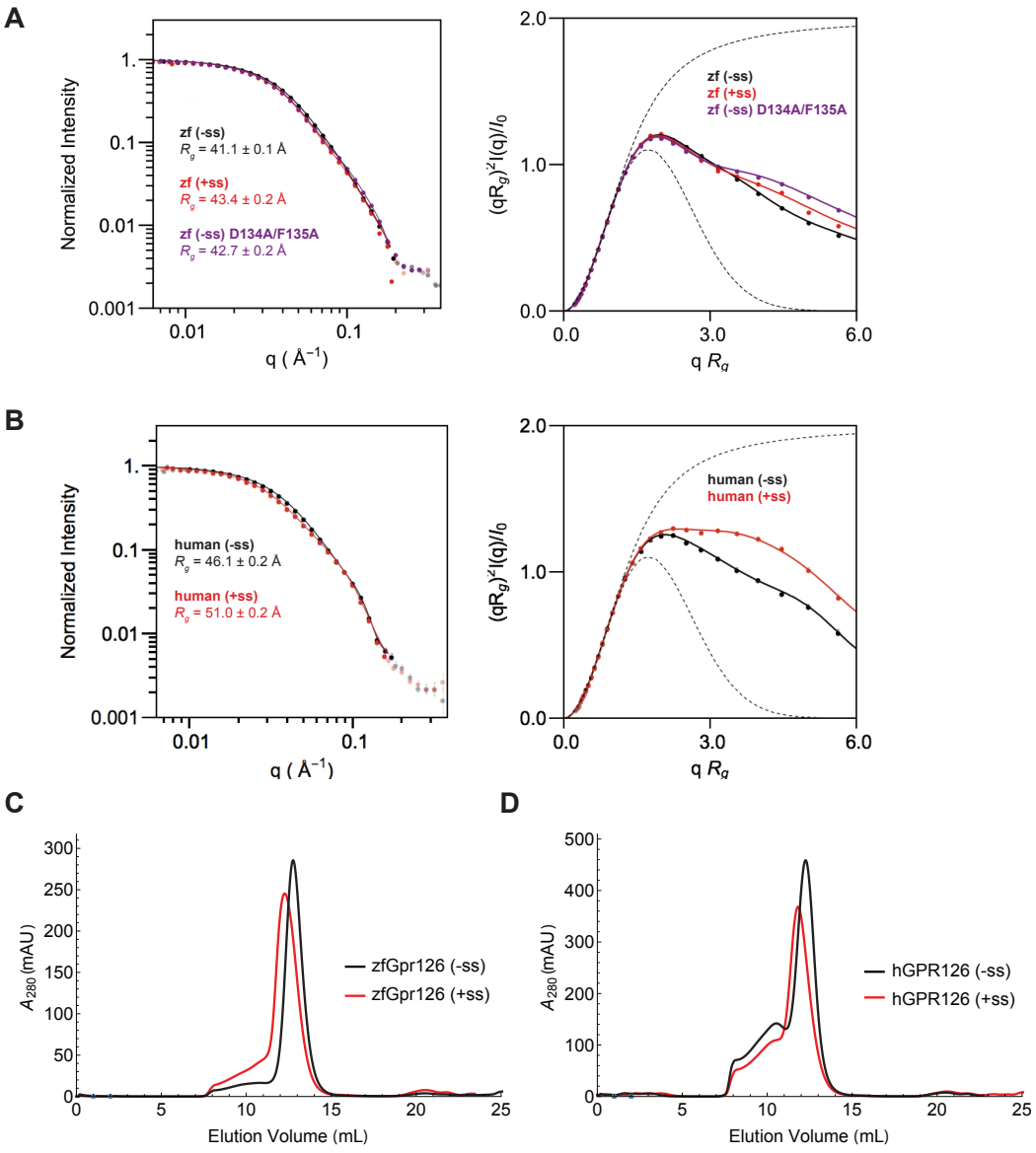


Figure 3.6 Differences in ECR conformation between Gpr126 splice isoforms

(A) (Left panel) SAXS curves and (right panel) normalized Kratky plots for zfGpr126 (-ss) (black), (+ss) (red), and (-ss) D134A/F135A (purple). (B) (Left panel) SAXS curves and (right panel) normalized Kratky plots for hGPR126 (-ss) (black) and (+ss) (red). Dashed lines show profiles expected for a random walk and compact (Guinier) particle located at the top and bottom, respectively of the plots. Solid lines show experimental fit with a smooth regularized scattering curve (i.e. the Fourier transform of p(r), solid black line, see methods). (C) Size exclusion column profiles for zfGpr126 (-ss) (black) and (+ss) (red). (D) Size exclusion column profiles for hGPR126 (-ss) (black) and (+ss) (red).

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. GPR126 has been shown previously to couple to and activate $G\alpha_s$, leading to production of cAMP (Mogha et al., 2013). Therefore, we used a cAMP signaling assay in which HEK293 cells were co-transfected with a full-length Gpr126 construct and a reporter luciferase that emits light upon binding to cAMP. Cell-surface expression level of Gpr126 constructs was quantified by flow cytometry analysis of cells stained by antibodies against N-terminal FLAG-tags (Figure 3.7A), and basal signaling results (Figure 3.7B) were normalized to expression level (Figure 3.7C).

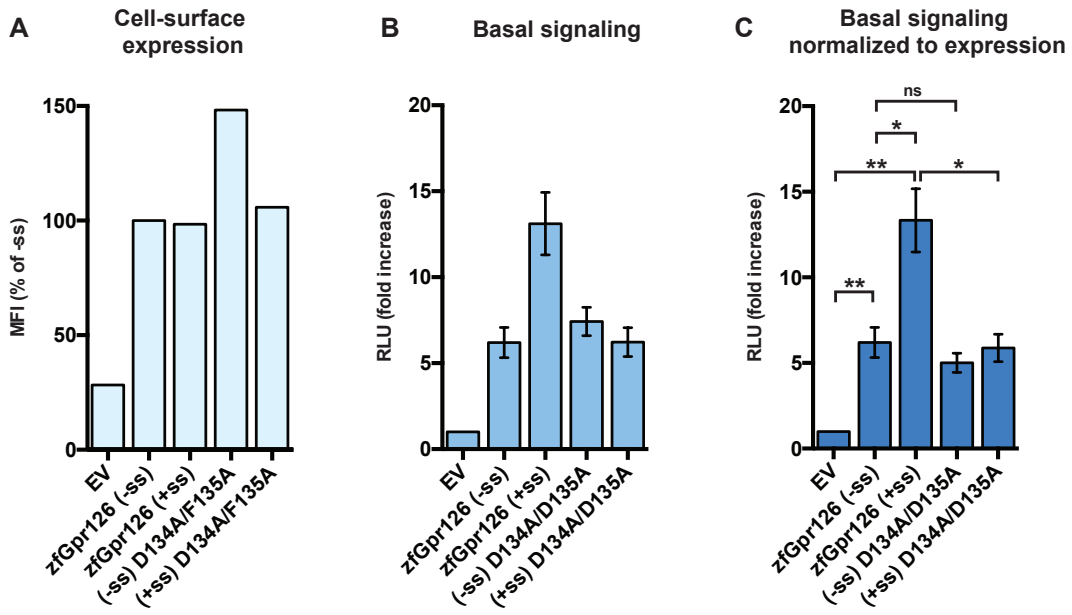


Figure 3.7 Alternative splice isoforms of Gpr126 modulate receptor signaling

(A) Cell-surface expression levels for empty vector (EV), zebrafish Gpr126 splice isoforms, and D134A/F135A mutants, 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. (B) Basal signaling measured by the cAMP signaling assay. Data are shown as fold increase over EV of RLU (relative luminescence units). (continued on next page)

(continued from previous page) (C) Basal cAMP signaling normalized to cell-surface expression. ns, $P > 0.05$; *. $P \leq 0.05$; **. $P \leq 0.01$ by unpaired two-tailed t test. Data in (B) and (C) are presented as mean ± SEM, n = 3, and are representative of at least three independent experiments.

Cells transfected with either (-ss) or (+ss) Gpr126 had higher cAMP levels compared to cells transfected with an empty vector (EV) (Figure 3.7C), demonstrating that basal activity of Gpr126 can be detected in this assay. As a positive control, a synthetic peptide agonist that targets the Gpr126 7TM activated Gpr126 signaling to a level consistent with similar, previously-published experiments (Figure 3.8A) (Liebscher et al., 2014) 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 (Figure 3.7C), and this result was consistent between both zebrafish and human (Figure 3.8D) constructs. 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.

To test the role of the calcium-binding site on signaling, 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 (Figure 3.8E), similar to the wild-type (+ss) isoform. The calcium-binding site mutation did not increase or decrease the cAMP signaling for the (-ss) Gpr126 isoform (Figure 3.7C). To test whether the calcium-binding site is

responsible for the increase in signaling in the (+ss) isoform, we mutated calcium-binding site residues D134A/F135A in the (+ss) isoform as well. This mutation resulted in lower cAMP levels compared to wild-type (+ss). Indeed, in all constructs in which the calcium-binding site is either obscured (wildtype -ss) or mutated (-ss D134A/D135A and +ss D134A/F135A), the cells had similarly low levels of cAMP compared to the cells transfected with a construct in which the calcium-binding site is likely exposed (wildtype +ss). 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 (Figure 3.7A). These results suggest that the calcium-binding site is important for increased signaling in the (+ss) isoform and is likely a functional site for Gpr126.

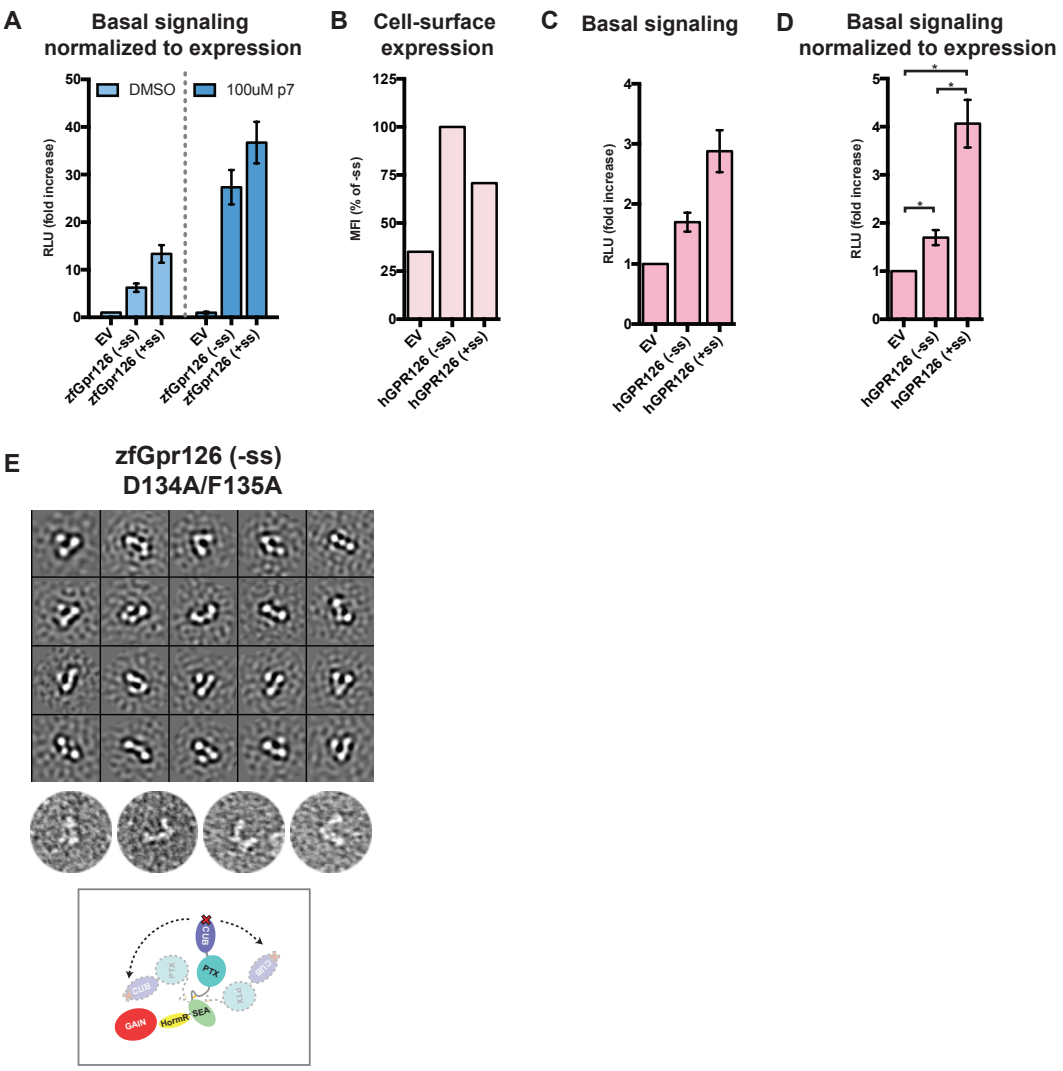


Figure 3.8 Calcium-binding site mutants alter ECR conformation

(A) Basal (left) and activated (right) cAMP signaling normalized to cell-surface expression for empty vector (EV) and zebrafish Gpr126 splice isoforms. (B) Cell-surface expression levels for EV and human 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. (C) Basal signaling for sam constructs in (B) measured by the cAMP signaling assay. Data are shown as fold increase over EV of RLU (relative luminescence units). (D) Basal cAMP signaling normalized to cell-surface expression. *P < 0.05 by unpaired two-tailed t test. Data in (A), (C), and (D) are presented as mean ± SEM, n = 3, and are representative of at least three independent experiments. (E) ECR conformation for -ss D134A/F135A Gpr126 ECR. Negative stain EM 2D class averages (top) and representative individual particles (middle) are shown. ECR conformations (bottom) based on negative stain EM are depicted as cartoons. Arrows with dashed lines indicate dynamic ECR conformation.

The calcium-binding site in the CUB domain is critical for PNS myelination in zebrafish

Ligand-binding sites and other functional sites on proteins are usually highly evolutionarily conserved. Since we identified the calcium-binding site in the CUB domain of Gpr126 as a potential functional site, we wanted to further evaluate its importance from an evolutionary perspective and assess its conservation level compared to the rest of the ECR. We used the ConSurf server (Ashkenazy et al., 2016) to perform surface conservation analyses on a diverse set of 90 Gpr126 protein sequences. The conservation score for each residue was mapped onto the Gpr126 ECR structure (Figure 3.9A), 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 (Figure 3.10A). The calcium-binding site is universally conserved among all species analyzed. The high level of conservation, along with the previous signaling results for calcium-binding site mutant constructs, further suggests that the calcium-binding site has an essential role in Gpr126 function.

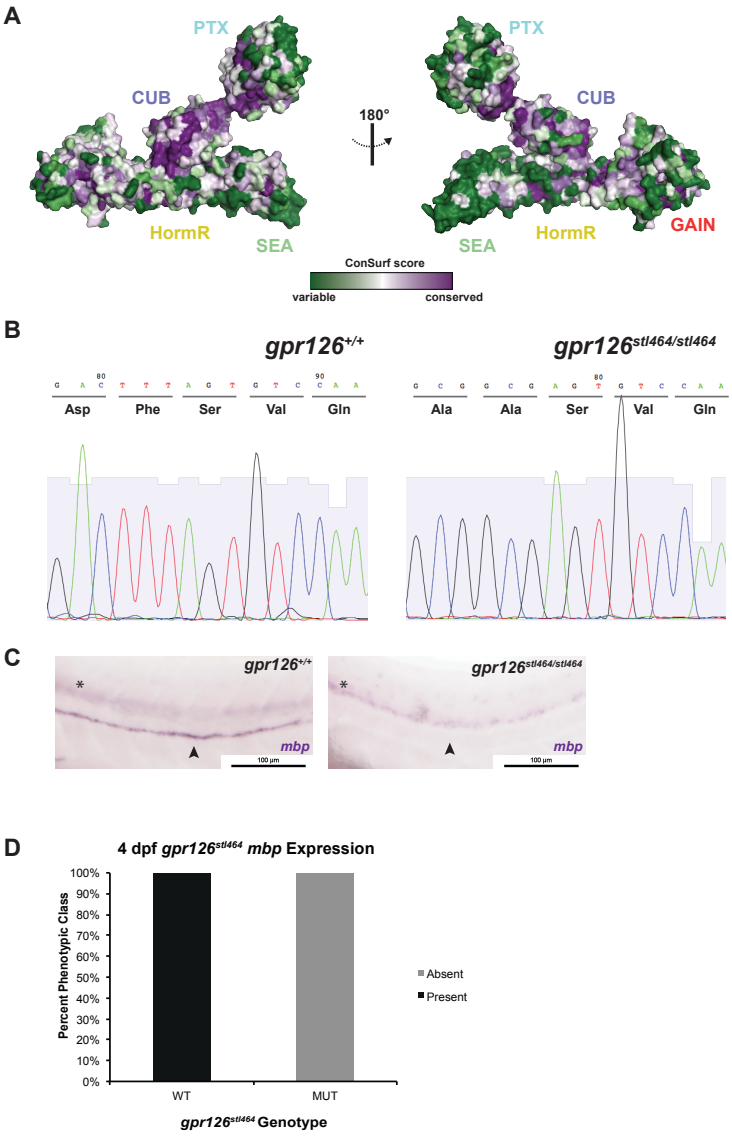


Figure 3.9 Surface conservation of Gpr126 ECR

(A) Conservation score of each residue mapped onto the surface of Gpr126 ECR. (B) Sequence traces for *gpr126*$^{+/+}$ *and gpr126*$^{stl464/stl464}$ zebrafish, showing D134A/F135A mutations in *gpr126*$^{stl464/stl464}$. (C) 4 dpf wild-type larvae express *mbp* throughout the posterior lateral line nerve (PLLn, arrowhead), whereas 4 dpf *gpr126*$^{stl464/stl464}$ larva lack *mbp* expression along the PLLn (arrowhead). Asterisks indicate CNS. (D) Quantification of presence or absence of *mbp* expression in wild-type and mutant larvae.

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 ear development in zebrafish through elevation of cAMP (Geng et al., 2013; Monk et al., 2009). Zebrafish *gpr126* mutations that impair G protein signaling result in abolished myelination of the peripheral axons by SC and cause "puffy" ears (Geng et al., 2013; Liebscher et al., 2014; Monk et al., 2009; Paavola et al., 2014; Petersen et al., 2015). 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 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 (Figures 3.3C and 3.10A). As a result, the mutant zebrafish, *gpr126*^{stl464}, harbor D134A and F135A mutations (Figures 3.10B and 3.9B). These mutations created a BstUI restriction enzyme site, which was used to genotype individual zebrafish (Figure 3.10C). Strikingly, compared to wild-type siblings, the *gpr126*^{stl464} mutant zebrafish developed the puffy ears (Figures 3.10D and 3.10E) that are indicative of a defect in Gpr126-mediated G protein signaling. In addition to the ear phenotype, mutant zebrafish did not express *mbp*, indicative of failed PNS myelination (Figures 3.10F, 3.10G, 3.9C, and 3.9D). These results show that D134 and F135 in the calcium-binding pocket of Gpr126 are essential for ear and SC development *in vivo*.

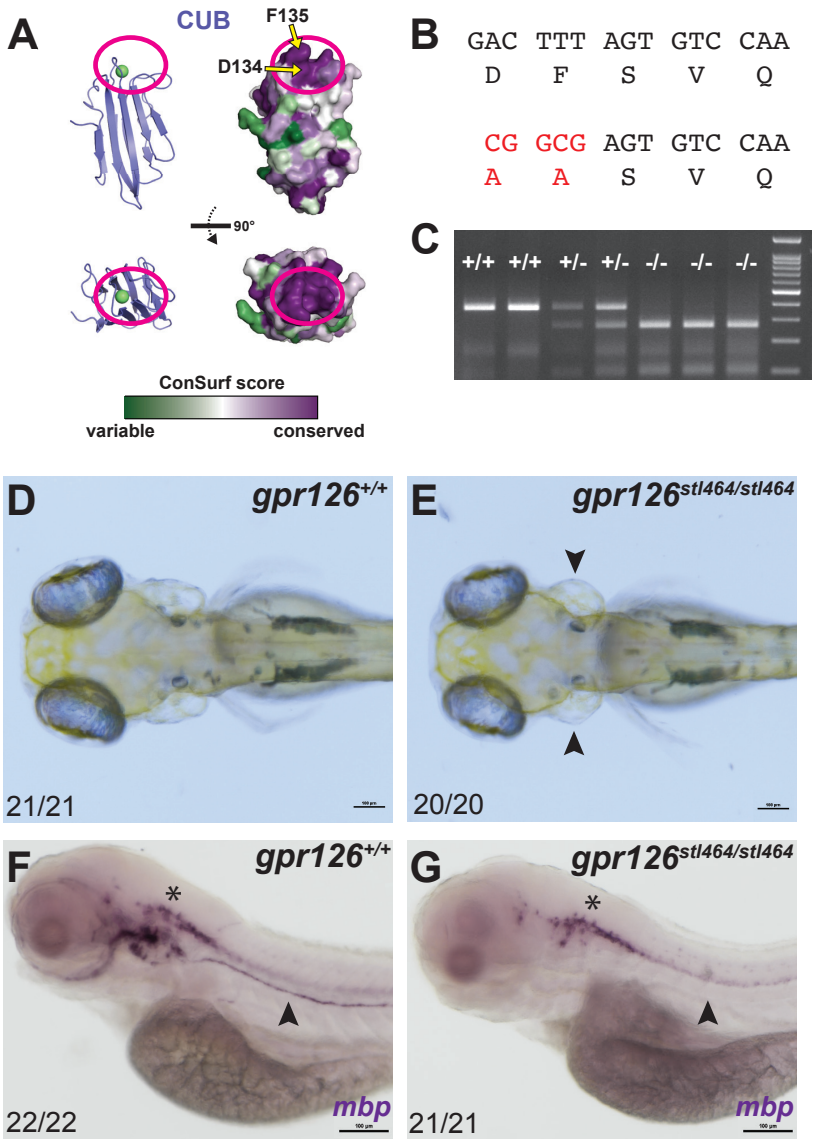


Figure 3.10 Zebrafish with mutations in the calcium-binding pocket in CUB domain show defective ear and Schwann cell development

(A) Surface conservation analysis 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). (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). Asterisks indicate CNS.

A proteolytic SEA domain may mediate domain-specific functions of GPR126

In addition to myelination and ear development, Gpr126 also plays an important role in heart development (Geng et al., 2013; Monk et al., 2009; Paavola et al., 2014; Patra et al., 2013; Waller-Evans et al., 2010). However, in contrast to the myelination and ear development functions of GPR126, which depend on intracellular cAMP signaling (Geng et al., 2013; Mogha et al., 2013; Monk et al., 2009), heart development does not depend on downstream G-protein signaling via the 7TM domain because ectopic expression of Gpr126 ECR suppresses the cardiac phenotype in *gpr126* knockdown in zebrafish (Patra et al., 2013), and *gpr126* mutant zebrafish with truncations after the CUB and PTX domains have normal heart development (Paavola et al., 2014). These observations suggest a versatile function and signaling property of Gpr126 that depends on only the ECR. Although furin-mediated cleavage had previously been proposed to release the CUB and PTX domains to act as ligands on other receptors (Moriguchi et al., 2004), the structural and mechanistic insight remained unclear.

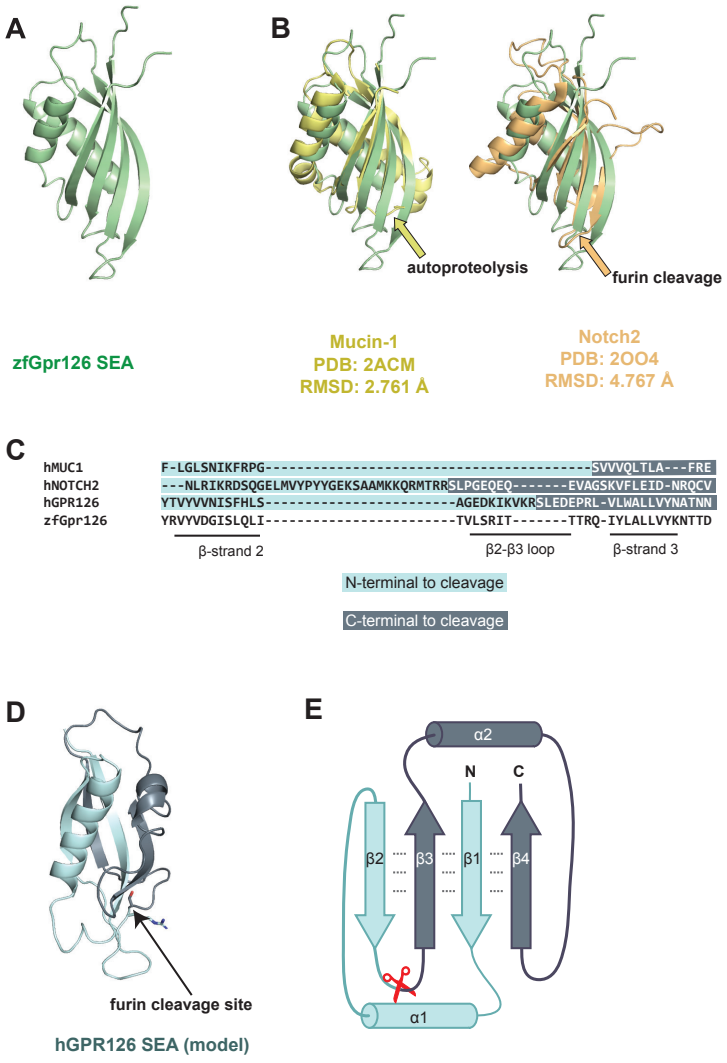


Figure 3.11 A furin cleavage site is mapped within the newly identified SEA domain

(A) Crystal structure of the SEA domain from zebrafish Gpr126. (B) Gpr126 SEA domain superimposed over Mucin-1 SEA domain (left, PDB: 2ACM) and over Notch2 SEA domain (right, PDB: 2OO4). The root-mean-square deviations (RMSDs) of atoms between overlaid structures are 2.761 Å and 4.767 Å, respectively. Arrows indicate cleavage sites in Mucin-1 and Notch2. (C) Sequence alignment of partial SEA domain from human Mucin-1, human Notch2, human GPR126, and zebrafish Gpr126. (D) Homology model of human GPR126 SEA model generated using SWISSMODEL. The arrow points to modelled furin-cleavage site. (E) Protein topology map of SEA domain. Furin cleavage site is indicated by red scissors. Residues N-terminal to cleavage site are dark blue and residues C-terminal to cleavage site are light blue. Dashed lines represent backbone hydrogen bonds between beta sheets.

As mentioned earlier, the previous unknown region in the GPR126 ECR contains a structured domain, which we revealed to be a SEA domain (Figure 3.11A) that is the site of furin cleavage. GPR126 SEA superimposes well over SEA domains from Mucin-1 and Notch-2 (Gordon et al., 2007; Macao et al., 2006), which are also cleaved (via autoproteolysis and furin, respectively), both in the same loop between β-strand 2 and β-strand 3 (Figure 3.11B). Although furin cleavage is not conserved in zebrafish Gpr126, it is conserved in many mammals and birds (Table 3.2 and Figure 3.2A), with a consensus sequence of (R/K)-X-K-R↓. Using sequence alignments (Figure 3.11C) and homology modeling, we mapped the furin-cleavage site in human GPR126 (Figure 3.11D) 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.

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 (Gordon et al., 2009; Macao et al., 2006) and shear forces likely unfold the domain and separate the protein into two fragments (Pelaseyed et al., 2013; van Putten and Strijbis, 2017). The Gpr126 SEA domain shows several noncovalent interdomain interactions, particularly between all four of the β-strands that form a β-sheet (Figure 3.11E), suggesting that separation of the SEA domain into two fragments would not readily occur immediately following cleavage, similar to the aforementioned SEA domains as well as to GAIN domain autoproteolysis. Instead, the two fragments likely stay associated non-covalently 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) (Figure 3.12A). HEK293 cells transfected with a mutant form of human GPR126 that is not able to be cleaved by furin were detected on the cell surface (Figure 3.12B), and therefore, the importance of furin cleavage is likely not primarily important for proper expression and trafficking but rather for a ligand-mediated function.

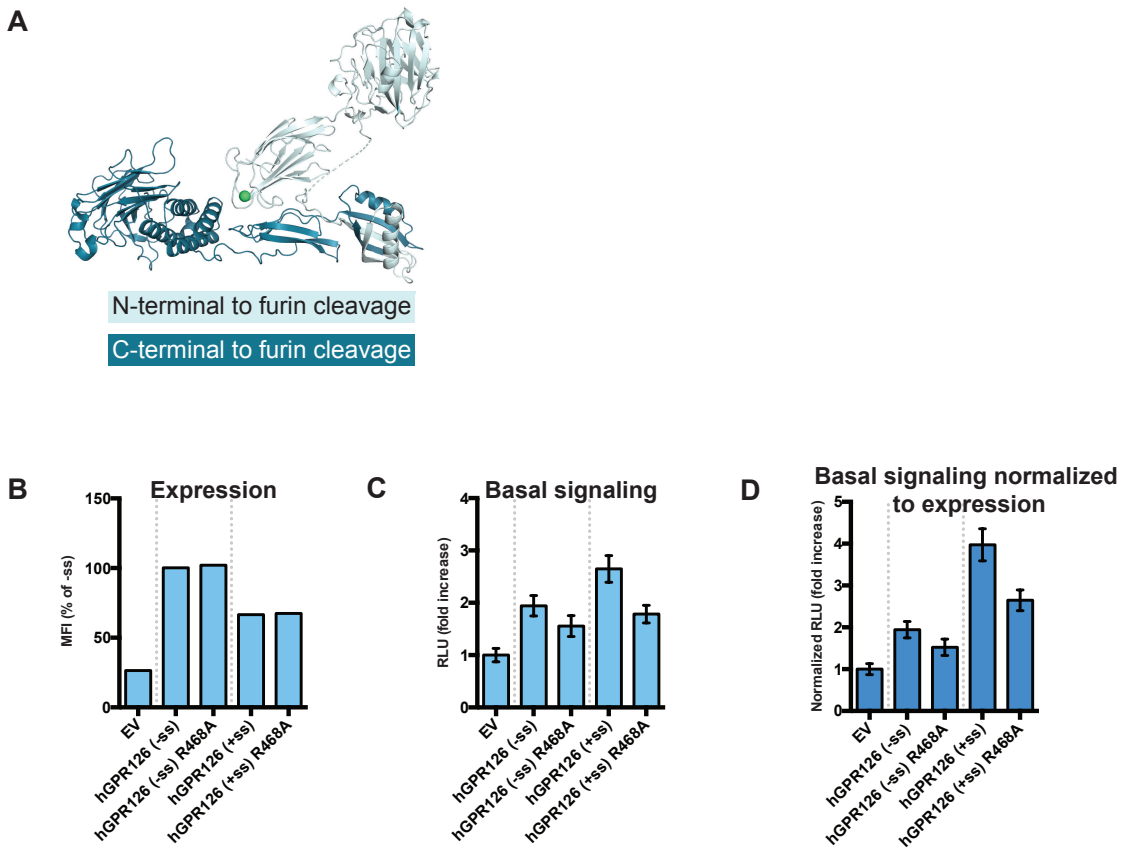


Figure 3.12 Furin cleavage of GPR126

(A) Structure of Gpr126 ECR. The region N-terminal to furin-cleavage site is colored light blue and region C-terminal to furin-cleavage site is colored dark blue. (B) Cell-surface expression levels for empty vector (EV) and various human GPR126 constructs. (C) Basal signaling for constructs in (B), shown as fold increase over EV. (D) Basal signaling normalized to cell-surface expression. Data are presented as mean ± SEM; n = 3.

Discussion

aGPCRs make up the second largest family of GPCRs with 33 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.

Importantly, Gpr126 may be regulating receptor function by modulating ECR conformation. The structure revealed that there is a highly conserved calcium-binding site through which the CUB domain is able to bind the HormR domain. 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* function of Gpr126. A subset of CUB domains from other proteins coordinate calcium in order to mediate ligand-binding (Gaboriaud et al., 2011). Thus, we speculate that Gpr126 uses this site to not only bind to itself in a closed

conformation (-ss isoform), effectively hiding the calcium-binding site from ligands, but also to bind to ligands when the ECR is in a possible open conformation (+ss isoform) in order to activate signaling. Ligand-induced activation mechanisms are currently not well-understood, but likely involve exerting force which leads to shedding at the autoproteolysis site and exposure of the tethered peptide agonist. Another possibility is that ligand binding rearranges ECR-7TM interactions which regulate receptor signaling. At present, three ligands have been reported for GPR126: type IV collagen (Paavola et al., 2014), laminin-211 (Petersen et al., 2015), and the prion protein (Kuffer et al., 2016). Several CUB domains from other proteins, such as C1s and MASP-1, have been reported to bind to collagen-like ligands through calcium coordination (Gingras et al., 2011; Venkatraman Girija et al., 2013), suggesting that Gpr126 may also bind to type IV collagen and the collagen-like region of prion protein in a similar manner.

Since Gpr126 is alternatively spliced in the region encoding the ECR, we examined the functional differences between isoforms. Alternative splicing in proteins is an important mechanism to greatly expand the functional capacity of metazoan genomes, and its regulatory role in brain function has been repeatedly demonstrated. For instance, DSCAMs, protocadherins, calcium channels, neurexins, and neuroligins have been shown to use alternative splicing for diversifying their functions (Aoto et al., 2013; Fuccillo et al., 2015; Irimia et al., 2014; Thalhammer et al., 2017). 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 at least 80 Å away from the 7 aa alternatively spliced site (Li et al., 2018). Our negative stain EM and

SAXS results suggest that alternative splicing between the PTX and SEA domains in Gpr126 perturbs the closed conformation and generates a population of ECR conformations that range from closed to extended. 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.

The structure also revealed the presence of a furin-cleaved SEA domain within the ECR of Gpr126. Cleaved SEA domains from other proteins have been shown to stay intact until a force is applied and pulls apart the fragments (Pelaseyed et al., 2013; van Putten and Strijbis, 2017). Similarly, we propose that Gpr126 regulates its activity by furin-dependent shedding in addition to the established GAIN-autoproteolysis-dependent shedding. Moreover, the released extracellular fragments may act as diffusible ligands and bind to other cell-surface receptors. Other aGPCRs that have SEA domains in their ECRs include ADGRF1/GPR110 and ADGRF5/GPR116 (Lum et al., 2010; Sandberg et al., 2008). Although these SEA domains are not cleaved by furin, they do contain the GSVVV (or GSIVA) motif that leads to autoproteolytic cleavage in the same loop (between β-strand 2 and β-strand 3) that is cleaved by furin in Gpr126. Therefore, SEA domain cleavage, whether by autoproteolysis or by furin, is a common feature in several aGPCRs and may have similar roles in regulating receptor function.

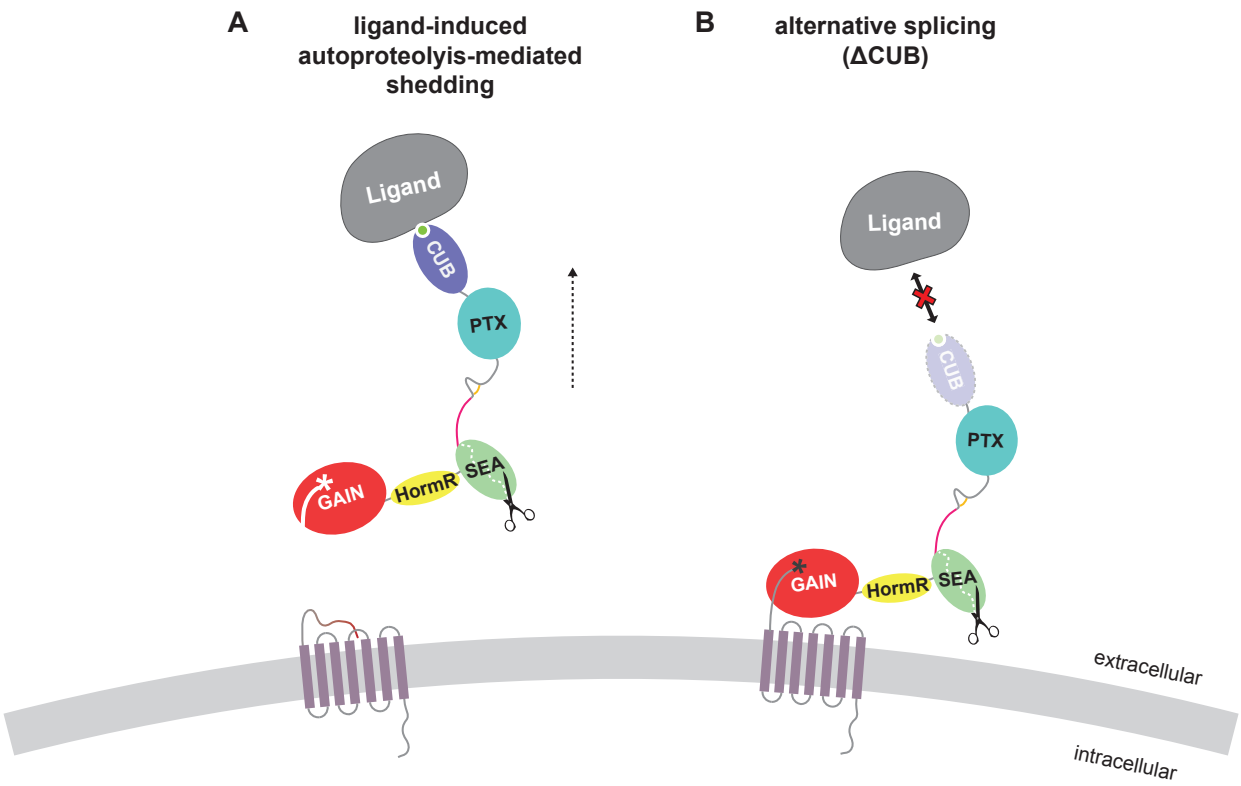


Figure 3.13 Additional models for activation and regulation of Gpr126 function

(A) Ligand binding to ECR of Gpr126 and mechanical force leads to exposure of tethered peptide and activation of the receptor. (B) An alternative splice isoform lacks the CUB domain, precluding ligand binding to calcium-coordination site and/or other regions of CUB domain.

Taken together, our results suggest that Gpr126 is a complex protein that makes use of its many domains to regulate its function. In addition to the autoproteolysis-dependent shedding mechanism (Figure 3.13A), Gpr126 uses several other mechanisms to regulate its function including modulation of the ECR conformation and furin-mediated cleavage. In the closed conformation (Figure 3.14A), Gpr126 signals less compared to when the ECR is in a more dynamic, open conformation (Figure 3.14B). The ECR can adopt a more open conformation by alternative splicing, allowing for access to the calcium-binding site by endogenous ligands (Figure 3.14C). Alternative

splicing which deletes the CUB domain may also preclude ligand binding and activation (Figure 3.13B). Mutation of the calcium-binding site leads to signaling defects *in vitro* and to ear and PNS defects *in vivo* (Figure 3.14D), consistent with our model. In addition, furin cleavage allows Gpr126 another mode of activation, whereby the extracellular fragment preceding the cleavage site may experience mechanical force and be shed, leading to activation of the receptor (Figure 3.14E). Release of extracellular fragments into the extracellular matrix may also regulate other functions, such as heart development, which has been shown to only require the CUB and PTX domains of Gpr126.

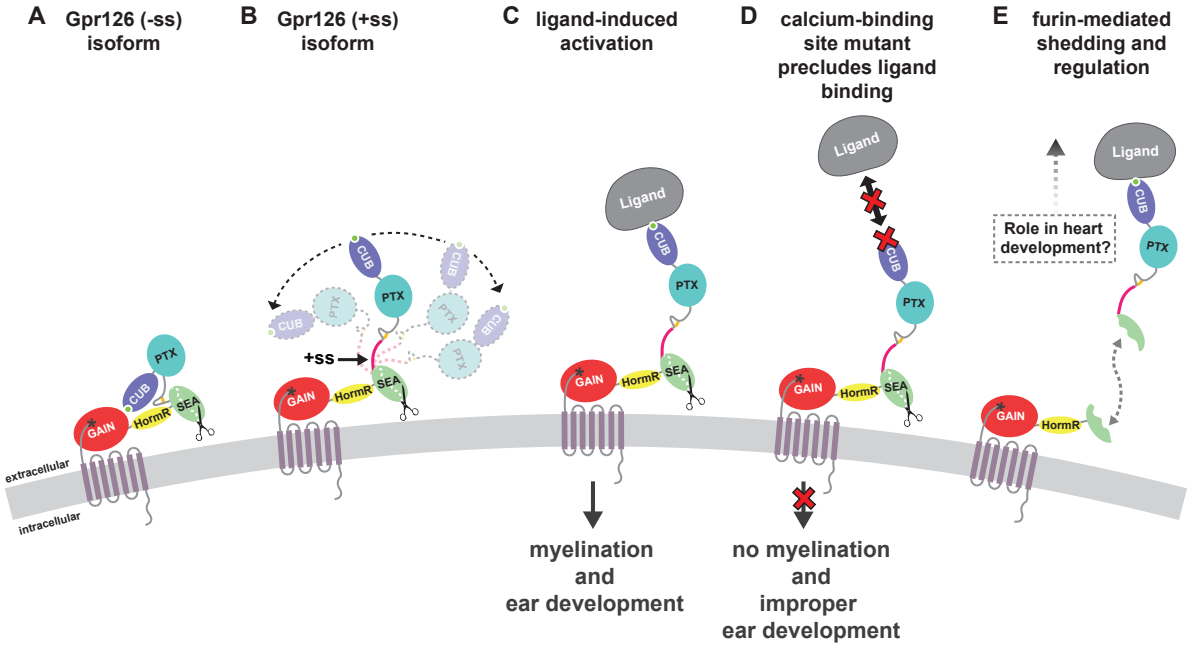


Figure 3.14 Model for ECR-dependent functions of Gpr126

The model depicts how Gpr126 ECR uses various mechanisms to mediate receptor function in PNS myelination, ear development and possibly in heart development. (A, B) Alternative splicing acts as a molecular switch to adopt different ECR conformations and have different basal levels of signaling. Gpr126 ECR that lacks the splice insert adopts a closed conformation and has basal activity (A), whereas Gpr126 ECR that includes the splice insert is more dynamic and open-like, and has enhanced basal activity. (continued on next page)

(continued from previous page) (C) In wild-type zebrafish, Gpr126 can bind to ligands via its calcium-binding site within the CUB domain and activate downstream G-protein signaling leading to wild-type PNS myelination and ear development. (D) Mutation of conserved residues within the calcium-binding site leads to defects in both myelination and ear development *in vivo*. (E) Gpr126 function may also be regulated by furin cleavage, by modulation of receptor signaling and/or release of extracellular fragments. SEA domain cleavage and release of the N-terminal domains is likely critical for heart development.

The Gpr126 closed conformation and hidden calcium-binding site is conceptually similar to EGFR. EGFR is in a closed, compact inactive conformation until ligand binding leads to a conformational change that extends the protein and reveals a hidden functional site that is important for its activation (Cho and Leahy, 2002; Li et al., 2005). Because this mechanism is key for drugging EGFR, the conceptual similarity provides an opportunity to also drug Gpr126. Drugs that alter the ECR conformation of Gpr126 or block ligand-binding sites, such as the calcium-binding site, may be useful for treating Gpr126-associated diseases. The ECRs of other aGPCRs are major players in mediating receptor functions as well. For example, using its ECR, ADGRA2/GPR124 regulates isoform-specific Wnt signaling (Cho et al., 2017; Eubelen et al., 2018; Vallon et al., 2018; Zhou and Nathans, 2014), the *C. elegans* ADGRL1/LAT-1 controls cell division planes during embryogenesis, and ADGRB1/BAI1 and ADGRL3/Lphn3 mediate synapse formation through interaction with other cell-surface proteins (Jackson et al., 2015; Jackson et al., 2016; Lu et al., 2015a; Tu et al., 2018). Thus, the ECRs of other aGPCR family members are also promising drug targets to treat numerous diseases once mechanistic details about their regulatory functions are understood.

Chapter 4: Structural Basis for Regulation of GPR56/ADGRG1 by Its Alternatively Spliced Extracellular Domains[1]

Introduction

GPR56, a regulator of oligodendrocyte and cortical development, belongs to the adhesion G protein-coupled receptor (aGPCR) family, a large family of chimeric proteins that have both adhesion and signaling functions (Hamann et al., 2015; Langenhan et al., 2013). As in the canonical GPCR families, aGPCRs have a seven-pass transmembrane helix bundle (7TM) that, for many aGPCRs, can be activated to initiate a signaling cascade via interactions with cytosolic G proteins. Unlike the canonical GPCR families, aGPCRs also have large and diverse extracellular regions (ECRs), mainly composed of domains generally involved in adhesion-related functions (Langenhan et al., 2013). Although this architecture suggests ECRs have functional importance, their biological roles are incompletely understood. In this study, we set out to determine the 3D structure of the entire ECR of GPR56 at atomic resolution. To this end, monobodies that recognize the ECR of GPR56 were engineered. Monobodies are

[1] Part of the text from this chapter was taken verbatim (with minor changes) from: Salzman GS, Ackerman SD, Ding C, Koide A, Leon K, Luo R, Stoveken HM, Fernandez CG, Tall GG, Piao X, Monk KR, Koide S, Araç D (2016). Structural Basis for Regulation of GPR56/ADGRG1 by Its Alternatively Spliced Extracellular Domains. *Neuron*. 91(6):1292-1304. My contributions to this work include cloning and expressing full length GPR56 in insect cells to be used in direct G protein coupling experiments (performed by Stoveken HM and Tall GG).

synthetic binding proteins based on the human fibronectin type-III (FN3) scaffold (Koide et al., 1998), which have recently emerged as powerful tools to facilitate structure determination as ‘‘crystallization chaperones.’’ Monobodies can also act as agonists or antagonists, further underscoring their utility in probing the function of a given protein (Sha et al., 2013; Stockbridge et al., 2015; Wojcik et al., 2010). The structure revealed the identity and boundaries of two extracellular domains: a previously unidentified N-terminal domain with low homology to all known folds and a short but functional GAIN domain. Notably, it was discovered that the entire newly defined N-terminal domain was deleted in GPR56 splice variant 4 (S4), but not the other variants, and deletion of this domain increased basal activity of the receptor. Finally, a highly conserved, surface exposed patch on the N-terminal domain was identified, mutation of which abolished GPR56 function *in vivo*. Together, these results elucidate the multifaceted manner by which the ECR regulates GPR56 function and broadens our understanding of aGPCR biology and oligodendrocyte development.

My contributions to this project are as follows: In order to test whether the GPR56 ECR regulates 7TM signaling, the full-length mouse GPR56 was cloned and expressed in insect cells. GPR56-containing cell membranes were then shown to activate G_{13}, which is upstream of RhoA and serum response element (SRE). This result then led to the development of the SRE-luciferase assay used to measure G_{13} G protein signaling of HEK293T cells overexpressing wild-type or mutant GPR56 constructs, a vital tool for studying GPR56 signaling.

Results

Expression and direct G protein coupling of mouse GPR56

An important first step in studying the biological function of aGPCRs is to clone and express the full-length receptors so that they can be used in biochemical and structural experiments. To aid in studying GPR56 function, the full-length mouse GPR56 was cloned, expressed, and extracted from insect cells.

In order to develop signaling assays to test G protein signaling of mouse GPR56, we first evaluated the G proteins to which mouse GPR56 can couple. Human and mouse GPR56 are both reported to activate G_{13} (Figure 4.1A), which is upstream of RhoA and serum response element (SRE) (Luo et al., 2011; Stoveken et al., 2015). Interestingly, we found mouse GPR56 weakly coupled to an additional G protein, G_q, but not $G_{i/o}$, which can couple to human GPR56 (Figure 4.1B-D) (Stoveken et al., 2015), likely illustrating different roles for GPR56 across species. Based on these results, we used an SRE-luciferase assay to measure G_{13} G protein signaling of HEK293T cells overexpressing WT or mutant GPR56 constructs to better understand how the ECR regulates GPR56 signaling.

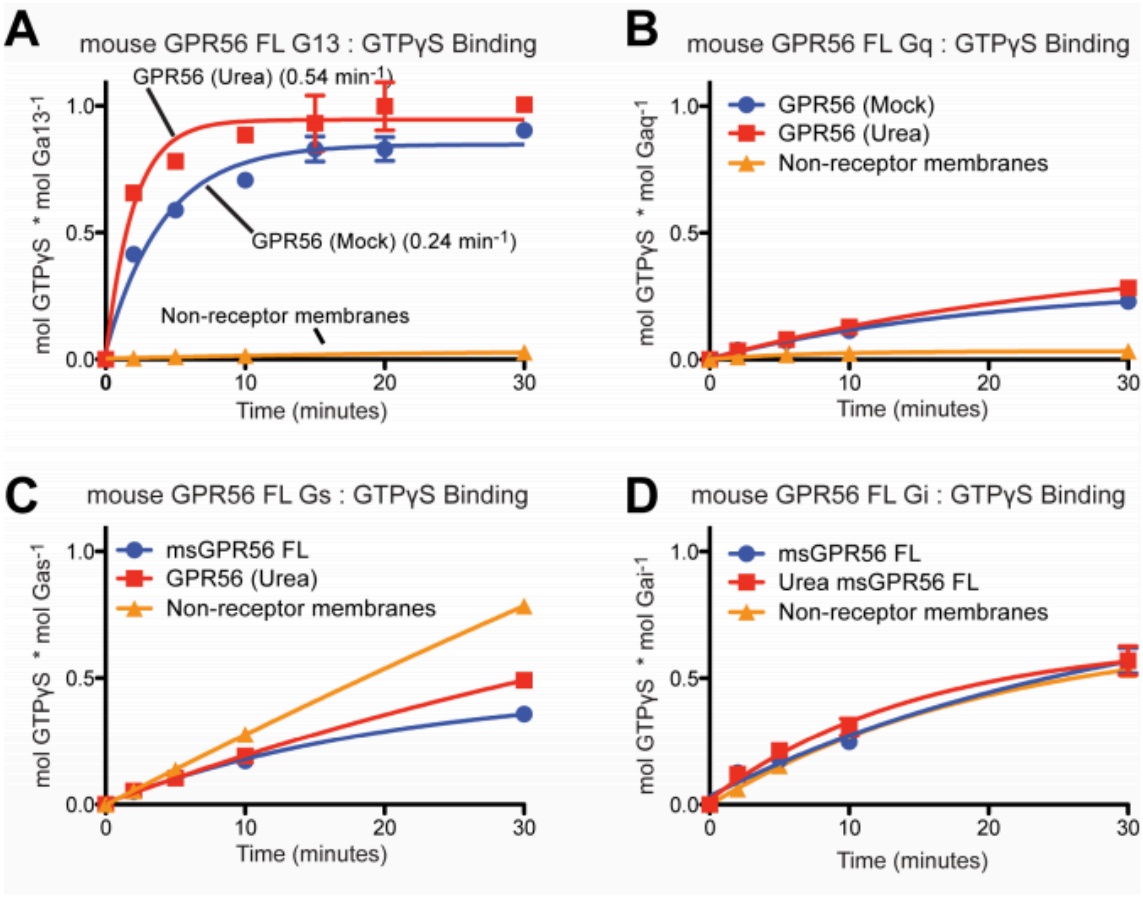


Figure 4.1 G protein coupling of mGPR56

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 βγ). The receptor-mediated G protein activation kinetics were measured using the [35S]-GTPγS binding assay. (A) GPR56 and G_{13}. (B) GPR56 and G_q. (C) GPR56 and G_s. (D) GPR56 and G_i. Data are presented as mean ± S.E.M.

Conclusion

Since aGPCRs are not as well-studied as other GPCR families, there is a lack of fundamental knowledge of the functional properties of aGPCRs. In this chapter, we focused on the important first steps required for studying the signaling function of an aGPCR with essential roles in oligodendrocyte and cortical development, GPR56. We first cloned and expressed the full-length mouse GPR56 in insect cells. Expression of

the GPR56 in this system allowed us to probe receptor function biochemically. For instance, we performed direct G protein coupling experiments with GPR56 expressed in insect cell membranes, which provided us with novel information about the signaling pathways of mouse GPR56. The strong coupling between GPR56 and G_{13} resulted in the development of an SRE-luciferase signaling assay, a powerful method for analyzing the effects of perturbations to the ECR of GPR56 on G protein signaling. In addition, the ability to express and extract the full-length cleaved receptor from insect cell membranes will benefit future structural and functional studies of GPR56 and other aGPCRs. For example, the detergent-solubilized and purified receptor will be used in crystallography or cryo-electron microscopy trials in order to determine the structure of the 7TM domain, which remains elusive for aGPCRs at present.

Chapter 5: Structural Basis for Teneurin Function in Circuit-Wiring: A Toxin Motif at the Synapse[1]

Introduction

Teneurins (TENs) are evolutionarily conserved cell-adhesion molecules that play a central role in embryogenesis, axon guidance, and synapse formation (Leamey and Sawatari, 2014; Tucker and Chiquet-Ehrismann, 2006). They are unusual because they are large type-II transmembrane proteins (>2,000 amino acids) that lack the requisite domains generally observed in cell-adhesion molecules, such as classical immunoglobulin (Ig), cadherin, LNS (laminin-a, neurexin, and sex hormone-binding globulin), or integrin domains. TENs have been implicated in synapse formation due to its strong binding to latrophilins (Lphns), adhesion G-protein-coupled receptors that are localized in synapses (Anderson et al., 2017; Boucard et al., 2014; Silva et al., 2011). Given that TENs are also localized to synapses and at least Lphn2 has been shown to be essential for hippocampal synapse formation (Anderson et al., 2017), a role for the heterophilic TEN-Lphn complex in synapse formation is plausible.

Despite their central importance in multiple physiological roles, the lack of information on the structure of TENs is one of the limiting factors in delineating their

[1] Part of the text from this chapter was taken verbatim (with minor changes) from: Li J, Shalev-Benami M, Sando R, Jiang X, Kibrom A, Wang J, Leon K, Katanski C, Nazarko O, Lu Y, Südhof T, Skiniotis G, Araç D (2018). Structural Basis for Teneurin Function in Circuit-Wiring: A Toxin Motif at the Synapse. *Cell*. 173(3):735-748. My contributions to this work include specimen screening by negative stain electron microscopy.

mechanisms of action. The 3.1 Å cryo-EM structure of the large ECR of human TEN2 was determined. The ECR has an unusual architecture whereby a large cylindrical β-barrel with clear similarity to bacterial Tc-toxins partially encapsulates a C-terminal toxin-like domain that emerges from the barrel and is tethered to its outer surface. The tethered C-terminal domain was found to mediate interactions with Lphn and these interactions activate trans signaling by controlling intracellular cAMP levels. In addition, a splice variant of TEN2, which involves a 7-amino-acid extension in a β-propeller domain region, was incapable of mediating interactions with Lphn and instead induces inhibitory postsynaptic specifications through interactions with alternative synaptic partners. These results reveal that TEN2, and by extension other TENs, forms highly unusual structures consistent with biological activities spanning a range of function.

My contributions to this project are as follows: To determine the structure of TENs, the first steps are to purify TEN ECR constructs and screen samples using negative stain EM to optimize sample conditions. Then, the optimized constructs and sample conditions can be applied to cryo-EM in order to obtain a high-resolution model of TEN.

Results

Negative stain electron microscopy of the extracellular region of teneurin

Various TEN ECR constructs, including the full length TEN ECR and the truncated TEN ECRΔ1 were purified and applied to negative stain EM grids and overlaid with uranyl formate. Examination of the grids with an electron microscope

revealed that both constructs were pure and homogenous (Figure 5.1). A concentration of 10 ug/mL protein was ideal for number of particles per grid and amount of spacing between particles.

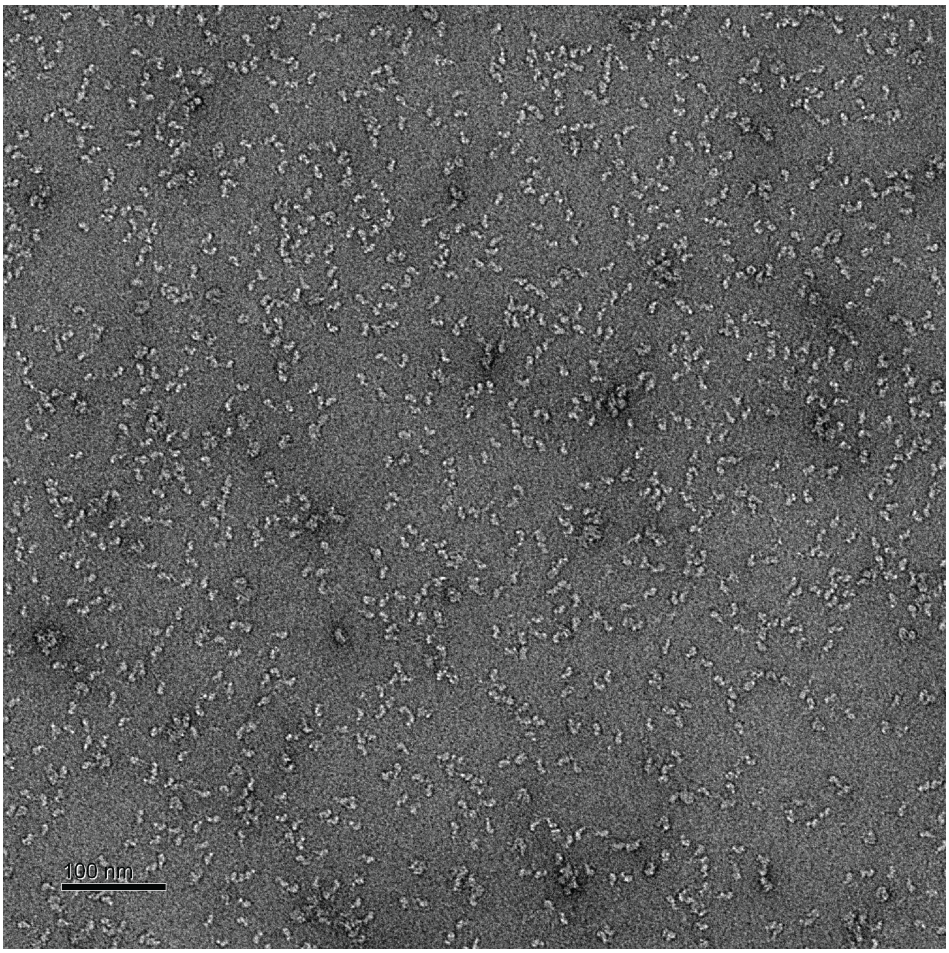


Figure 5.1 Initial negative stain electron micrograph of TEN2 ECR

Purified TEN2 ECR (10 ug/mL) applied to an EM grid and overlaid with uranyl formate, imaged by electron microscope at 23,000x magnification. Scale bar is 100 nm.

TENs form constitutive cis-dimers via highly conserved disulfide bonds formed between their ECRs in close proximity to the transmembrane pass (Figure 5.2A) (Feng et al., 2002; Oohashi et al., 1999; Vysokov et al., 2016). Further examination of the

purified human TEN2 ECR by negative stain EM showed cis-dimerizaton with two large globular domains extended by thin flexible linkers that are tied together through a stalk-like region at their end (Figure 5.2B). Examination of the TEN ECRΔ1 construct, which deletes the stalk-like region yielded a soluble monomeric protein (Figure 5.2C) with a molecular mass of 215 kDa that was suitable for high-resolution structural studies.

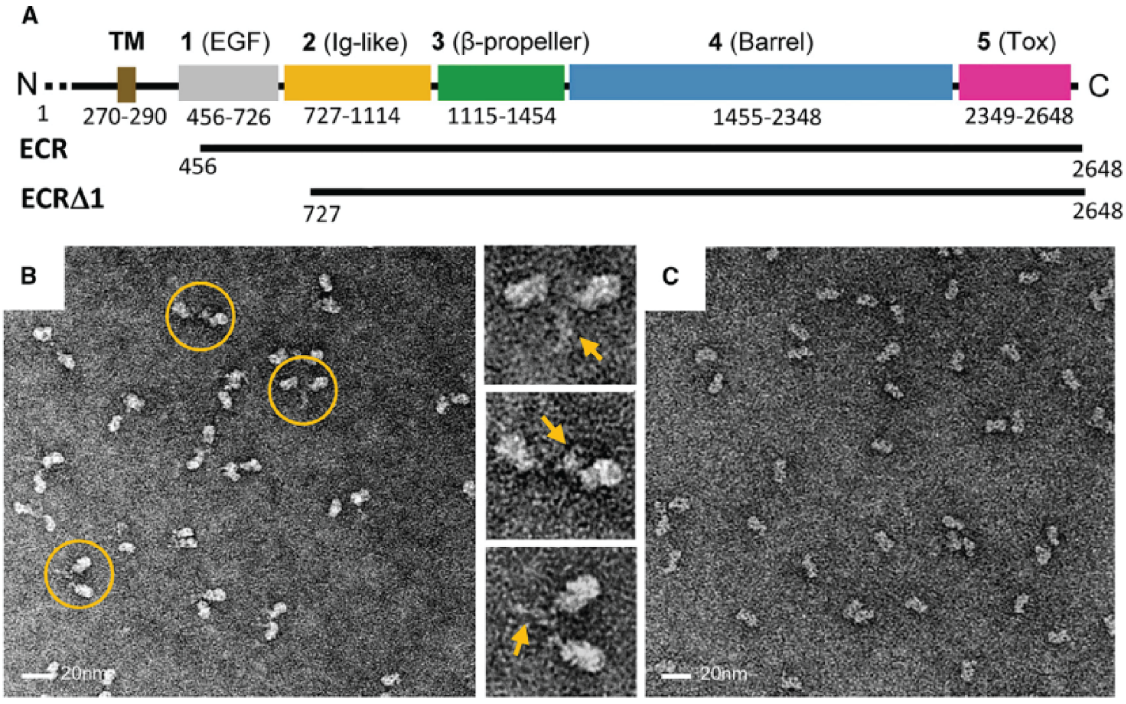


Figure 5.2 Comparison of TEN2 ECR constructs using negative stain electron microscopy

(A) Domain composition of TEN2 and constructs used in the present study. Extracellular domains are colored gray, yellow, green, blue, and magenta for domains 1–5, respectively; transmembrane region (TM) in gold. Domain numbers and descriptions are indicated in black above scheme. Domain borders are indicated below. (B and C) Representative negative-stain EM micrographs of TEN2 ECR (B) and ECRΔ1 (C) indicating dimerization through the ECR domain 1 (B). Scale bar, 20 nm. Close up views of selected dimers (1–3) are indicated to the right with domain 1 dimerization pointed with yellow arrow. ECRΔ1 is of monomeric nature (C).

Conclusion

The identification of endogenous ligands that can activate GPCRs is an important step for understanding GPCR biology. The lack of information about endogenous ligands for many aGPCRs has hindered the study of this family of receptors, and thus, the identification and characterization of aGPCR ligands is one of the major goals for the aGPCR field. The discovery that TENs are ligands for Lphns led to many cellular and biochemical experiments that shed light on TEN-Lphn function. However, the lack of information on the structure of TENs is one of the limiting factors in delineating their mechanisms of action. In this chapter, we focused on the early steps of determining the high-resolution cryo-EM structure of TEN2 ECR. Purified TEN2 ECR, which includes a stalk-like region which mediates dimerization, and TEN2 ECRΔ1, which does not include this region, were screened using negative stain EM. This initial screening allowed us to evaluate the purity and overall quality of the protein samples. In addition, we found that the TEN ECRΔ1 construct was more suitable for high-resolution structural studies due to its monomeric form. This initial screening then led to a cryo-EM structure, which provided new information about the architecture of TEN2, which interestingly has homology with bacterial toxins, as well as the consequence of alternative splicing of TEN on ligand interactions and synapse formation.

Chapter 6: A Comprehensive Mutagenesis Screen of the Adhesion GPCR Latrophilin-1/ADGRL1[1]

Introduction

Adhesion G protein-coupled receptors (aGPCRs) are cell-surface molecules that mediate intercellular communication via cell-cell and cell-matrix interactions (Hamann et al., 2015; Promel et al., 2013). With 33 members in humans, they make up the second largest GPCR family, but are the least studied and least understood (Fredriksson et al., 2003b). Genetic studies suggest critical roles for aGPCRs in development and immunity and especially in neurobiology (such as brain development (Bae et al., 2014; Piao et al., 2004), synapse maturation and elimination (Bolliger et al., 2011), myelination of neurons (Monk et al., 2009), central nervous system angiogenesis (Kuhnert et al., 2010), and neural tube development (Chae et al., 1999; Langenhan et al., 2009; Shima et al., 2004; Usui et al., 1999)), and link them to numerous diseases including neurodevelopmental disorders, deafness, male infertility, schizophrenia, and immune disorders (Langenhan et al., 2013; Promel et al., 2013). In addition, many aGPCRs are found to be over- or underexpressed in various cancers (Kan et al., 2010; Shashidhar et al., 2005; Xu et al., 2006), and a recent study reports that aGPCRs are some of the frequently mutated

[1] Part of the text from this chapter was taken verbatim (with minor changes) from: Nazarko O, Kibrom A, Winkler J, Leon K, Stoveken H, Salzman G, Merdas K, Lu Y, Narkhede P, Tall G, Prömel S, Araç D (2018). A comprehensive mutagenesis screen of the adhesion GPCR Latrophilin-1/ADGRL-1. *iScience*. 3:264-278. My contributions to this work include cloning and expressing full length Lphn3 in insect cells to be used in direct G protein coupling experiments and assisting with peptide-binding experiments.

genes in cancerous tumors (O'Hayre et al., 2013). Considering that many drugs target the 7TMs of GPCRs to regulate receptor activity, thereby eliciting the desirable therapeutic effects, aGPCRs may be promising targets for drugs to treat numerous diseases including cancer. Currently, there is no high-resolution structure for the 7TM domain of an aGPCR. GPCRs from the secretin family have the highest 7TM domain similarity to aGPCRs, suggesting that aGPCR 7TM domains might be activated via similar mechanisms (Fredriksson et al., 2003b; Hollenstein et al., 2013; Rasmussen et al., 2011b; Rosenbaum et al., 2007; Siu et al., 2013). In addition, the signaling pathways of aGPCRs are largely unknown, making any functional studies difficult to perform. Altogether, further studies of aGPCRs progress slowly owing to these obstacles, and the molecular mechanisms underlying aGPCR signal transduction remain unknown.

In this study, latrophilin-1 (Lphn1)/ADGRL1, a key molecule in synapse formation and brain development, was used as a model aGPCR to study aGPCR function. Two robust in vitro assays were established to monitor receptor signaling and showed that the endogenous agonist of Lphns, a 14-amino-acid peptide, binds to and activates the receptor. A large set of bioinformatics-based point mutations and disease mutations on the 7TM region of Lphns were studied using these assays. In addition, mutants that are constitutively active, constitutively inactive, or nonresponsive to the agonist peptide were identified. Intriguingly, these experiments identified a cancer-associated mutation that exhibited high basal activity and abolished the rescue of the embryonic development phenotype in transgenic worms. These results provide the basic groundwork for future drug design against aGPCRs.

My contributions to this project are as follows: To better study the molecular mechanism of Lphn signaling, human Lphn was cloned and expressed in insect cells. In G protein coupling assays, the Lphn-containing membranes were shown to couple to G_i, which inhibits the production of cAMP by adenylyl cyclase. We also crosslinked Lphn to a synthetic peptide agonist which mimics the endogenous *Stachel* peptide, which confirmed the binding of the peptide to the receptor. From these results, we were able to develop a comprehensive mutagenesis screen using a cAMP assay in order to identify transmembrane domain residues essential for Lphn basal activity and for agonist peptide response.

Results

Expression and direct G protein coupling of human Lphn

An important first step in studying the biological function of aGPCRs is to clone and express the receptors so that they can be used in biochemical and structural experiments. To aid in studying Lphn function, a human Lphn construct was cloned, expressed, and extracted from insect cells.

In order to develop signaling assays to test G protein signaling of human Lphn, we performed direct G protein coupling experiments using human Lphn embedded in insect cell membranes. Comparison of empty insect cell membranes with those embedded with Lphn showed an increase in binding to purified G_i proteins, suggesting that Lphn couples to G_i (Figure 6.1). When membranes were treated with urea to induce dissociation of the ECR and *Stachel* exposure, we observed increased coupling to G_i,

suggesting that Lphns are sensitive to *Stachel*-mediated activation in this assay. Based on these results, we used a cAMP-luciferase assay to measure G_i G protein signaling of HEK293 cells overexpressing WT or mutant Lphn constructs to better understand how the ECR regulates Lphn signaling.

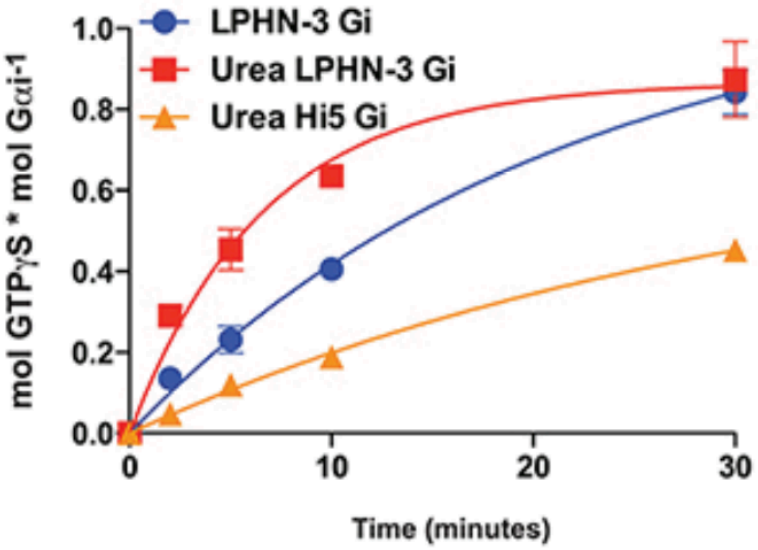


Figure 6.1 G protein coupling of human Lphn

High-Five cell membranes with Lphn were subject to urea treatment to induce ECR-shedding and compared to membranes with no receptor or no treatment. Membranes were reconstituted with purified heterotrimeric G proteins, including the αi subunit and the βγ subunits. The receptor-mediated G protein activation kinetics were measured using the [35S]-GTPγS binding assay for G_i. Data are presented as mean ± S.E.M.

Stachel peptide binds and activates the 7TM domain of human Lphn

Several aGPCRs such as GPR126, GPR133, GPR64, GPR114, GPR56, GPR110, and GPR116, have been reported to be activated by their *Stachel* peptides, which function as tethered agonists (Demberg et al., 2015, 2017; Kishore et al., 2016; Liebscher et al., 2014; Muller et al., 2015; Scholz et al., 2017; Stoveken et al., 2015) To

test whether Lphns also bind to their *Stachel* peptides and to optimize our signaling assays for detecting *Stachel* peptide-dependent activation, we wanted to confirm the direct interaction between the *Stachel* peptide and the 7TM domain of Lphn.

Human Lphn containing only residues C-terminal to the autoproteolysis site was cloned, expressed, and extracted from insect cells. The detergent-solubilized and purified receptor was mixed with synthesized *Stachel* peptide in the presence or absence of a cross-linking reagent, and peptide binding was detected by western blot analysis (Figure 6.2) against both the biotin-tag on the *Stachel* peptide and the His-tag on Lphn. The peptide was indeed found to cross-link with Lphn in the presence of cross-linking reagent, which suggests that Lphn is similar to other aGPCRs that have been confirmed to bind their *Stachel* peptides. These results allowed us to then optimize cAMP-based signaling assays, which utilize the synthesized *Stachel* peptide as an activating ligand.

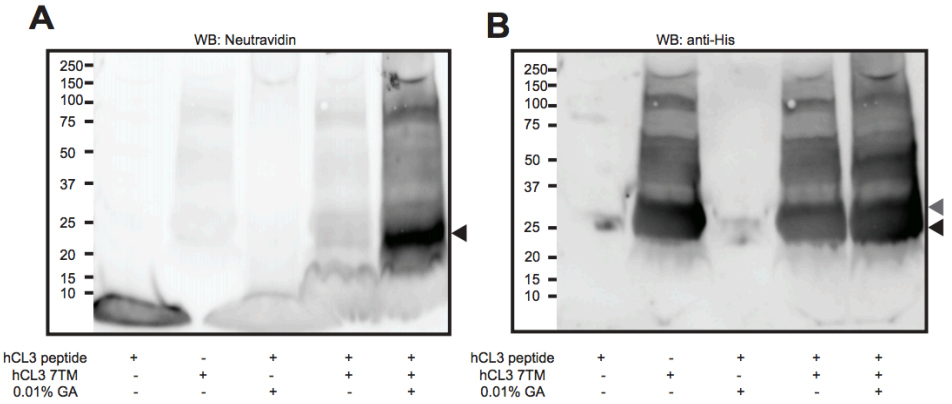


Figure 6.2 *Stachel* peptide binds and activates the 7TM domain of human Lphn

(A) Synthesized human Lphn *Stachel* peptide can be crosslinked to purified human Lphn 7TM in the presence of glutaraldehyde. Western blot of Lphn peptide, Lphn 7TM, and mixture of Lphn peptide and Lphn 7TM in presence or absence of 0.01% glutaraldehyde blotted with (A) labeled NeutrAvidin to detect biotinylated *Stachel* peptide, black arrow, or (B) anti-His antibody to detect His-tagged hLphn 7TM, black and gray arrows indicate Lphn 7TM with differing levels of N-linked glycosylation.

Conclusion

GPCRs are the largest family of cell-surface receptors and are the targets for approximately 30% of currently marketed drugs that treat a very broad spectrum of diseases including neuropsychiatric, cardiovascular, pulmonary, and metabolic disorders; cancer; and acquired immune deficiency syndrome (AIDS). Adhesion GPCRs recently emerged as a novel subfamily of GPCRs with important roles in diverse cellular processes, and they carry the potential to be the targets for the next generation of drugs. However, a molecular understanding of their activation mechanisms is a first step for moving forward. In this chapter, we focused on a model aGPCR subfamily, Lphns, and performed the initial biochemical characterizations towards understanding the signaling mechanisms of this family of receptors.

First, we cloned and expressed human Lphn in insect cell membranes and performed direct G protein coupling experiments, which showed that Lphn couples to and activates G_i proteins. This result allowed us to develop a cAMP-based signaling assay which measures the signaling activities of Lphn constructs expressed in HEK293 cells. The signaling assay is key for studying Lphn function, since we can now test the effects of mutations of residues essential for receptor activity as well as mutations caused by cancers. In addition, we can test the response to the *Stachel* agonist peptide and delineate which residues are important for receptor activation.

We also cloned and expressed into insect cells a truncated Lphn. We extracted and purified the protein from the membranes and were able to cross-link it to its synthetic *Stachel* peptide, which confirmed for the first time that there is an interaction

between the Lphn 7TM and its *Stachel* peptide and provided insight into the activation mechanism of Lphn.

Altogether, these results provide the groundwork for a thorough characterization of the model aGPCR family, Lphns. The data shed light on signaling pathways and *Stachel* peptide-mediated activation mechanism of aGPCRs. In addition, the ability to express and extract Lphn constructs from insect cell membranes will benefit future structural and functional studies of Lphn3 and other aGPCRs. The detergent-solubilized and purified receptor will be used in crystallography or cryo-electron microscopy trials in order to determine the structure of the 7TM domain, which remains elusive for aGPCRs at present.

Chapter 7: Structural Studies on the Transmembrane Domains of aGPCRs[1]

Introduction

In recent years, several high-resolution crystal structures of the 7TM domains of GPCRs from many families have given insight into their signaling mechanisms (Hollenstein et al., 2013; Kruse et al., 2012; Rasmussen et al., 2011a; Siu et al., 2013; Wang et al., 2013a; White et al., 2012; Yin et al., 2014; Zhang et al., 2012). However, there is currently no structural information available for the aGPCR family, and therefore, its signaling mechanism remains largely unknown. It is likely that aGPCRs have distinct signaling mechanisms due to their uniquely large extracellular domains and lack of conserved motifs found in the 7TMs of the other families (Krishnan et al., 2012).

The structural organization of aGPCRs consists of a large ECR, the 7TM, and an intracellular region (Lagerstrom and Schioth, 2008). Present in almost every aGPCR is the recently discovered and conserved GAIN domain that resides within the extracellular region directly preceding the 7TM (Arac et al., 2012). During protein biosynthesis, the GAIN domain is required for cleavage between the ECR and the 7TM at a conserved site within the GAIN domain called the GPCR proteolysis site. Cleavage creates two fragments: an N-terminal fragment (NTF) consisting of the extracellular

[1] The negative stain EM experiments in this chapter were performed and analyzed by Jeffrey Tarrasch and Georgios Skiniotis at the University of Michigan.

region up to the cleavage site and a C-terminal fragment (CTF) consisting of ~15 residues (hereafter referred to as “tethered peptide”) attached to the 7TM and intracellular region (Lin et al., 2004).

The crystal structures of GAIN domains from aGPCRs Lphn1, BAI3 (Arac et al., 2012), and GPR56 (Salzman et al., 2016) indicate that autoproteolysis does not lead to dissociation of the NTF and the CTF. Rather, the tethered peptide is a structured β-strand that is held tightly within the GAIN domain through hydrogen bonds and hydrophobic interactions which allows the NTF and CTF to remain non-covalently associated (Arac et al., 2012). The current model of aGPCR receptor activation suggests that binding of a protein ligand to the NTF exposes the tethered peptide, which can interact with the 7TM to induce conformational changes (Liebscher et al., 2014; Luo et al., 2014), a mechanism similar to that of the protease-activated receptors from the rhodopsin family (Zhang et al., 2012).

Although the GAIN domain has been shown to be necessary and sufficient for autoproteolysis (Arac et al., 2012), the GAIN domain may also play additional roles in aGPCR signaling. Because the GAIN domain always precedes the 7TM and because there are GAIN-domain containing aGPCRs that are not cleaved, such as GPR124, it is possible that the GAIN domain has dual functions that regulate signaling. For example, the GAIN domain may interact with the 7TM and act as an antagonist, precluding agonist binding to the receptor. Structural information for the GAIN and 7TM domains would shed light on whether there is indeed a regulatory interaction between the domains.

Regulation of receptor signaling may also occur through receptor oligomerization (Szidonya et al., 2008). This concept has been well established for many growth factor receptors and cytokine receptors (Bouvier, 2001). However, the existence of oligomerization in GPCRs has been controversial. Recently, there has been convincing evidence for functionally relevant GPCR oligomers (Liebscher et al., 2013). Examples of oligomers affecting signaling include cooperativity in ligand-binding and changes in G-protein selectivity (Szidonya et al., 2008). There is currently no evidence of oligomerization for any aGPCR, but because aGPCR 7TM domains have not been extensively studied, oligomerization remains a potential regulatory mechanism for aGPCR function.

Numerous biological functions are associated with aGPCRs, including brain development, neutrophil migration, and angiogenesis (Paavola and Hall, 2012). Though the exact functions of many aGPCRs remain unclear at present, there is overwhelming evidence that mutations in aGPCRs that disrupt their functions lead to various diseases (O'Hayre et al., 2013). For example, Lphn3 and GPR56 mutations are linked to attention deficit hyperactivity disorder (Arcos-Burgos et al., 2010) and bilateral frontoparietal polymicrogyria (Chiang et al., 2011), respectively. Obtaining structural information for these receptors will aid in uncovering their signaling mechanisms, delineating how the disease mutations affect signaling, and engineering therapeutic drugs.

In this chapter, we discuss the initial steps towards structural characterization of aGPCRs, including 1) the cloning, expression, and purification of aGPCR constructs in insect cells and 2) negative stain electron microscopy studies of an aGPCR construct.

These steps will lay the groundwork for future studies leading to high-resolution structures of the 7TM domains of aGPCRs.

Results

Optimization of expression and purification of aGPCR transmembrane domains

Key to pursuing structural studies of membrane proteins is the ability to express and isolate sufficient amounts of pure, stable protein samples. We screened sixteen different aGPCRs (human ETL, human GPR133, human GPR126, xenopus GPR124, human GPR116, mouse GPR110, human GPR64, human EMR2, human CD97, rat Lphn1, human Lphn3, drosophila Lat-1, C. elegans Lat-1, human BAI3, mouse GPR56, and C. elegan FMI) and different truncated versions of these aGPCRs to find the best expressing aGPCR in our insect cell baculovirus system.

Insect cell membranes containing the best constructs were solubilized with various detergents at different concentrations to obtain the optimal detergent for extraction and for providing stability to the protein. From western blot analysis against the N-terminal FLAG tag (detecting both uncleaved receptor and cleaved N-terminal fragment), all detergents were found to extract Lphn3 from the membrane to various degrees.

The receptor in each of these detergents was then purified on an antiFLAG resin column followed by a size exclusion column. The size exclusion column profiles were evaluated to assess overall purity and stability, with an optimal profile being a monodisperse peak at ~12 mL. All detergents had significant peaks at the void volume,

indicative of a population of aggregated protein. However, most of the samples also had peaks at ~12 mL with varying degrees of uniformity and sharpness among the different detergents.

The peak fraction for each detergent-solubilized receptor was incubated at 4C for one week, when they were re-run on size exclusion column to assess the stability of the receptor in these detergents over time. In addition, the peak fractions were visualized on negative stain EM grids to assess the level of heterogeneity of the samples. From the size exclusion column and negative stain EM results, the best detergent for the receptor was determined.

Negative stain electron microscopy analysis

Since the initial negative stain EM screen of the receptor solubilized and purified in the chosen detergent was successful, several negative stain EM micrographs were collected for this sample in order to analyze the protein architecture of this construct. The individual particles were grouped into 2D class averages based on their relative orientations. The averaged views show the protein with two major lobes: one representing the extracellular region and the other representing the 7TM domain inside of the detergent micelle. There is no evidence of oligomerization of the protein in this particular detergent. In addition, analysis of 2D class averages suggest that the orientation of extracellular domains is not very flexible with respect to the 7TM domain. Rather, it seems that there is rigid orientation between the ECR and the 7TM, which suggests that there is an interaction between these two regions.

Conclusion

The ability to produce large amounts of pure, monodisperse protein is critical for pursuing structural studies of proteins, especially membrane proteins. Because membrane proteins are difficult to purify in large quantities, it is important to optimize the extraction and purification steps in order to increase the chances of successful crystallization and/or cryo-EM trials. In this chapter, we focused on the optimization of extraction and purification of an aGPCR from insect cell membranes, specifically focusing on the appropriate detergents to be used during these steps.

The receptor, expressed in insect cell membranes, was solubilized in a variety of detergents to select the best one for extracting the most protein out of the membrane. We found that most detergents were equally successful at extracting similar amounts of protein. To test how stable the proteins were in each of the detergents, the proteins were purified on antiFLAG resin columns and run on size exclusion columns. The size exclusion column peaks were useful indicators of the heterogeneity and stability of the receptor in each detergent. To further test homogeneity of the samples, the peak fractions were visualized using negative stain EM. Using this technique, one detergent became a less desirable detergent due to the amount of large unidentifiable aggregates present on the grid. Fortunately, in other detergents the sample was more homogenous and lacked aggregates. The peak fractions for each detergent were also incubated for one week at 4C and re-run on a size exclusion column to test long-term protein stability after purification.

Taken together, the detergent screen was an important step in optimizing the solubilization and purification steps for this aGPCR, which will increase the chances of successful structural studies. Based on all of the results, a strategy for producing sufficient amounts of pure sample for crystallographic and cryo-EM studies was obtained.

Since negative stain EM does not require large amounts of proteins, we were able to use this technique to investigate the protein architecture of the receptor. Single particles were analyzed and grouped into 2D class averages, which showed averaged views of the protein from all orientations. These class averages showed that the receptor does not appear to be oligomerized. The data also showed that the protein is not very flexible between the ECR and the 7TM. Instead, the extracellular domains seem to be in a rigid orientation with respect to the 7TM. These results suggest that there may be an interaction between the ECR and the 7TM which contributes to this rigid orientation and that this interaction may be important for regulating receptor function. However, high-resolution structural studies are critical for gaining further insight into the structure and function of this and other aGPCRs. Importantly, the extraction and purification conditions for this receptor have been optimized and will be a helpful contribution towards the goal of obtaining a high-resolution aGPCR 7TM structure.

Chapter 8: Conclusions and Future Direction

Regulation of aGPCR function by its extracellular region

As the aGPCR field of study is relatively new compared to other GPCRs, the mechanism by which aGPCRs signal and are regulated are not well understood. A useful way to study the functions of aGPCRs is to first better understand the structures of these receptors. Obtaining high-resolution structures of aGPCRs, both their ECRs and 7TM domains, provide many avenues for proposing models of activation and regulation. In fact, a major discovery in this field came from the crystal structures of GAIN domains from two aGPCRs (Arac et al., 2012), which led to the development of aGPCR activation models.

Prior to the studies described in this thesis, there were crystal structures of ECR fragments including GAIN domains from various aGPCRs (Arac et al., 2012; Salzman et al., 2016) and other adhesion domains (Jackson et al., 2015; Lu et al., 2015a; Ranaivoson et al., 2015) from Lphns. In 2016, the structure of the first full ECR of an aGPCR was published (Salzman et al., 2016). Although the GPR56 ECR is one of the smaller ECRs in the aGPCR family, the structure led to a lot of new information about its function. For example, the presence of a disulfide bond linking the two ECR domains suggested that rigidity of the ECR plays an important role in its function. In addition, the structure helped identify a conserved patch on the ECR, likely a ligand-binding site. Clearly, there is much value in studying the ECR structures of aGPCRs. Since the aGPCR family is very diverse in their ECRs, the GPR56 structure is likely very different

from the other aGPCRs although there may be commonalities between them such as the importance of ECR domain architecture and rigidity.

Indeed, the full-length ECR of Gpr126 has a very distinct architecture that is important for its function. Although there are no disulfide bonds linking any of the five domains in the ECR of Gpr126, a calcium-binding site in the CUB domain of Gpr126 (-ss isoform) allows it to bind to the HormR domain, forming a compact, closed ECR conformation. An alternative splice variant (+ss isoform) which includes an additional 23 aa was shown, through negative stain EM experiments, to destabilize this closed conformation into a more dynamic, open conformation that exposes the calcium-binding site. This conformational shift is conceptually similar to EGFR, the ECR of which is in a closed, compact inactive conformation, and upon ligand binding, extends its ECR and reveals a hidden functional site that is important for its activation (Kovacs et al., 2015). From G protein signaling assays, we also found that the two alternatively spliced Gpr126 isoforms signal differently, which demonstrates that the ECR of Gpr126 regulates receptor signaling. We also mutated the calcium-binding site in zebrafish and showed that the mutant zebrafish have malformed ears and do not express *mbp*, a critical structural component of the myelin sheath. Because ear and myelination development both require the G protein signaling activity of Gpr126, we showed that the calcium-binding site specifically, despite being far in sequence from the 7TM domain, regulates receptor activity.

The regulation of signaling may be due to ECR conformational changes affecting regulatory ECR-7TM interactions or affecting the accessibility of ligand-binding sites such as the calcium-binding site. Our signaling results support the latter mechanism but

would be strengthened by additional experiments showing that ligands can bind to and activate the open (+ss) isoform better compared to the closed (-ss) isoform.

Reported ligands for Gpr126 are type IV collagen, laminin-211, and the prion protein (Kuffer et al., 2016; Paavola et al., 2014; Petersen et al., 2015). Type IV collagen has been shown to bind somewhere in the CUB-PTX region of Gpr126 (Paavola et al., 2014), where the calcium-binding site is located. Other collagen molecules have been found to bind to calcium-binding sites on the CUB domains of their ligands (Gingras et al., 2011; Venkatraman Girija et al., 2013), which suggests that type IV collagen also binds to the calcium-binding site on Gpr126 in a similar manner. The prion protein, which has a collagen-like region that is sufficient to activate Gpr126, likely also binds to the calcium-binding site similarly (Kuffer et al., 2016). Laminin-211 was reported to bind to GPR126 at a different site, somewhere within the SEA-HormR-GAIN region (Petersen et al., 2015). Laminin-211 activates Gpr126 signaling only when a mechanical force is applied, which suggests that laminin-211 works to pull on the ECR and induce shedding and exposure of the *Stachel* peptide in order to activate signaling. Since mechanical forces were not required to induce activation of Gpr126 by type IV collagen and the prion protein, these ligands may act through a different mechanism from laminin-211. It is possible that binding of type IV collagen and the prion protein can modulate the ECR conformation without inducing shedding in order to regulate signaling. Future signaling assay experiments where ligands are applied to autoproteolysis-mutant Gpr126 constructs would determine whether type IV collagen and the prion protein act through *Stachel*-dependent mechanisms. Nevertheless, it is clear that the ECR plays an important role in receptor signaling and analysis of the

crystal structure has provided great insight into how the ECR might regulate this receptor activity.

From the crystal structure, we also determined that furin protease cleaves Gpr126 within its SEA domain, leading to a fractured but intact domain which likely comes apart upon a mechanical force by a ligand, another possible mechanism by which the ECR regulates receptor activity. Interestingly, there are other aGPCRs (ADGRF1/GPR110 and ADGRF5/GPR116) that have SEA domains that are similarly proteolyzed, and these SEA domains likely have the same function as the Gpr126 SEA domain. At present, we do not know the effects of furin cleavage on Gpr126 function, except that furin cleavage is not a requirement for proper expression and trafficking of Gpr126 to the cell surface. Future signaling assay experiments comparing furin-mutant and wild-type constructs would tell us whether furin cleavage affects receptor signaling. In addition, it is possible that shedding at the furin cleavage site occurs less readily in the (-ss) isoform compared to the (+ss) isoform due to its more compact structure which holds the domains together. To test whether this is true, we could express both isoforms (with N-terminal FLAG tags) on HEK293 cells and probe both the cells and the medium with antiFLAG antibodies. This would tell us whether the (+ss) leads to more disproportionate shedding at the furin cleavage site compared to (-ss).

Lastly, we found in the crystal structure that the angle between the HormR and GAIN domains in Gpr126 is shifted compared to three other HormR+GAIN structures ((Arac et al., 2012), which suggests that conformational changes occur in aGPCR ECRs and may have functional importance. Currently, we are working on crystallizing the ECR of (+ss) Gpr126. This isoform does not crystallize as readily or in the same conditions

as (-ss) Gpr126, which hints that the two isoforms have different shapes and crystal-packing abilities. Hopefully, the crystal structure of (+ss) will show us a conformational change compared to (-ss), perhaps including a shift in the GAIN and HormR domain orientations.

These new findings about the Gpr126 ECR showed that ECR conformation is important for receptor function, consistent with GPR56. Because there are only two full-length aGPCR ECR structures, there is still much to learn and compare with other aGPCRs. Interestingly, SAXS experiments on CD97 have shown that glycosylation has an effect on ECR conformation, and this is correlated with a difference in ECR functionality (Yang et al., 2017). Thus, ECR conformation is likely to be an important feature for more aGPCRs, and we will benefit from continuing to purse structural studies on these receptors.

Alternative splicing of *Gpr126*

Alternative splicing is a powerful way to generate protein diversity and function. Regulation of alternative splicing can be tissue-specific, meaning different ratios of alternative spliced isoforms are expressed in different tissues, or can be a result of cell signaling (Nilsen and Graveley, 2010), such as a response to depolarization of neurons. About a decade ago, many mRNA transcripts encoding aGPCRs were found to be alternatively spliced, producing both functional and nonfunctional receptors (Bjarnadottir et al., 2007). It was unclear how alternative splicing affects function in aGPCRs, although in many cases alternative splicing has been shown to be abundant, particularly in the nervous system (Yeo et al., 2004), and important for biological function (Su et al.,

2018). For example, neurexins are important presynaptic proteins that bind to the postsynaptic neuroligins to form functional synapses. Alternative splicing of neurexins has been shown to mediate synaptic specificity by tuning the affinity of neurexins to neuroligins and other protein partners, such as leucine-rich repeat containing postsynaptic cell adhesion molecule (LRRTM2) and cerebellin1 precursor protein (Cbln1) and dystroglycan (Williams et al., 2010). The ability for neurexins for interact with many different partners through alternative splicing is thought to provide diversity to the many functions of synapses in the nervous system.

Analysis of the structure of the ECR of Gpr126 allowed us to explore the structural and functional consequences of alternative splicing of *Gpr126*. The two isoforms we focused on here, (-ss) and (+ss), differ by the absence/presence of 23 amino acids in the linker region. From the structure of (-ss) ECR, we hypothesized that the additional amino acids would alter the ECR conformation from a closed conformation to a more open and dynamic conformation, which was confirmed by negative stain EM analysis. In addition, the two isoforms also signal differently in our cAMP signaling assay, demonstrating that alternative splicing of *Gpr126* regulates the structure and function of Gpr126. Currently, the expression pattern of the different splice isoforms of Gpr126 is not known. Obtaining this information in the future would greatly expand our understanding of Gpr126 functions in different tissues, such as in Schwann cells, heart, and ear.

In addition to the alternative splicing described above, a different splice isoform of Gpr126 was found to precisely delete the CUB domain (Patra et al., 2013). HEK293 cells expressing a CUB-deleted Gpr126 signaled much lower compared to the wild-type

Gpr126. Since the calcium-binding site, a potential ligand-binding site, is located within the CUB domain, it is possible that Gpr126 function is regulated through the presence of the calcium-binding site. Yet another alternative splicing event leads to different amino acid sequences at the C-terminus of Gpr126 (Moriguchi et al., 2004). Since this is the intracellular region of the protein, our ECR structure does not provide any insight into this isoform. Because this is a region that is exposed to the cytoplasm, it perhaps regulates binding of G proteins to Gpr126, but this remains yet to be explored.

The structure of the ECR has provided a structural framework by which we can start to understand the implications of alternative splicing. Because the ECRs of aGPCRs are very diverse, studying the structures of other aGPCRs would only contribute more to our understanding of how these receptors are regulated by their ECRs.

Extracellular region conformation of Gpr126 and similarities to other receptors

ECR conformation has been shown to play important roles in the functions of several receptors. For example, the ECR EGFR is in a closed, compact inactive conformation until EGF binds to the ECR, leading to a conformational change into a more extended ECR that exposes the dimerization arm of the ECR and promotes EGFR dimerization and activation of signaling (Kovacs et al., 2015). Several drugs have been developed to control EGFR ECR conformation in order to modulate its function. For example, the drug cetuximab blocks ligand binding and prevents extension of the ECR into the active conformation, thus inhibiting receptor activity. In addition, the drug

pertuzumab binds to the dimerization arm of HER2, an EGFR family member, to block dimerization and activation.

A similar mechanism is also found in integrins. In the inactive state, integrin ECRs are compact and bent and form closely associated dimers. Extracellular ligand binding stabilizes extended conformations of the monomers. This leads to movement of the transmembrane and cytoplasmic regions, which is thought to activate intracellular signaling (Campbell and Humphries, 2011).

These mechanisms bear resemblance to what we know so far about the Gpr126 ECR. As described in the previous sections, the Gpr126 ECR conformation can be modulated by alternative splicing. In the (-ss) isoform, the conformation is more compact and adopts a closed conformation in which the N-terminal CUB domain interacts with the more C-terminal HormR domain, similar to the inactive EGFR and integrin structures. In the (+ss) isoform, the additional 23 aa likely destabilizes the interactions present in (-ss) and allows the ECR conformation to be more dynamic and open. This more open conformation presumably exposes a functional calcium-binding site to Gpr126 ligands in the extracellular environment, akin to the exposure of the dimerization arm of EGFR upon conformational change. Although we have observed this conformational change between alternatively spliced isoforms of Gpr126, it is possible that conformational changes occur in other ways. For example, a ligand binding to Gpr126 may induce opening of a closed ECR. Much more work still needs to be done to elucidate the mechanism by which Gpr126 is regulated by its ECR conformation.

Towards drug development against Gpr126 and other aGPCRs

GPCRs are involved in numerous essential biological functions, and because of their importance to human health, many therapeutic drugs have been developed to target and modulate the activities of these receptors. Because the aGPCR family was largely ignored until recently, they lag behind the other GPCR families in terms of drug development. Interestingly, because aGPCRs have large and diverse ECRs, unlike other GPCRs, one strategy for drugging aGPCRs is to target their ECRs instead of their 7TMs. This would provide drug specificity for multiple aGPCRs which may have similar 7TM domains but very different ECRs, thus avoiding harmful side effects. In addition, because drugs that target GPCRs have historically been small molecules and peptides (Sriram and Insel, 2018) that bind to the core of the 7TM region, targeting ECRs would allow for other types of therapeutics such as antibodies or synthetic antibody-like proteins that can target the many binding surfaces provided by ECRs. This has already been shown to be successful for the ECR of GPR56, for which several antibody-like proteins were engineered to activate or inhibit receptor signaling (Salzman et al., 2016; Salzman et al., 2017).

An important step in developing drugs against aGPCRs is to understand their structures. In that way, we can make smarter decisions about how to target their functions. For example, the structure of the ECR of Gpr126 revealed that the ECR conformation is important for regulating receptor activity, and a more open ECR conformation is linked to increased signaling. In addition, the structure identified the calcium-binding site as an important site for both mediating a closed ECR conformation as well as for binding Gpr126 ligands. Therefore, drugs can be engineered to modulate

Gpr126 activity with goals of: 1) stabilizing a closed ECR conformation to inhibit signaling, 2) stabilizing an open ECR conformation to activate signaling, and 3) blocking ligand binding to the calcium-binding site to inhibit activation by endogenous ligands. By targeting Gpr126 activity with these kinds of therapeutics, we can begin to treat the many diseases associated with Gpr126, including several cancers, adolescent idiopathic scoliosis, and arthrogryposis multiplex congenita.

Structures of the transmembrane domains of aGPCRs

Obtaining a high-resolution structure of an aGPCR 7TM domain is still a major goal that has yet to be achieved in the aGPCR field. This is due to many factors including their relatively recent discovery, the lack of stabilizing ligands that would aid in crystallography trials, and the difficulty in purifying these receptors. The work described in this thesis begins to address these factors and hopefully will lead to a structure of an aGPCR 7TM in the future.

Obtaining sufficient amounts of purified protein is an important step for structural studies, both crystallography and cryo-EM. Chapter 7 details the purification optimization trials for an aGPCR. From this work, we are now able to extract and purify this protein to use in future studies. In addition, negative stain EM analysis of the protein showed that the sample is of good quality and can next be screened for cryo-EM studies.

With recent advancements in cryo-EM, we can pursue both crystallization and cryo-EM studies in order to obtain an aGPCR structure. The benefit of cryo-EM is that we do not need to grow crystals of the receptor. However, the protein of interest must

have a large molecular mass (ideally >100 kDa). aGPCR 7TMs by themselves are relatively small, about 35 kDa, but because the ECRs of aGPCRs can be quite large, full-length aGPCRs may be more suitable for cryo-EM studies. In addition, we can form complexes with aGPCRs to make the protein particles even larger. For example, a complex can be formed between full-length Lphn3 (161 kDa) and its extracellular ligand, teneurin (215 kDa, discussed in Chapter 5). Moreover, several recent GPCR structures have been obtained by forming a GPCR-G-protein complex. This not only increases the size of the protein particle, but the G protein also stabilizes the receptor in an active form. In Chapter 6, we determined that Lphn3 couples to and activates G_i, and therefore, we may also pursue a Lphn3-G_i cryo-EM structure.

Of course, even after the first structure of an aGPCR is obtained, there will still be many other structures to pursue. For example, both active and inactive conformations of aGPCR 7TM domains will provide insight into the activation mechanisms of aGPCRs and how they compare to other GPCR families. A structure of an aGPCR bound to its *Stachel* peptide will allow us to understand how *Stachel* peptide binding mediates receptor activation. In addition, structures of aGPCRs with different ECR conformations may show different ECR-7TM interactions that are important for regulating function. There are many aspects of aGPCRs still to be understood, and this thesis helps lay the groundwork for future studies.

Chapter 9: Materials and Methods

Methods used in Chapter 3

Cloning and purification of GPR126 from insect cells

The ECR (residues S38-A816) of zebrafish Gpr126 (-ss), along with a C-terminal 8XHis-tag, was cloned into the pAcGP67a vector. A baculovirus expression system was used to express Gpr126 ECR in High Five cells as previously described (Arac et al., 2012). SeMet-labeled Gpr126 ECR was expressed as previously described (Dong et al., 2009). The insect cell culture was centrifuged for 15 minutes at 900 g at room temperature, and the supernatant containing the secreted proteins was collected. To increase purity of the supernatant, a precipitation method was performed. To each liter of supernatant, 50 mL of 1 M Tris pH 8.0, 1 mL $CaCl_2$, and 1 mL $NiCl_2$ were added and stirred for 30 minutes. The precipitated mixture was centrifuged for 30 minutes at 8000 g at room temperature. The supernatant was collected and incubated with nickel-nitrilotriacetic agarose resin (Qiagen) for 3 hours. The resin was transferred to a glass funnel and washed with 10 mM HEPES 7.2, 150 mM NaCl, 20 mM imidazole pH 7.2, then eluted with 10 mM HEPES 7.2, 150 mM NaCl, 200 mM imidazole pH 7.2. The eluate was concentrated and loaded onto a size exclusion column (Superdex 200 10/300 GL; GE Healthcare), with a final buffer of 10 mM HEPES pH 7.2, 150 mM NaCl.

X-ray crystallography

Purified Gpr126 ECR (both native and SeMet-labeled) was concentrated to 3 mg/mL and screened with several 96-condition sparse matrix crystallography screens. Crystals were obtained in JCSG Core I Suite (Qiagen) position H4: 50mM potassium dihydrogen phosphate, 20% (w/v) PEG 8000. Both native and SeMet-labeled datasets were collected to 2.4 Å at the Advanced Photon Source at Argonne National Laboratory (beamline 23-ID-D). The data sets were processed with HKL2000 and an initial model was determined by SAD phasing using Crank2 in CCP4. Refinement was performed with both REFMAC5 (CCP4) and phenix.refine (PHENIX).

Negative stain electron microscopy

Purified Gpr126 ECR was diluted to ~5 ug/mL and applied to a glow-discharged EM grid (Electron Microscopy Sciences, CF400-Cu) using a conventional negative-stain protocol (Booth et al., 2011). Briefly, 3 uL of protein sample was applied to the grid for 30 seconds, then blotted with filter paper. The grid was put into contact with a 25 uL drop of water and then blotted with filter paper, and this was repeated once more. The grid was then put into contact with a 25 uL drop of 0.75 % uranyl formate solution for 20 seconds, then blotted with filter paper. The grid was air dried and then imaged on a Tecnai G2 F30 operated at 300kV. Image processing was performed with EMAN2 (Tang et al., 2007).

cAMP signaling assay

Full-length, truncated, and mutant Gpr126 constructs were cloned into pCMV5. All constructs include N-terminal FLAG-tags for measuring cell-surface expression levels.

HEK293 cells 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 µg Gpr126 DNA, 0.35 µg GloSensor reporter plasmid (Promega, E2301), and 2.8 µL transfection reagent Fugene 6 (Promega, PRE2693). After a 24-hour incubation, the transfected cells were detached with trypsin-EDTA (Thermo Fisher, 25200056) and seeded (50,000 cells per well) in a white 96-well assay plate. Following another 24-hour incubation, the DMEM was replaced with 100 µL Opti-MEM (Gibco, 31985079) and incubated for 30 minutes. To each well was then added 1 µL GloSensor substrate and 11µL FBS. Basal-level luminescence measurements were taken after 30 minutes to allow for equilibration. The cells were then treated with either 100 uM p7 synthetic peptide (or 1 mM p14) or vehicle DMSO for 15 minutes. Measurements were taken with a Synergy HTX BioTeck plate reader at 25°C.

Flow cytometry to measure cell-surface expression of Gpr126 constructs

HEK293 cells were transfected as previously described and incubated for 24 hours. The cells were then detached with trypsin-EDTA (Thermo Fisher, 25200056) and seeded in a 24-well plate. Following another 24-hour incubation, the cells were detached with citric saline and washed with 1x PBS. The cells were then stained with mouse anti-FLAG primary antibody (1:1000 dilution; Sigma-Aldrich, F3165), washed with 1x PBS + 0.1% BSA twice, stained with donkey anti-mouse Alexa Fluor 488 secondary antibody (1:500 dilution; Invitrogen, A21202), and washed with 1x PBS + 0.1% BSA twice. For data

collection, the cells were resuspended in 1x PBS + 0.1% BSA and 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. The *gpr126*^{stl464} 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 $CaCl_2$, 0.33 mM $MgSO_4$). Larvae were staged at days post fertilization (dpf). *Gpr126*^{stl464} larvae can be identified at 4 dpf by a puffy ear phenotype.

Genotyping

To identify carriers of the *gpr126*^{stl464} allele, the following primers were used to amplify the 381 base pair (bp) locus of interest: F: 5'-GTTGTCGTCAAGACCGGCAC-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 276 and 105 bp products, and the mutant product is 381 bp. BstUI cleaves mutant PCR product into 277 and 104 bp products, and the wild-type product is 381 bp. We recommend using BstUI for genotyping. Any larvae identified with the puffy ear phenotype were always genotyped as *gpr126*^{stl464} 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 vector (Hwang et al., 2013). 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 μM 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 minutes, then cooled to 25° C over a 45 minute time interval. The pDR274 vector was linearized with BsaI and oligonucleotides were ligated into the vector with T4 ligase (NEB) for 10 minutes 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/μl 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 hour, heat inactivated at 80° C for 20 minutes, 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 ~330 ng/μl gRNA, ~371 ng/μl of *Cas9* mRNA (obtained from the Hope Center Transgenic Core at Washington University in St. Louis), and 50 ng/μl of the ssODN. The 150 bp ssODN was ordered from IDT and contained a 5 bp mutation: 5'-atcataaacataccttgcttgtaactgatatggaagcctttcttttggacactCGCCGcggagttaaagaaaacctccat cacatttccagtggagttgag-3'. 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. Whole mount *in situ* hybridization was performed as previously described (Cunningham and Monk, 2018). The previously characterized *mbp* riboprobe was utilized in this experiment (Lyons et al., 2005). Larvae were scored for either presence or absence of *mbp* expression along the PLLn.

Small-angle X-ray scattering

SAXS measurements were performed at the Advanced Photon Source at Argonne National Laboratory 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 was

analyzed using autorg and datgnom using the commands ‘‘autorg –sminrg 0.55 –smaxrg 1.1’’ and ‘‘datgnom ‘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 2.83 (Petoukhov et al., 2012).

Methods used in Chapter 4

Full-length mouse GPR56 was cloned into the pAcGP67A baculovirus transfer vector. A baculovirus expression system was used to express mouse GPR56 in insect cells as previously described (Arac et al., 2012). To confirm expression, insect cell membranes containing mouse GPR56 were harvested and solubilized in buffer containing 30mM HEPES pH 7.5, 500mM NaCl, 10% glycerol, and detergent. Solubilized protein in detergent was run on an SDS-PAGE gel and transferred to a PVDF membrane for western blotting against the receptor. Insect cell membranes containing mouse GPR56 were then used to detect direct G_{13} G protein coupling as previously described (Stoveken et al., 2015).

Methods used in Chapter 5

Various ECR constructs of human TEN2 were cloned into a pAcGP67a vector and expressed in High-Five insect cells using the baculovirus expression system as previously described (Arac¸ et al., 2012). For structural studies, TEN ECR (residues N456-R2648) and TEN ECRΔ1 (residues T727-R2648) were cloned with N-terminal 8XHis-DesG-tags (Skinner et al., 2014). 72 hr after viral infection of High-Five cells, the

medium containing secreted glycosylated proteins was collected and centrifuged at 900 g for 15 min at room temperature.

Purified TEN2 constructs were diluted to ~10 ug/mL and applied to an EM grid (Electron Microscopy Sciences, CF400-Cu,) using a conventional negative-stain protocol (Booth et al., 2011). The sample was imaged on a Tecnai G2 F30 operated at 300kV. Image processing was performed with EMAN2 (Tang et al., 2007).

Methods used in Chapter 6

For G protein coupling experiments, human Lphn was cloned into the pAcGP67A baculovirus transfer vector. A baculovirus expression system was used to express Lphn in insect cells as previously described (Arac et al., 2012). To confirm expression, insect cell membranes containing Lphn were harvested and solubilized in buffer containing 30mM HEPES pH 7.5, 500mM NaCl, 10% glycerol, and detergent. Solubilized protein in detergent was run on an SDS-PAGE gel and transferred to a PVDF membrane for western blotting against the receptor. Insect cell membranes containing Lphn were then used to detect direct G_i G protein coupling as previously described (Stoveken et al., 2015).

For peptide-binding experiments, purified Lphn 7TM (40uM) and biotinylated human Lphn3 peptide (80uM) were incubated for 30 minutes at 4°C. Glutaraldehyde (Sigma-Aldrich #G5882, 0.01%) was added to the sample and incubated for 5 minutes at room temperature. The crosslinking reaction was quenched with 0.5M Tris pH 8.0. The sample was subjected to SDS-PAGE and transferred to PVDF membrane for western blotting. Peptide was detected with NeutrAvidin DyLight 650 (Invitrogen

#84607, 1:2000 dilution) and Lphn was detected with mouse anti-His primary antibody (Qiagen #34660, 1:2000 dilution) and donkey antimouse Alexa Fluor 488 secondary antibody (Invitrogen #A-21202, 1:5000 dilution).

Methods used in Chapter 7

aGPCR constructs for sixteen aGPCRs were cloned and expressed in insect cells. To assess expression levels, insect cell membranes containing aGPCR constructs were harvested and solubilized in buffer containing 30mM HEPES pH 7.5, 500mM NaCl, 10% glycerol, and detergent. Solubilized protein samples were run on an SDS-PAGE gel and transferred to a PVDF membrane for western blotting against the receptors. The best-expressing aGPCR was chosen to continue with purification optimization.

A detergent screen in order to choose an optimal detergent for extraction in terms of amount of protein extracted as well as stability in the specific detergent. Membranes containing the aGPCR were solubilized with various detergents at various concentrations and the solubilized samples were run on an SDS-PAGE gel and transferred to a PVDF membrane for western blotting against the receptors.

Solubilized samples were then purified with antiFLAG resin and run on a size exclusion column. The peak fractions were kept to re-run on size exclusion column after a one-week incubation at 4C to assess protein stability, and peak fraction (diluted to 5 ug/mL) samples wers also applied to a negative stain EM grid (Electron Microscopy Sciences, CF400-Cu) and overlaid with uranyl formate. The EM grids were visualized with a Tecnai T12 electron microscope operated at 120 kV. Images were recorded at a

magnification of ×71,138 and a defocus value of ~1.5 μm on a Gatan US4000 CCD camera.

Based on the negative stain EM images of the protein in each detergent, the protein in the best detergent was chosen to continue analysis. Images were processed in EMAN2 (cite) and 2D class averages were performed using ISAC (Yang et al., 2012).

References

Admani, S., Feldstein, S., Gonzalez, E.M., and Friedlander, S.F. (2014). Beta blockers: an innovation in the treatment of infantile hemangiomas. The Journal of clinical and aesthetic dermatology *7*, 37-45.

Andersen, C.B., Madsen, M., Storm, T., Moestrup, S.K., and Andersen, G.R. (2010). Structural basis for receptor recognition of vitamin-B(12)-intrinsic factor complexes. Nature *464*, 445-448.

Anderson, G.R., Maxeiner, S., Sando, R., Tsetsenis, T., Malenka, R.C., and Sudhof, T.C. (2017). Postsynaptic adhesion GPCR latrophilin-2 mediates target recognition in entorhinal-hippocampal synapse assembly. The Journal of cell biology *216*, 3831-3846.

Aoto, J., Martinelli, D.C., Malenka, R.C., Tabuchi, K., and Südhof, T.C. (2013). Presynaptic neurexin-3 alternative splicing trans-synaptically controls postsynaptic AMPA receptor trafficking. Cell *154*, 75-88.

Arac, D., Boucard, A.A., Bolliger, M.F., Nguyen, J., Soltis, S.M., Sudhof, T.C., and Brunger, A.T. (2012). A novel evolutionarily conserved domain of cell-adhesion GPCRs mediates autoproteolysis. The EMBO journal *31*, 1364-1378.

Arcos-Burgos, M., Jain, M., Acosta, M.T., Shively, S., Stanescu, H., Wallis, D., Domene, S., Velez, J.I., Karkera, J.D., Balog, J.*, et al.* (2010). A common variant of the latrophilin 3 gene, LPHN3, confers susceptibility to ADHD and predicts effectiveness of stimulant medication. Molecular psychiatry *15*, 1053-1066.

Asherson, P., and Gurling, H. (2012). Quantitative and molecular genetics of ADHD. Current topics in behavioral neurosciences *9*, 239-272.

Ashkenazy, H., Abadi, S., Martz, E., Chay, O., Mayrose, I., Pupko, T., and Ben-Tal, N. (2016). ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. Nucleic acids research *44*, W344-350.

Audet, M., and Bouvier, M. (2012). Restructuring G-protein- coupled receptor activation. Cell *151*, 14-23.

Austyn, J.M., and Gordon, S. (1981). F4/80, a monoclonal antibody directed specifically against the mouse macrophage. European journal of immunology *11*, 805-815.

Bae, B.I., Tietjen, I., Atabay, K.D., Evrony, G.D., Johnson, M.B., Asare, E., Wang, P.P., Murayama, A.Y., Im, K., Lisgo, S.N.*, et al.* (2014). Evolutionarily dynamic alternative splicing of GPR56 regulates regional cerebral cortical patterning. Science (New York, NY) *343*, 764-768.

Baker, J.G., and Hill, S.J. (2007). Multiple GPCR conformations and signalling pathways: implications for antagonist affinity estimates. Trends in pharmacological sciences *28*, 374-381.

Balenga, N., Azimzadeh, P., Hogue, J.A., Staats, P.N., Shi, Y., Koh, J., Dressman, H., and Olson, J.A., Jr. (2017). Orphan Adhesion GPCR GPR64/ADGRG2 Is Overexpressed in Parathyroid Tumors and Attenuates Calcium-Sensing Receptor-Mediated Signaling. Journal of bone and mineral research : the official journal of the American Society for Bone and Mineral Research *32*, 654-666.

Barisione, G., Baroffio, M., Crimi, E., and Brusasco, V. (2010). Beta-Adrenergic Agonists. Pharmaceuticals (Basel, Switzerland) *3*, 1016-1044.

Basson, M.A. (2012). Signaling in cell differentiation and morphogenesis. Cold Spring Harbor perspectives in biology *4*.

Baud, V., Chissoe, S.L., Viegas-Pequignot, E., Diriong, S., N'Guyen, V.C., Roe, B.A., and Lipinski, M. (1995). EMR1, an unusual member in the family of hormone receptors with seven transmembrane segments. Genomics *26*, 334-344.

Bayin, N.S., Frenster, J.D., Kane, J.R., Rubenstein, J., Modrek, A.S., Baitalmal, R., Dolgalev, I., Rudzenski, K., Scarabottolo, L., Crespi, D.*, et al.* (2016). GPR133 (ADGRD1), an adhesion G-protein-coupled receptor, is necessary for glioblastoma growth. Oncogenesis *5*, e263.

Bhat, R.R., Yadav, P., Sahay, D., Bhargava, D.K., Creighton, C.J., Yazdanfard, S., Al-Rawi, A., Yadav, V., Qin, L., Nanda, S.*, et al.* (2018). GPCRs profiling and identification of GPR110 as a potential new target in HER2+ breast cancer. Breast cancer research and treatment *170*, 279-292.

Bjarnadottir, T.K., Geirardsdottir, K., Ingemansson, M., Mirza, M.A., Fredriksson, R., and Schioth, H.B. (2007). Identification of novel splice variants of Adhesion G protein-coupled receptors. Gene *387*, 38-48.

Bolliger, M.F., Martinelli, D.C., and Sudhof, T.C. (2011). The cell-adhesion G protein-coupled receptor BAI3 is a high-affinity receptor for C1q-like proteins. Proceedings of the National Academy of Sciences of the United States of America *108*, 2534-2539.

Booe, J.M., Walker, C.S., Barwell, J., Kuteyi, G., Simms, J., Jamaluddin, M.A., Warner, M.L., Bill, R.M., Harris, P.W., Brimble, M.A.*, et al.* (2015). Structural Basis for Receptor Activity-Modifying Protein-Dependent Selective Peptide Recognition by a G Protein-Coupled Receptor. Molecular cell *58*, 1040-1052.

Booth, D.S., Avila-Sakar, A., and Cheng, Y. (2011). Visualizing proteins and macromolecular complexes by negative stain EM: from grid preparation to image acquisition. Journal of visualized experiments : JoVE.

Bortolato, A., Doré, A.S., Hollenstein, K., Tehan, B.G., Mason, J.S., and Marshall, F.H. (2014). Structure of Class B GPCRs: new horizons for drug discovery. British Journal of Pharmacology *171*, 3132-3145.

Boucard, A.A., Ko, J., and Sudhof, T.C. (2012). High affinity neurexin binding to cell adhesion G-protein-coupled receptor CIRL1/latrophilin-1 produces an intercellular adhesion complex. The Journal of biological chemistry *287*, 9399-9413.

Boucard, A.A., Maxeiner, S., and Sudhof, T.C. (2014). Latrophilins function as heterophilic cell-adhesion molecules by binding to teneurins: regulation by alternative splicing. The Journal of biological chemistry *289*, 387-402.

Bouvier, M. (2001). Oligomerization of G-protein coupled transmitter receptors. Nature Reviews Neuroscience *2*.

Buckley, C.D., Rainger, G.E., Bradfield, P.F., Nash, G.B., and Simmons, D.L. (1998). Cell adhesion: more than just glue (review). Molecular membrane biology *15*, 167-176.

Bustanji, Y., and Samori, B. (2002). The mechanical properties of human angiostatin can be modulated by means of its disulfide bonds: A single-molecule force-spectroscopy study. Angewandte Chemie-International Edition *41*, 1546-1548.

Byrne, E.F.X., Sircar, R., Miller, P.S., Hedger, G., Luchetti, G., Nachtergaele, S., Tully, M.D., Mydock-McGrane, L., Covey, D.F., Rambo, R.P.*, et al.* (2016). Structural basis of Smoothened regulation by its extracellular domains. Nature *535*, 517-522.

Cabrera-Vera, T.M., Vanhauwe, J., Thomas, T.O., Medkova, M., Preininger, A., Mazzoni, M.R., and Hamm, H.E. (2003). Insights into G protein structure, function, and regulation. Endocrine reviews *24*, 765-781.

Calderon-Zamora, L., Ruiz-Hernandez, A., Romero-Nava, R., Leon-Sicairos, N., Canizalez-Roman, A., Hong, E., Huang, F., and Villafana, S. (2017). Possible involvement of orphan receptors GPR88 and GPR124 in the development of hypertension in spontaneously hypertensive rat. Clinical and experimental hypertension (New York, NY : 1993) *39*, 513-519.

Calderwood, D.A. (2004). Integrin activation. Journal of cell science *117*, 657.

Campbell, I.D., and Humphries, M.J. (2011). Integrin structure, activation, and interactions. Cold Spring Harbor perspectives in biology *3*.

Carr, J.C., Boese, E.A., Spanheimer, P.M., Dahdaleh, F.S., Martin, M., Calva, D., Schafer, B., Thole, D.M., Braun, T., O'Dorisio, T.M.*, et al.* (2012). Differentiation of small bowel and pancreatic neuroendocrine tumors by gene-expression profiling. Surgery *152*, 998-1007.

Cavallaro, U., and Dejana, E. (2011). Adhesion molecule signalling: not always a sticky business. Nature Reviews Molecular Cell Biology *12*, 189.

Chae, J., Kim, M.J., Goo, J.H., Collier, S., Gubb, D., Charlton, J., Adler, P.N., and Park, W.J. (1999). The Drosophila tissue polarity gene starry night encodes a member of the protocadherin family. Development (Cambridge, England) *126*, 5421-5429.

Chang, G.W., Stacey, M., Kwakkenbos, M.J., Hamann, J., Gordon, S., and Lin, H.H. (2003). Proteolytic cleavage of the EMR2 receptor requires both the extracellular stalk and the GPS motif. FEBS letters *547*, 145-150.

Chiang, N.Y., Chang, G.W., Huang, Y.S., Peng, Y.M., Hsiao, C.C., Kuo, M.L., and Lin, H.H. (2016). Heparin interacts with the adhesion GPCR GPR56, reduces receptor shedding, and promotes cell adhesion and motility. Journal of cell science *129*, 2156-2169.

Chiang, N.Y., Hsiao, C.C., Huang, Y.S., Chen, H.Y., Hsieh, I.J., Chang, G.W., and Lin, H.H. (2011). Disease-associated GPR56 mutations cause bilateral frontoparietal polymicrogyria via multiple mechanisms. The Journal of biological chemistry *286*, 14215-14225.

Chiang, N.Y., Peng, Y.M., Juang, H.H., Chen, T.C., Pan, H.L., Chang, G.W., and Lin, H.H. (2017). GPR56/ADGRG1 Activation Promotes Melanoma Cell Migration via NTF Dissociation and CTF-Mediated Galpha12/13/RhoA Signaling. J Invest Dermatol *137*, 727-736.

Cho, C., Smallwood, P.M., and Nathans, J. (2017). 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.

Cho, H.S., and Leahy, D.J. (2002). Structure of the extracellular region of HER3 reveals an interdomain tether. Science (New York, NY) *297*, 1330-1333.

Christopoulos, A. (2014). Advances in G protein-coupled receptor allostery: from function to structure. Mol Pharmacol *86*, 463-478.

Chun, L., Zhang, W.H., and Liu, J.F. (2012). Structure and ligand recognition of class C GPCRs. Acta pharmacologica Sinica *33*, 312-323.

Cohen-Dvashi, H., Kilimnik, I., and Diskin, R. (2018). Structural basis for receptor recognition by Lujo virus. Nature microbiology *3*, 1153-1160.

Coin, I., Katritch, V., Sun, T., Xiang, Z., Siu, F.Y., Beyermann, M., Stevens, R.C., and Wang, L. (2013). Genetically encoded chemical probes in cells reveal the binding path of urocortin-I to CRF class B GPCR. Cell *155*, 1258-1269.

Cork, S.M., and Van Meir, E.G. (2011). Emerging roles for the BAI1 protein family in the regulation of phagocytosis, synaptogenesis, neurovasculature, and tumor development. Journal of molecular medicine (Berlin, Germany) *89*, 743-752.

Crockett-Torabi, E. (1998). Selectins and mechanisms of signal transduction. Journal of leukocyte biology *63*, 1-14.

Cullen, M., Elzarrad, M.K., Seaman, S., Zudaire, E., Stevens, J., Yang, M.Y., Li, X., Chaudhary, A., Xu, L., Hilton, M.B.*, et al.* (2011). GPR124, an orphan G protein-coupled receptor, is required for CNS-specific vascularization and establishment of the blood-brain barrier. Proceedings of the National Academy of Sciences of the United States of America *108*, 5759-5764.

Cunningham, R.L., and Monk, K.R. (2018). Whole Mount In Situ Hybridization and Immunohistochemistry for Zebrafish Larvae. Methods in molecular biology (Clifton, NJ) *1739*, 371-384.

Curtin, J.A., Quint, E., Tsipouri, V., Arkell, R.M., Cattanach, B., Copp, A.J., Henderson, D.J., Spurr, N., Stanier, P., Fisher, E.M.*, et al.* (2003). Mutation of Celsr1 disrupts planar polarity of inner ear hair cells and causes severe neural tube defects in the mouse. Current biology : CB *13*, 1129-1133.

d'Ydewalle, C., Benoy, V., and Van Den Bosch, L. (2012). Charcot-Marie-Tooth disease: emerging mechanisms and therapies. The international journal of biochemistry & cell biology *44*, 1299-1304.

de Graaf, C., Song, G., Cao, C., Zhao, Q., Wang, M.W., Wu, B., and Stevens, R.C. (2017). Extending the Structural View of Class B GPCRs. Trends in biochemical sciences *42*, 946-960.

De Pascalis, C., and Etienne-Manneville, S. (2017). Single and collective cell migration: the mechanics of adhesions. Molecular Biology of the Cell *28*, 1833-1846.

Deak, F., Liu, X., Khvotchev, M., Li, G., Kavalali, E.T., Sugita, S., and Sudhof, T.C. (2009). Alpha-latrotoxin stimulates a novel pathway of Ca2+-dependent synaptic exocytosis independent of the classical synaptic fusion machinery. The Journal of neuroscience : the official journal of the Society for Neuroscience *29*, 8639-8648.

Demberg, L.M., Rothemund, S., Schoneberg, T., and Liebscher, I. (2015). Identification of the tethered peptide agonist of the adhesion G protein-coupled receptor GPR64/ADGRG2. Biochemical and biophysical research communications *464*, 743-747.

DeRosse, P., Lencz, T., Burdick, K.E., Siris, S.G., Kane, J.M., and Malhotra, A.K. (2008). The genetics of symptom-based phenotypes: toward a molecular classification of schizophrenia. Schizophrenia bulletin *34*, 1047-1053.

Dong, G., Wearsch, P.A., Peaper, D.R., Cresswell, P., and Reinisch, K.M. (2009). Insights into MHC class I peptide loading from the structure of the tapasin-ERp57 thiol oxidoreductase heterodimer. Immunity *30*, 21-32.

Dorsam, R.T., and Gutkind, J.S. (2007). G-protein-coupled receptors and cancer. Nature reviews Cancer *7*, 79-94.

Eubelen, M., Bostaille, N., Cabochette, P., Gauquier, A., Tebabi, P., Dumitru, A.C., Koehler, M., Gut, P., Alsteens, D., Stainier, D.Y.R*., et al.* (2018). A molecular mechanism for Wnt ligand-specific signaling. Science (New York, NY) *361*.

Feng, K., Zhou, X.H., Oohashi, T., Morgelin, M., Lustig, A., Hirakawa, S., Ninomiya, Y., Engel, J., Rauch, U., and Fassler, R. (2002). All four members of the Ten-m/Odz family of transmembrane proteins form dimers. The Journal of biological chemistry *277*, 26128-26135.

Francavilla, C., Loeffler, S., Piccini, D., Kren, A., Christofori, G., and Cavallaro, U. (2007). Neural cell adhesion molecule regulates the cellular response to fibroblast growth factor. Journal of cell science *120*, 4388-4394.

Fredriksson, R., Gloriam, D.E., Hoglund, P.J., Lagerstrom, M.C., and Schioth, H.B. (2003a). There exist at least 30 human G-protein-coupled receptors with long Ser/Thr-rich N-termini. Biochemical and biophysical research communications *301*, 725-734.

Fredriksson, R., Lagerstrom, M.C., Lundin, L.G., and Schioth, H.B. (2003b). The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints. Molecular Pharmacology *63*, 1256-1272.

Freemont, A.J., and Hoyland, J.A. (1996). Cell adhesion molecules. Clinical molecular pathology *49*, M321-M330.

Fuccillo, M.V., Foldy, C., Gokce, O., Rothwell, P.E., Sun, G.L., Malenka, R.C., and Sudhof, T.C. (2015). Single-Cell mRNA Profiling Reveals Cell-Type-Specific Expression of Neurexin Isoforms. Neuron *87*, 326-340.

Fukuzawa, T., Ishida, J., Kato, A., Ichinose, T., Ariestanti, D.M., Takahashi, T., Ito, K., Abe, J., Suzuki, T., Wakana, S*., et al.* (2013). Lung surfactant levels are regulated by Ig-Hepta/GPR116 by monitoring surfactant protein D. PloS one *8*, e69451.

Gaboriaud, C., Gregory-Pauron, L., Teillet, F., Thielens, N.M., Bally, I., and Arlaud, G.J. (2011). Structure and properties of the Ca(2+)-binding CUB domain, a widespread ligand-recognition unit involved in major biological functions. The Biochemical journal *439*, 185-193.

Garinet, S., Pignot, G., Vacher, S., Le Goux, C., Schnitzler, A., Chemlali, W., Nanor, S., Delongchamps, N.B., Zerbib, M., Sibony, M*., et al.* (2018). High prevalence of a hotspot of non-coding somatic mutations in intron 6 of GPR126 in bladder cancer. Molecular cancer research : MCR.

Geng, F.S., Abbas, L., Baxendale, S., Holdsworth, C.J., Swanson, A.G., Slanchev, K., Hammerschmidt, M., Topczewski, J., and Whitfield, T.T. (2013). Semicircular canal morphogenesis in the zebrafish inner ear requires the function of gpr126 (lauscher), an

adhesion class G protein-coupled receptor gene. Development (Cambridge, England) *140*, 4362-4374.

Gibson, N.J. (2011). Cell adhesion molecules in context: CAM function depends on the neighborhood. Cell adhesion & migration *5*, 48-51.

Gingras, A.R., Girija, U.V., Keeble, A.H., Panchal, R., Mitchell, D.A., Moody, P.C., and Wallis, R. (2011). Structural basis of mannan-binding lectin recognition by its associated serine protease MASP-1: implications for complement activation. Structure (London, England : 1993) *19*, 1635-1643.

Gordon, W.R., Vardar-Ulu, D., Histen, G., Sanchez-Irizarry, C., Aster, J.C., and Blacklow, S.C. (2007). Structural basis for autoinhibition of Notch. Nature structural & molecular biology *14*, 295-300.

Gordon, W.R., Vardar-Ulu, D., L'Heureux, S., Ashworth, T., Malecki, M.J., Sanchez-Irizarry, C., McArthur, D.G., Histen, G., Mitchell, J.L., Aster, J.C.*, et al.* (2009). Effects of S1 cleavage on the structure, surface export, and signaling activity of human Notch1 and Notch2. PloS one *4*, e6613.

Grace, C.R., Perrin, M.H., Cantle, J.P., Vale, W.W., Rivier, J.E., and Riek, R. (2007). Common and divergent structural features of a series of corticotropin releasing factor-related peptides. Journal of the American Chemical Society *129*, 16102-16114.

Gray, J.X., Haino, M., Roth, M.J., Maguire, J.E., Jensen, P.N., Yarme, A., Stetler-Stevenson, M.A., Siebenlist, U., and Kelly, K. (1996). CD97 is a processed, seven-transmembrane, heterodimeric receptor associated with inflammation. Journal of immunology (Baltimore, Md : 1950) *157*, 5438-5447.

Gumbiner, B.M. (1996). Cell adhesion: the molecular basis of tissue architecture and morphogenesis. Cell *84*, 345-357.

Gurevich, V.V., and Gurevich, E.V. (2017). Molecular Mechanisms of GPCR Signaling: A Structural Perspective. International journal of molecular sciences *18*, 2519.

Hamann, J., Aust, G., Arac, D., Engel, F.B., Formstone, C., Fredriksson, R., Hall, R.A., Harty, B.L., Kirchhoff, C., Knapp, B.*, et al.* (2015). International Union of Basic and Clinical Pharmacology. XCIV. Adhesion G protein-coupled receptors. Pharmacological reviews *67*, 338-367.

Hamann, J., Eichler, W., Hamann, D., Kerstens, H.M., Poddighe, P.J., Hoovers, J.M., Hartmann, E., Strauss, M., and van Lier, R.A. (1995). Expression cloning and chromosomal mapping of the leukocyte activation antigen CD97, a new seven-span transmembrane molecule of the secretion receptor superfamily with an unusual extracellular domain. Journal of immunology (Baltimore, Md : 1950) *155*, 1942-1950.

Hamann, J., Vogel, B., van Schijndel, G.M., and van Lier, R.A. (1996). The seven-span transmembrane receptor CD97 has a cellular ligand (CD55, DAF). The Journal of experimental medicine *184*, 1185-1189.

Harding, M.M. (2000). The geometry of metal-ligand interactions relevant to proteins. II. Angles at the metal atom, additional weak metal-donor interactions. Acta crystallographica Section D, Biological crystallography *56*, 857-867.

Hoare, S.R. (2005). Mechanisms of peptide and nonpeptide ligand binding to Class B G-protein-coupled receptors. Drug discovery today *10*, 417-427.

Hollenstein, K., Kean, J., Bortolato, A., Cheng, R.K., Dore, A.S., Jazayeri, A., Cooke, R.M., Weir, M., and Marshall, F.H. (2013). Structure of class B GPCR corticotropin-releasing factor receptor 1. Nature *499*, 438-443.

Holm, L., and Laakso, L.M. (2016). Dali server update. Nucleic acids research *44*, W351-355.

Hu, P., and Luo, B.H. (2013). Integrin bi-directional signaling across the plasma membrane. Journal of cellular physiology *228*, 306-312.

Huttenlocher, A., and Horwitz, A.R. (2011). Integrins in cell migration. Cold Spring Harbor perspectives in biology *3*, a005074.

Hwang, W.Y., Fu, Y., Reyon, D., Maeder, M.L., Tsai, S.Q., Sander, J.D., Peterson, R.T., Yeh, J.R., and Joung, J.K. (2013). Efficient genome editing in zebrafish using a CRISPR-Cas system. Nature biotechnology *31*, 227-229.

Ichtchenko, K., Bittner, M.A., Krasnoperov, V., Little, A.R., Chepurny, O., Holz, R.W., and Petrenko, A.G. (1999). A novel ubiquitously expressed alpha-latrotoxin receptor is a member of the CIRL family of G-protein-coupled receptors. The Journal of biological chemistry *274*, 5491-5498.

Iguchi, T., Sakata, K., Yoshizaki, K., Tago, K., Mizuno, N., and Itoh, H. (2008). Orphan G protein-coupled receptor GPR56 regulates neural progenitor cell migration via a G alpha 12/13 and Rho pathway. J Biol Chem *283*, 14469-14478.

Irimia, M., Weatheritt, R.J., Ellis, J.D., Parikshak, N.N., Gonatopoulos-Pournatzis, T., Babor, M., Quesnel-Vallieres, M., Tapial, J., Raj, B., O'Hanlon, D.*, et al.* (2014). A highly conserved program of neuronal microexons is misregulated in autistic brains. Cell *159*, 1511-1523.

Jackson, V.A., del Toro, D., Carrasquero, M., Roversi, P., Harlos, K., Klein, R., and Seiradake, E. (2015). Structural basis of latrophilin-FLRT interaction. Structure (London, England : 1993) *23*, 774-781.

Jackson, V.A., Mehmood, S., Chavent, M., Roversi, P., Carrasquero, M., del Toro, D., Seyit-Bremer, G., Ranaivoson, F.M., Comoletti, D., Sansom, M.S.P.*, et al.* (2016).

Super-complexes of adhesion GPCRs and neural guidance receptors. Nature communications *7*.

Jackson, V.A., Meijer, D.H., Carrasquero, M., van Bezouwen, L.S., Lowe, E.D., Kleanthous, C., Janssen, B.J.C., and Seiradake, E. (2018). Structures of Teneurin adhesion receptors reveal an ancient fold for cell-cell interaction. Nature communications *9*, 1079.

Janeway, C.A., Jr., Bottomly, K., Horowitz, J., Kaye, J., Jones, B., and Tite, J. (1985). Modes of cell:cell communication in the immune system. Journal of immunology (Baltimore, Md : 1950) *135*, 739s-742s.

Kakegawa, W., Mitakidis, N., Miura, E., Abe, M., Matsuda, K., Takeo, Y.H., Kohda, K., Motohashi, J., Takahashi, A., Nagao, S.*, et al.* (2015). Anterograde C1ql1 signaling is required in order to determine and maintain a single-winner climbing fiber in the mouse cerebellum. Neuron *85*, 316-329.

Kan, Z., Jaiswal, B.S., Stinson, J., Janakiraman, V., Bhatt, D., Stern, H.M., Yue, P., Haverty, P.M., Bourgon, R., Zheng, J.*, et al.* (2010). Diverse somatic mutation patterns and pathway alterations in human cancers. Nature *466*, 869-873.

Kang, Y., Kuybeda, O., de Waal, P.W., Mukherjee, S., Van Eps, N., Dutka, P., Zhou, X.E., Bartesaghi, A., Erramilli, S., Morizumi, T.*, et al.* (2018). Cryo-EM structure of human rhodopsin bound to an inhibitory G protein. Nature *558*, 553-558.

Karpus, O.N., Veninga, H., Hoek, R.M., Flierman, D., van Buul, J.D., Vandenakker, C.C., vanBavel, E., Medof, M.E., van Lier, R.A., Reedquist, K.A.*, et al.* (2013). Shear stress-dependent downregulation of the adhesion-G protein-coupled receptor CD97 on circulating leukocytes upon contact with its ligand CD55. Journal of immunology (Baltimore, Md : 1950) *190*, 3740-3748.

Katritch, V., Cherezov, V., and Stevens, R.C. (2013). Structure-Function of the G Protein–Coupled Receptor Superfamily. Annual Review of Pharmacology and Toxicology *53*, 531-556.

Kishore, A., and Hall, R.A. (2017). Disease-associated extracellular loop mutations in the adhesion G protein-coupled receptor G1 (ADGRG1; GPR56) differentially regulate downstream signaling. The Journal of biological chemistry *292*, 9711-9720.

Kishore, A., Purcell, R.H., Nassiri-Toosi, Z., and Hall, R.A. (2016). Stalk-dependent and Stalk-independent Signaling by the Adhesion G Protein-coupled Receptors GPR56 (ADGRG1) and BAI1 (ADGRB1). The Journal of biological chemistry *291*, 3385-3394.

Knall, C., and Johnson, G.L. (1998). G-protein regulatory pathways: Rocketing into the twenty-first century. Journal of cellular biochemistry *72 Suppl 30-31*, 137-146.

Koehl, A., Hu, H., Maeda, S., Zhang, Y., Qu, Q., Paggi, J.M., Latorraca, N.R., Hilger, D., Dawson, R., Matile, H.*, et al.* (2018). Structure of the μ-opioid receptor–Gi protein complex. Nature *558*, 547-552.

Koh, J.T., Kook, H., Kee, H.J., Seo, Y.W., Jeong, B.C., Lee, J.H., Kim, M.Y., Yoon, K.C., Jung, S., and Kim, K.K. (2004). Extracellular fragment of brain-specific angiogenesis inhibitor 1 suppresses endothelial cell proliferation by blocking alphavbeta5 integrin. Experimental cell research *294*, 172-184.

Koide, A., Bailey, C.W., Huang, X., and Koide, S. (1998). The fibronectin type III domain as a scaffold for novel binding proteins. Journal of molecular biology *284*, 1141-1151.

Koirala, S., Jin, Z., Piao, X., and Corfas, G. (2009). GPR56-regulated granule cell adhesion is essential for rostral cerebellar development. The Journal of neuroscience : the official journal of the Society for Neuroscience *29*, 7439-7449.

Kou, I., Takahashi, Y., Johnson, T.A., Takahashi, A., Guo, L., Dai, J., Qiu, X., Sharma, S., Takimoto, A., Ogura, Y.*, et al.* (2013). Genetic variants in GPR126 are associated with adolescent idiopathic scoliosis. Nature genetics *45*, 676-679.

Kovacs, E., Zorn, J.A., Huang, Y., Barros, T., and Kuriyan, J. (2015). A structural perspective on the regulation of the epidermal growth factor receptor. Annual review of biochemistry *84*, 739-764.

Krasnoperov, V.G., Beavis, R., Chepurny, O.G., Little, A.R., Plotnikov, A.N., and Petrenko, A.G. (1996). The calcium-independent receptor of alpha-latrotoxin is not a neurexin. Biochemical and biophysical research communications *227*, 868-875.

Krasnoperov, V.G., Bittner, M.A., Beavis, R., Kuang, Y., Salnikow, K.V., Chepurny, O.G., Little, A.R., Plotnikov, A.N., Wu, D., Holz, R.W.*, et al.* (1997). alpha-Latrotoxin stimulates exocytosis by the interaction with a neuronal G-protein-coupled receptor. Neuron *18*, 925-937.

Krishnan, A. (2012). The Origin of GPCRs: Identification of Mammalian like Rhodopsin, Adhesion, Glutamate and Frizzled GPCRs in Fungi. *7*.

Krishnan, A., Almen, M.S., Fredriksson, R., and Schioth, H.B. (2012). The origin of GPCRs: identification of mammalian like Rhodopsin, Adhesion, Glutamate and Frizzled GPCRs in fungi. PloS one *7*, e29817.

Kruse, A.C., Hu, J., Pan, A.C., Arlow, D.H., Rosenbaum, D.M., Rosemond, E., Green, H.F., Liu, T., Chae, P.S., Dror, R.O.*, et al.* (2012). Structure and dynamics of the M3 muscarinic acetylcholine receptor. Nature *482*, 552-556.

Kuffer, A., Lakkaraju, A.K., Mogha, A., Petersen, S.C., Airich, K., Doucerain, C., Marpakwar, R., Bakirci, P., Senatore, A., Monnard, A.*, et al.* (2016). The prion protein is an agonistic ligand of the G protein-coupled receptor Adgrg6. Nature *536*, 464-468.

Kuhnert, F., Mancuso, M.R., Shamloo, A., Wang, H.T., Choksi, V., Florek, M., Su, H., Fruttiger, M., Young, W.L., Heilshorn, S.C*., et al.* (2010). Essential regulation of CNS angiogenesis by the orphan G protein-coupled receptor GPR124. Science (New York, NY) *330*, 985-989.

Lagerstrom, M.C., Rabe, N., Haitina, T., Kalnina, I., Hellstrom, A.R., Klovins, J., Kullander, K., and Schioth, H.B. (2007). The evolutionary history and tissue mapping of GPR123: specific CNS expression pattern predominantly in thalamic nuclei and regions containing large pyramidal cells. Journal of neurochemistry *100*, 1129-1142.

Lagerstrom, M.C., and Schioth, H.B. (2008). Structural diversity of G protein-coupled receptors and significance for drug discovery. Nat Rev Drug Discov *7*, 339-357.

Langenhan, T., Aust, G., and Hamann, J. (2013). Sticky signaling--adhesion class G protein-coupled receptors take the stage. Science signaling *6*, re3.

Langenhan, T., Piao, X., and Monk, K.R. (2016). Adhesion G protein-coupled receptors in nervous system development and disease. Nature reviews Neuroscience *17*, 550-561.

Langenhan, T., Promel, S., Mestek, L., Esmaeili, B., Waller-Evans, H., Hennig, C., Kohara, Y., Avery, L., Vakonakis, I., Schnabel, R*., et al.* (2009). Latrophilin signaling links anterior-posterior tissue polarity and oriented cell divisions in the C. elegans embryo. Developmental cell *17*, 494-504.

Leamey, C.A., and Sawatari, A. (2014). The teneurins: new players in the generation of visual topography. Seminars in cell & developmental biology *35*, 173-179.

Lee, J.W., Huang, B.X., Kwon, H., Rashid, M.A., Kharebava, G., Desai, A., Patnaik, S., Marugan, J., and Kim, H.Y. (2016). Orphan GPR110 (ADGRF1) targeted by N-docosahexaenoylethanolamine in development of neurons and cognitive function. Nature communications *7*, 13123.

Leemans, J.C., te Velde, A.A., Florquin, S., Bennink, R.J., de Bruin, K., van Lier, R.A., van der Poll, T., and Hamann, J. (2004). The epidermal growth factor-seven transmembrane (EGF-TM7) receptor CD97 is required for neutrophil migration and host defense. Journal of immunology (Baltimore, Md : 1950) *172*, 1125-1131.

Leja, J., Essaghir, A., Essand, M., Wester, K., Oberg, K., Totterman, T.H., Lloyd, R., Vasmatzis, G., Demoulin, J.B., and Giandomenico, V. (2009). Novel markers for enterochromaffin cells and gastrointestinal neuroendocrine carcinomas. Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc *22*, 261-272.

Lelianova, V.G., Davletov, B.A., Sterling, A., Rahman, M.A., Grishin, E.V., Totty, N.F., and Ushkaryov, Y.A. (1997). Alpha-latrotoxin receptor, latrophilin, is a novel member of the secretin family of G protein-coupled receptors. The Journal of biological chemistry *272*, 21504-21508.

Levine, A., Bashan-Ahrend, A., Budai-Hadrian, O., Gartenberg, D., Menasherow, S., and Wides, R. (1994). Odd Oz: a novel Drosophila pair rule gene. Cell *77*, 587-598.

Ley, K., Rivera-Nieves, J., Sandborn, W.J., and Shattil, S. (2016). Integrin-based therapeutics: biological basis, clinical use and new drugs. Nature reviews Drug discovery *15*, 173-183.

Li, J., Shalev-Benami, M., Sando, R., Jiang, X., Kibrom, A., Wang, J., Leon, K., Katanski, C., Nazarko, O., Lu, Y.C.*, et al.* (2018). Structural Basis for Teneurin Function in Circuit-Wiring: A Toxin Motif at the Synapse. Cell *173*, 735-748.e715.

Li, S., Schmitz, K.R., Jeffrey, P.D., Wiltzius, J.J., Kussie, P., and Ferguson, K.M. (2005). Structural basis for inhibition of the epidermal growth factor receptor by cetuximab. Cancer cell *7*, 301-311.

Li, X., Roszko, I., Sepich, D.S., Ni, M., Hamm, H.E., Marlow, F.L., and Solnica-Krezel, L. (2013). Gpr125 modulates Dishevelled distribution and planar cell polarity signaling. Development (Cambridge, England) *140*, 3028-3039.

Liang, Y.-L., Khoshouei, M., Radjainia, M., Zhang, Y., Glukhova, A., Tarrasch, J., Thal, D.M., Furness, S.G.B., Christopoulos, G., Coudrat, T.*, et al.* (2017). Phase-plate cryo-EM structure of a class B GPCR–G-protein complex. Nature *546*, 118.

Liebscher, I., Schon, J., Petersen, S.C., Fischer, L., Auerbach, N., Demberg, L.M., Mogha, A., Coster, M., Simon, K.U., Rothemund, S.*, et al.* (2014). A tethered agonist within the ectodomain activates the adhesion G protein-coupled receptors GPR126 and GPR133. Cell reports *9*, 2018-2026.

Liebscher, I., Schoneberg, T., and Promel, S. (2013). Progress in demystification of adhesion G protein-coupled receptors. Biological chemistry *394*, 937-950.

Lin, H.H., Chang, G.W., Davies, J.Q., Stacey, M., Harris, J., and Gordon, S. (2004). Autocatalytic cleavage of the EMR2 receptor occurs at a conserved G protein-coupled receptor proteolytic site motif. The Journal of biological chemistry *279*, 31823-31832.

Liu, G., Liu, S., Lin, M., Li, X., Chen, W., Zuo, Y., Liu, J., Niu, Y., Zhao, S., Long, B.*, et al.* (2018). Genetic polymorphisms of GPR126 are functionally associated with PUMC classifications of adolescent idiopathic scoliosis in a Northern Han population. Journal of cellular and molecular medicine *22*, 1964-1971.

Lossie, A.C., Nakamura, H., Thomas, S.E., and Justice, M.J. (2005). Mutation of l7Rn3 Shows That Odz4 Is Required for Mouse Gastrulation. Genetics *169*, 285-299.

Lu, S., Liu, S., Wietelmann, A., Kojonazarov, B., Atzberger, A., Tang, C., Schermuly, R.T., Grone, H.J., and Offermanns, S. (2017). Developmental vascular remodeling defects and postnatal kidney failure in mice lacking Gpr116 (Adgrf5) and Eltd1 (Adgrl4). PloS one *12*, e0183166.

Lu, Y.C., Nazarko, O.V., Sando, R., 3rd, Salzman, G.S., Li, N.S., Sudhof, T.C., and Arac, D. (2015a). Structural Basis of Latrophilin-FLRT-UNC5 Interaction in Cell Adhesion. Structure (London, England : 1993) *23*, 1678-1691.

Lu, Y.C., Nazarko, O.V., Sando, R., 3rd, Salzman, G.S., Sudhof, T.C., and Arac, D. (2015b). Structural Basis of Latrophilin-FLRT-UNC5 Interaction in Cell Adhesion. Structure *23*, 1678-1691.

Lum, A.M., Wang, B.B., Beck-Engeser, G.B., Li, L., Channa, N., and Wabl, M. (2010). Orphan receptor GPR110, an oncogene overexpressed in lung and prostate cancer. BMC cancer *10*, 40.

Luo, B.H., Carman, C.V., and Springer, T.A. (2007). Structural basis of integrin regulation and signaling. Annual review of immunology *25*, 619-647.

Luo, R., Jeong, S.J., Yang, A., Wen, M., Saslowsky, D.E., Lencer, W.I., Arac, D., and Piao, X. (2014). Mechanism for adhesion G protein-coupled receptor GPR56-mediated RhoA activation induced by collagen III stimulation. PloS one *9*, e100043.

Luttrell, L.M., and Lefkowitz, R.J. (2002). The role of beta-arrestins in the termination and transduction of G-protein-coupled receptor signals. Journal of cell science *115*, 455-465.

Lyons, D.A., Pogoda, H.M., Voas, M.G., Woods, I.G., Diamond, B., Nix, R., Arana, N., Jacobs, J., and Talbot, W.S. (2005). erbb3 and erbb2 are essential for schwann cell migration and myelination in zebrafish. Current biology : CB *15*, 513-524.

Ma, L., and Pei, G. (2007). Beta-arrestin signaling and regulation of transcription. Journal of cell science *120*, 213-218.

Macao, B., Johansson, D.G., Hansson, G.C., and Hard, T. (2006). Autoproteolysis coupled to protein folding in the SEA domain of the membrane-bound MUC1 mucin. Nature structural & molecular biology *13*, 71-76.

Maiga, A., Lemieux, S., Pabst, C., Lavallee, V.P., Bouvier, M., Sauvageau, G., and Hebert, J. (2016). Transcriptome analysis of G protein-coupled receptors in distinct genetic subgroups of acute myeloid leukemia: identification of potential disease-specific targets. Blood cancer journal *6*, e431.

Manglik, A., and Kruse, A.C. (2017). Structural Basis for G Protein-Coupled Receptor Activation. Biochemistry *56*, 5628-5634.

Masiero, M., Simoes, F.C., Han, H.D., Snell, C., Peterkin, T., Bridges, E., Mangala, L.S., Wu, S.Y., Pradeep, S., Li, D.*, et al.* (2013). A core human primary tumor angiogenesis signature identifies the endothelial orphan receptor ELTD1 as a key regulator of angiogenesis. Cancer cell *24*, 229-241.

Mattes, B., and Scholpp, S. (2018). Emerging role of contact-mediated cell communication in tissue development and diseases. Histochemistry and cell biology *150*, 431-442.

McEver, R.P. (2015). Selectins: initiators of leucocyte adhesion and signalling at the vascular wall. Cardiovascular research *107*, 331-339.

McGee, J., Goodyear, R.J., McMillan, D.R., Stauffer, E.A., Holt, J.R., Locke, K.G., Birch, D.G., Legan, P.K., White, P.C., Walsh, E.J.*, et al.* (2006). The very large G-protein-coupled receptor VLGR1: a component of the ankle link complex required for the normal development of auditory hair bundles. The Journal of neuroscience : the official journal of the Society for Neuroscience *26*, 6543-6553.

McKnight, A.J., and Gordon, S. (1996). EGF-TM7: a novel subfamily of seven-transmembrane-region leukocyte cell-surface molecules. Immunol Today *17*, 283-287.

McMillan, D.R., and White, P.C. (2010). Studies on the very large G protein-coupled receptor: from initial discovery to determining its role in sensorineural deafness in higher animals. Advances in experimental medicine and biology *706*, 76-86.

Mogha, A., Benesh, A.E., Patra, C., Engel, F.B., Schoneberg, T., Liebscher, I., and Monk, K.R. (2013). Gpr126 functions in Schwann cells to control differentiation and myelination via G-protein activation. The Journal of neuroscience : the official journal of the Society for Neuroscience *33*, 17976-17985.

Monk, K.R., Naylor, S.G., Glenn, T.D., Mercurio, S., Perlin, J.R., Dominguez, C., Moens, C.B., and Talbot, W.S. (2009). A G protein-coupled receptor is essential for Schwann cells to initiate myelination. Science (New York, NY) *325*, 1402-1405.

Monk, K.R., Oshima, K., Jors, S., Heller, S., and Talbot, W.S. (2011). Gpr126 is essential for peripheral nerve development and myelination in mammals. Development (Cambridge, England) *138*, 2673-2680.

Moriguchi, T., Haraguchi, K., Ueda, N., Okada, M., Furuya, T., and Akiyama, T. (2004). DREG, a developmentally regulated G protein-coupled receptor containing two conserved proteolytic cleavage sites. Genes to cells : devoted to molecular & cellular mechanisms *9*, 549-560.

Mosca, T.J. (2015). On the Teneurin track: a new synaptic organization molecule emerges. Frontiers in Cellular Neuroscience *9*.

Mosca, T.J., Hong, W., Dani, V.S., Favaloro, V., and Luo, L. (2012). Trans-synaptic Teneurin signalling in neuromuscular synapse organization and target choice. Nature *484*, 237-241.

Muller, A., Winkler, J., Fiedler, F., Sastradihardja, T., Binder, C., Schnabel, R., Kungel, J., Rothemund, S., Hennig, C., Schoneberg, T.*, et al.* (2015). Oriented Cell Division in

the C. elegans Embryo Is Coordinated by G-Protein Signaling Dependent on the Adhesion GPCR LAT-1. PLoS genetics *11*, e1005624.

Myers, K.A., Nasioulas, S., Boys, A., McMahon, J.M., Slater, H., Lockhart, P., Sart, D.D., and Scheffer, I.E. (2018). ADGRV1 is implicated in myoclonic epilepsy. Epilepsia *59*, 381-388.

Nakamura, H., Cook, R.N., and Justice, M.J. (2013). Mouse Tenm4 is required for mesoderm induction. BMC developmental biology *13*, 9.

Nave, K.A., Sereda, M.W., and Ehrenreich, H. (2007). Mechanisms of disease: inherited demyelinating neuropathies--from basic to clinical research. Nature clinical practice Neurology *3*, 453-464.

Nazarko, O., Kibrom, A., Winkler, J., Leon, K., Stoveken, H., Salzman, G., Merdas, K., Lu, Y., Narkhede, P., Tall, G.*, et al.* (2018). A Comprehensive Mutagenesis Screen of the Adhesion GPCR Latrophilin-1/ADGRL1. iScience *3*, 264-278.

Ni, Y.Y., Chen, Y., Lu, S.Y., Sun, B.Y., Wang, F., Wu, X.L., Dang, S.Y., Zhang, G.H., Zhang, H.X., Kuang, Y.*, et al.* (2014). Deletion of Gpr128 results in weight loss and increased intestinal contraction frequency. World journal of gastroenterology *20*, 498-508.

Nik-Zainal, S., Davies, H., Staaf, J., Ramakrishna, M., Glodzik, D., Zou, X., Martincorena, I., Alexandrov, L.B., Martin, S., Wedge, D.C.*, et al.* (2016). Landscape of somatic mutations in 560 breast cancer whole-genome sequences. Nature *534*, 47-54.

Nilsen, T.W., and Graveley, B.R. (2010). Expansion of the eukaryotic proteome by alternative splicing. Nature *463*, 457-463.

Nordstrom, K.J., Lagerstrom, M.C., Waller, L.M., Fredriksson, R., and Schioth, H.B. (2009). The Secretin GPCRs descended from the family of Adhesion GPCRs. Molecular biology and evolution *26*, 71-84.

O'Hayre, M., Vazquez-Prado, J., Kufareva, I., Stawiski, E.W., Handel, T.M., Seshagiri, S., and Gutkind, J.S. (2013). The emerging mutational landscape of G proteins and G-protein-coupled receptors in cancer. Nature reviews Cancer *13*, 412-424.

O'Sullivan, M.L., de Wit, J., Savas, J.N., Comoletti, D., Otto-Hitt, S., Yates, J.R., 3rd, and Ghosh, A. (2012). FLRT proteins are endogenous latrophilin ligands and regulate excitatory synapse development. Neuron *73*, 903-910.

O'Sullivan, M.L., Martini, F., von Daake, S., Comoletti, D., and Ghosh, A. (2014). LPHN3, a presynaptic adhesion-GPCR implicated in ADHD, regulates the strength of neocortical layer 2/3 synaptic input to layer 5. Neural development *9*, 7.

Ohta, S., Sakaguchi, S., Kobayashi, Y., Mizuno, N., Tago, K., and Itoh, H. (2015). Agonistic antibodies reveal the function of GPR56 in human glioma U87-MG cells. Biol Pharm Bull *38*, 594-600.

Oliveira, D.M., Laudanna, C., Migliozzi, S., Zoppoli, P., Santamaria, G., Grillone, K., Elia, L., Mignogna, C., Biamonte, F., Sacco, R.*, et al.* (2018). Identification of different mutational profiles in cancers arising in specific colon segments by next generation sequencing. Oncotarget *9*, 23960-23974.

Oohashi, T., Zhou, X.H., Feng, K., Richter, B., Morgelin, M., Perez, M.T., Su, W.D., Chiquet-Ehrismann, R., Rauch, U., and Fassler, R. (1999). Mouse ten-m/Odz is a new family of dimeric type II transmembrane proteins expressed in many tissues. The Journal of cell biology *145*, 563-577.

Ozer, B., and Sezerman, U. (2017). Analysis of the interplay between methylation and expression reveals its potential role in cancer aetiology. Functional & integrative genomics *17*, 53-68.

Paavola, K.J., and Hall, R.A. (2012). Adhesion G Protein-Coupled Receptors: Signaling, Pharmacology, and Mechanisms of Activation. Molecular Pharmacology *82*, 777-783.

Paavola, K.J., Sidik, H., Zuchero, J.B., Eckart, M., and Talbot, W.S. (2014). Type IV collagen is an activating ligand for the adhesion G protein-coupled receptor GPR126. Science signaling *7*, ra76.

Paavola, K.J., Stephenson, J.R., Ritter, S.L., Alter, S.P., and Hall, R.A. (2011). The N terminus of the adhesion G protein-coupled receptor GPR56 controls receptor signaling activity. The Journal of biological chemistry *286*, 28914-28921.

Patra, C., van Amerongen, M.J., Ghosh, S., Ricciardi, F., Sajjad, A., Novoyatleva, T., Mogha, A., Monk, K.R., Muhlfeld, C., and Engel, F.B. (2013). Organ-specific function of adhesion G protein-coupled receptor GPR126 is domain-dependent. Proceedings of the National Academy of Sciences of the United States of America *110*, 16898-16903.

Peeters, M.C., Mos, I., Lenselink, E.B., Lucchesi, M., AP, I.J., and Schwartz, T.W. (2016). Getting from A to B-exploring the activation motifs of the class B adhesion G protein-coupled receptor subfamily G member 4/GPR112. FASEB journal : official publication of the Federation of American Societies for Experimental Biology *30*, 1836-1848.

Pelaseyed, T., Zäch, M., Petersson Å, C., Svensson, F., Johansson, D.G.A., and Hansson, G.C. (2013). Unfolding dynamics of the mucin SEA domain probed by force spectroscopy suggest that it acts as a cell protective device. The FEBS journal *280*.

Pereira, J.A., Lebrun-Julien, F., and Suter, U. (2012). Molecular mechanisms regulating myelination in the peripheral nervous system. Trends in neurosciences *35*, 123-134.

Petersen, S.C., Luo, R., Liebscher, I., Giera, S., Jeong, S.J., Mogha, A., Ghidinelli, M., Feltri, M.L., Schoneberg, T., Piao, X.*, et al.* (2015). The adhesion GPCR GPR126 has distinct, domain-dependent functions in Schwann cell development mediated by interaction with laminin-211. Neuron *85*, 755-769.

Petoukhov, M.V., Franke, D., Shkumatov, A.V., Tria, G., Kikhney, A.G., Gajda, M., Gorba, C., Mertens, H.D., Konarev, P.V., and Svergun, D.I. (2012). New developments in the ATSAS program package for small-angle scattering data analysis. Journal of applied crystallography *45*, 342-350.

Piao, X., Hill, R.S., Bodell, A., Chang, B.S., Basel-Vanagaite, L., Straussberg, R., Dobyns, W.B., Qasrawi, B., Winter, R.M., Innes, A.M.*, et al.* (2004). G protein-coupled receptor-dependent development of human frontal cortex. Science (New York, NY) *303*, 2033-2036.

Pickering, C., Hagglund, M., Szmydynger-Chodobska, J., Marques, F., Palha, J.A., Waller, L., Chodobski, A., Fredriksson, R., Lagerstrom, M.C., and Schioth, H.B. (2008). The Adhesion GPCR GPR125 is specifically expressed in the choroid plexus and is upregulated following brain injury. BMC neuroscience *9*, 97.

Piraino, S.W., and Furney, S.J. (2017). Identification of coding and non-coding mutational hotspots in cancer genomes. BMC genomics *18*, 17.

Prashar, P., Yadav, P.S., Samarjeet, F., and Bandyopadhyay, A. (2014). Microarray meta-analysis identifies evolutionarily conserved BMP signaling targets in developing long bones. Developmental biology *389*, 192-207.

Promel, S., Frickenhaus, M., Hughes, S., Mestek, L., Staunton, D., Woollard, A., Vakonakis, I., Schoneberg, T., Schnabel, R., Russ, A.P.*, et al.* (2012a). The GPS motif is a molecular switch for bimodal activities of adhesion class G protein-coupled receptors. Cell reports *2*, 321-331.

Promel, S., Langenhan, T., and Arac, D. (2013). Matching structure with function: the GAIN domain of adhesion-GPCR and PKD1-like proteins. Trends in pharmacological sciences *34*, 470-478.

Promel, S., Waller-Evans, H., Dixon, J., Zahn, D., Colledge, W.H., Doran, J., Carlton, M.B., Grosse, J., Schoneberg, T., Russ, A.P.*, et al.* (2012b). Characterization and functional study of a cluster of four highly conserved orphan adhesion-GPCR in mouse. Dev Dyn *241*, 1591-1602.

Qin, X., Xu, L., Xia, C., Zhu, W., Sun, W., Liu, Z., Qiu, Y., and Zhu, Z. (2017). Genetic Variant of GPR126 Gene is Functionally Associated With Adolescent Idiopathic Scoliosis in Chinese Population. Spine *42*, E1098-e1103.

Rana, R., Carroll, C.E., Lee, H.J., Bao, J., Marada, S., Grace, C.R., Guibao, C.D., Ogden, S.K., and Zheng, J.J. (2013). Structural insights into the role of the Smoothened cysteine-rich domain in Hedgehog signalling. Nature communications *4*, 2965.

Ranaivoson, F.M., Liu, Q., Martini, F., Bergami, F., von Daake, S., Li, S., Lee, D., Demeler, B., Hendrickson, W.A., and Comoletti, D. (2015). Structural and mechanistic insights into the latrophilin3 - FLRT3 complex that mediates glutamatergic synapse development. Structure (London, England : 1993) *23*, 1665-1677.

Rasmussen, S.G., Choi, H.J., Fung, J.J., Pardon, E., Casarosa, P., Chae, P.S., Devree, B.T., Rosenbaum, D.M., Thian, F.S., Kobilka, T.S.*, et al.* (2011a). Structure of a nanobody-stabilized active state of the beta(2) adrenoceptor. Nature *469*, 175-180.

Rasmussen, S.G., DeVree, B.T., Zou, Y., Kruse, A.C., Chung, K.Y., Kobilka, T.S., Thian, F.S., Chae, P.S., Pardon, E., Calinski, D.*, et al.* (2011b). Crystal structure of the beta2 adrenergic receptor-Gs protein complex. Nature *477*, 549-555.

Ravenscroft, G., Nolent, F., Rajagopalan, S., Meireles, A.M., Paavola, K.J., Gaillard, D., Alanio, E., Buckland, M., Arbuckle, S., Krivanek, M.*, et al.* (2015). Mutations of GPR126 are responsible for severe arthrogryposis multiplex congenita. American journal of human genetics *96*, 955-961.

Robertson, J.L. (2018). The lipid bilayer membrane and its protein constituents. The Journal of General Physiology *150*, 1472.

Rosenbaum, D.M., Cherezov, V., Hanson, M.A., Rasmussen, S.G., Thian, F.S., Kobilka, T.S., Choi, H.J., Yao, X.J., Weis, W.I., Stevens, R.C.*, et al.* (2007). GPCR engineering yields high-resolution structural insights into beta2-adrenergic receptor function. Science (New York, NY) *318*, 1266-1273.

Safaee, M., Clark, A.J., Ivan, M.E., Oh, M.C., Bloch, O., Sun, M.Z., Oh, T., and Parsa, A.T. (2013). CD97 is a multifunctional leukocyte receptor with distinct roles in human cancers (Review). International journal of oncology *43*, 1343-1350.

Safaee, M., Ivan, M.E., Oh, M.C., Oh, T., Sayegh, E.T., Kaur, G., Sun, M.Z., Bloch, O., and Parsa, A.T. (2014). The role of epidermal growth factor-like module containing mucin-like hormone receptor 2 in human cancers. Oncology reviews *8*, 242.

Salzman, G.S., Ackerman, S.D., Ding, C., Koide, A., Leon, K., Luo, R., Stoveken, H.M., Fernandez, C.G., Tall, G.G., Piao, X.*, et al.* (2016). Structural Basis for Regulation of GPR56/ADGRG1 by Its Alternatively Spliced Extracellular Domains. Neuron *91*, 1292-1304.

Salzman, G.S., Zhang, S., Gupta, A., Koide, A., Koide, S., and Arac, D. (2017). Stachel-independent modulation of GPR56/ADGRG1 signaling by synthetic ligands directed to its extracellular region. Proceedings of the National Academy of Sciences of the United States of America *114*, 10095-10100.

Sandberg, A., Johansson, D.G., Macao, B., and Hard, T. (2008). SEA domain autoproteolysis accelerated by conformational strain: energetic aspects. Journal of molecular biology *377*, 1117-1129.

Scaltriti, M., and Baselga, J. (2006). The Epidermal Growth Factor Receptor Pathway: A Model for Targeted Therapy. Clinical Cancer Research *12*, 5268.

Scheiffele, P. (2003). Cell-cell signaling during synapse formation in the CNS. Annual review of neuroscience *26*, 485-508.

Schlyer, S., and Horuk, R. (2006). I want a new drug: G-protein-coupled receptors in drug development. Drug Discov Today *11*, 481-493.

Scholz, N., Gehring, J., Guan, C., Ljaschenko, D., Fischer, R., Lakshmanan, V., Kittel, R.J., and Langenhan, T. (2015). The Adhesion GPCR Latrophilin/CIRL Shapes Mechanosensation. Cell Rep *11*, 866-874.

Seiradake, E., del Toro, D., Nagel, D., Cop, F., Hartl, R., Ruff, T., Seyit-Bremer, G., Harlos, K., Border, E.C., Acker-Palmer, A.*, et al.* (2014). FLRT structure: balancing repulsion and cell adhesion in cortical and vascular development. Neuron *84*, 370-385.

Sha, F., Gencer, E.B., Georgeon, S., Koide, A., Yasui, N., Koide, S., and Hantschel, O. (2013). Dissection of the BCR-ABL signaling network using highly specific monobody inhibitors to the SHP2 SH2 domains. Proceedings of the National Academy of Sciences of the United States of America *110*, 14924-14929.

Shashidhar, S., Lorente, G., Nagavarapu, U., Nelson, A., Kuo, J., Cummins, J., Nikolich, K., Urfer, R., and Foehr, E.D. (2005). GPR56 is a GPCR that is overexpressed in gliomas and functions in tumor cell adhesion. Oncogene *24*, 1673-1682.

Shi, L., Lehto, S.G., Zhu, D.X., Sun, H., Zhang, J., Smith, B.P., Immke, D.C., Wild, K.D., and Xu, C. (2016). Pharmacologic Characterization of AMG 334, a Potent and Selective Human Monoclonal Antibody against the Calcitonin Gene-Related Peptide Receptor. The Journal of pharmacology and experimental therapeutics *356*, 223-231.

Shima, Y., Kengaku, M., Hirano, T., Takeichi, M., and Uemura, T. (2004). Regulation of dendritic maintenance and growth by a mammalian 7-pass transmembrane cadherin. Dev Cell *7*, 205-216.

Shimaoka, M., and Springer, T.A. (2003). Therapeutic antagonists and conformational regulation of integrin function. Nature Reviews Drug Discovery *2*, 703.

Shin, D., Lin, S.T., Fu, Y.H., and Ptacek, L.J. (2013). Very large G protein-coupled receptor 1 regulates myelin-associated glycoprotein via Galphas/Galphaq-mediated protein kinases A/C. Proceedings of the National Academy of Sciences of the United States of America *110*, 19101-19106.

Silva, J.P., Lelianova, V.G., Ermolyuk, Y.S., Vysokov, N., Hitchen, P.G., Berninghausen, O., Rahman, M.A., Zangrandi, A., Fidalgo, S., Tonevitsky, A.G.*, et al.* (2011). Latrophilin 1 and its endogenous ligand Lasso/teneurin-2 form a high-affinity transsynaptic receptor pair with signaling capabilities. Proc Natl Acad Sci U S A *108*, 12113-12118.

Simundza, J., and Cowin, P. (2013). Adhesion G-Protein Coupled Receptors: Elusive Hybrids Come of Age. Cell communication & adhesion *20*, 213-225.

Siu, F.Y., He, M., de Graaf, C., Han, G.W., Yang, D., Zhang, Z., Zhou, C., Xu, Q., Wacker, D., Joseph, J.S.*, et al.* (2013). Structure of the human glucagon class B G-protein-coupled receptor. Nature *499*, 444-449.

Sriram, K., and Insel, P.A. (2018). G Protein-Coupled Receptors as Targets for Approved Drugs: How Many Targets and How Many Drugs? Molecular Pharmacology *93*, 251-258.

Stacey, M., Chang, G.W., Sanos, S.L., Chittenden, L.R., Stubbs, L., Gordon, S., and Lin, H.H. (2002). EMR4, a novel epidermal growth factor (EGF)-TM7 molecule up-regulated in activated mouse macrophages, binds to a putative cellular ligand on B lymphoma cell line A20. The Journal of biological chemistry *277*, 29283-29293.

Stehlik, C., Kroismayr, R., Dorfleutner, A., Binder, B.R., and Lipp, J. (2004). VIGR--a novel inducible adhesion family G-protein coupled receptor in endothelial cells. FEBS letters *569*, 149-155.

Stockbridge, R.B., Kolmakova-Partensky, L., Shane, T., Koide, A., Koide, S., Miller, C., and Newstead, S. (2015). Crystal structures of a double-barrelled fluoride ion channel. Nature *525*, 548-551.

Stoveken, H.M., Hajduczok, A.G., Xu, L., and Tall, G.G. (2015). Adhesion G protein-coupled receptors are activated by exposure of a cryptic tethered agonist. Proceedings of the National Academy of Sciences of the United States of America *112*, 6194-6199.

Su, C.-H., D, D., and Tarn, W.-Y. (2018). Alternative Splicing in Neurogenesis and Brain Development. Frontiers in molecular biosciences *5*, 12-12.

Sudhof, T.C. (2001). alpha-Latrotoxin and its receptors: neurexins and CIRL/latrophilins. Annual review of neuroscience *24*, 933-962.

Sudhof, T.C. (2008). Neuroligins and neurexins link synaptic function to cognitive disease. Nature *455*, 903-911.

Sugita, S., Ichtchenko, K., Khvotchev, M., and Sudhof, T.C. (1998). alpha-Latrotoxin receptor CIRL/latrophilin 1 (CL1) defines an unusual family of ubiquitous G-protein-linked receptors. G-protein coupling not required for triggering exocytosis. The Journal of biological chemistry *273*, 32715-32724.

Sugita, S., Khvochtev, M., and Sudhof, T.C. (1999). Neurexins are functional alpha-latrotoxin receptors. Neuron *22*, 489-496.

Sun, Z., Guo, S.S., and Fassler, R. (2016). Integrin-mediated mechanotransduction. The Journal of cell biology *215*, 445-456.

Szidonya, L., Cserzo, M., and Hunyady, L. (2008). Dimerization and oligomerization of G-protein-coupled receptors: debated structures with established and emerging functions. The Journal of endocrinology *196*, 435-453.

Tang, G., Peng, L., Baldwin, P.R., Mann, D.S., Jiang, W., Rees, I., and Ludtke, S.J. (2007). EMAN2: an extensible image processing suite for electron microscopy. Journal of structural biology *157*, 38-46.

Tang, X., Jin, R., Qu, G., Wang, X., Li, Z., Yuan, Z., Zhao, C., Siwko, S., Shi, T., Wang, P.*, et al.* (2013). GPR116, an adhesion G-protein-coupled receptor, promotes breast cancer metastasis via the Galphaq-p63RhoGEF-Rho GTPase pathway. Cancer research *73*, 6206-6218.

Terzikhan, N., Sun, F., Verhamme, F.M., Adams, H.H.H., Loth, D., Bracke, K.R., Stricker, B.H.C., Lahousse, L., Dupuis, J., Brusselle, G.G.*, et al.* (2018). Heritability and genome-wide association study of diffusing capacity of the lung. The European respiratory journal *52*.

Thal, D.M., Glukhova, A., Sexton, P.M., and Christopoulos, A. (2018). Structural insights into G-protein-coupled receptor allostery. Nature *559*, 45-53.

Thalhammer, A., Contestabile, A., Ermolyuk, Y.S., Ng, T., Volynski, K.E., Soong, T.W., Goda, Y., and Cingolani, L.A. (2017). Alternative Splicing of P/Q-Type Ca(2+) Channels Shapes Presynaptic Plasticity. Cell reports *20*, 333-343.

Traynor, K. (2018). FDA approves licensing of erenumab-aooe to prevent migraine. American journal of health-system pharmacy : AJHP : official journal of the American Society of Health-System Pharmacists *75*, 929-930.

Tu, Y.K., Duman, J.G., and Tolias, K.F. (2018). The Adhesion-GPCR BAI1 Promotes Excitatory Synaptogenesis by Coordinating Bidirectional Trans-synaptic Signaling. The Journal of neuroscience : the official journal of the Society for Neuroscience *38*, 8388-8406.

Tucker, R.P., and Chiquet-Ehrismann, R. (2006). Teneurins: a conserved family of transmembrane proteins involved in intercellular signaling during development. Developmental biology *290*, 237-245.

Uings, I.J., and Farrow, S.N. (2000). Cell receptors and cell signalling. Molecular pathology : MP *53*, 295-299.

Usui, T., Shima, Y., Shimada, Y., Hirano, S., Burgess, R.W., Schwarz, T.L., Takeichi, M., and Uemura, T. (1999). Flamingo, a seven-pass transmembrane cadherin, regulates planar cell polarity under the control of Frizzled. Cell *98*, 585-595.

Vallon, M., Yuki, K., Nguyen, T.D., Chang, J., Yuan, J., Siepe, D., Miao, Y., Essler, M., Noda, M., Garcia, K.C.*, et al.* (2018). A RECK-WNT7 Receptor-Ligand Interaction

Enables Isoform-Specific Regulation of Wnt Bioavailability. Cell reports *25*, 339-349.e339.

Valtcheva, N., Primorac, A., Jurisic, G., Hollmen, M., and Detmar, M. (2013). The orphan adhesion G protein-coupled receptor GPR97 regulates migration of lymphatic endothelial cells via the small GTPases RhoA and Cdc42. The Journal of biological chemistry *288*, 35736-35748.

van Putten, J.P.M., and Strijbis, K. (2017). Transmembrane Mucins: Signaling Receptors at the Intersection of Inflammation and Cancer. Journal of innate immunity *9*, 281-299.

Venkatakrishnan, A.J., Deupi, X., Lebon, G., Tate, C.G., Schertler, G.F., and Babu, M.M. (2013). Molecular signatures of G-protein-coupled receptors. Nature *494*, 185-194.

Venkatraman Girija, U., Gingras, A.R., Marshall, J.E., Panchal, R., Sheikh, M.A., Harper, J.A., Gal, P., Schwaeble, W.J., Mitchell, D.A., Moody, P.C.*, et al.* (2013). Structural basis of the C1q/C1s interaction and its central role in assembly of the C1 complex of complement activation. Proceedings of the National Academy of Sciences of the United States of America *110*, 13916-13920.

Vysokov, N.V., Silva, J.P., Lelianova, V.G., Ho, C., Djamgoz, M.B., Tonevitsky, A.G., and Ushkaryov, Y.A. (2016). The Mechanism of Regulated Release of Lasso/Teneurin-2. Frontiers in molecular neuroscience *9*, 59.

Waller-Evans, H., Promel, S., Langenhan, T., Dixon, J., Zahn, D., Colledge, W.H., Doran, J., Carlton, M.B., Davies, B., Aparicio, S.A.*, et al.* (2010). The orphan adhesion-GPCR GPR126 is required for embryonic development in the mouse. PloS one *5*, e14047.

Wang, C., Wu, H., Katritch, V., Han, G.W., Huang, X.P., Liu, W., Siu, F.Y., Roth, B.L., Cherezov, V., and Stevens, R.C. (2013a). Structure of the human smoothened receptor bound to an antitumour agent. Nature *497*, 338-343.

Wang, J., Wang, X., Chen, X., Lu, S., Kuang, Y., Fei, J., and Wang, Z. (2018). Gpr97/Adgrg3 ameliorates experimental autoimmune encephalomyelitis by regulating cytokine expression. Acta biochimica et biophysica Sinica *50*, 666-675.

Wang, J.J., Zhang, L.L., Zhang, H.X., Shen, C.L., Lu, S.Y., Kuang, Y., Wan, Y.H., Wang, W.G., Yan, H.M., Dang, S.Y.*, et al.* (2013b). Gpr97 is essential for the follicular versus marginal zone B-lymphocyte fate decision. Cell death & disease *4*, e853.

Wee, P., and Wang, Z. (2017). Epidermal Growth Factor Receptor Cell Proliferation Signaling Pathways. Cancers *9*, 52.

Weis, W.I., and Kobilka, B.K. (2014). The Molecular Basis of G Protein–Coupled Receptor Activation. Annual review of biochemistry *87*, 897-919.

Weston, M.D., Luijendijk, M.W., Humphrey, K.D., Moller, C., and Kimberling, W.J. (2004). Mutations in the VLGR1 gene implicate G-protein signaling in the pathogenesis of Usher syndrome type II. American journal of human genetics *74*, 357-366.

White, J.F., Noinaj, N., Shibata, Y., Love, J., Kloss, B., Xu, F., Gvozdenovic-Jeremic, J., Shah, P., Shiloach, J., Tate, C.G.*, et al.* (2012). Structure of the agonist-bound neurotensin receptor. Nature *490*, 508-513.

White, J.P., Wrann, C.D., Rao, R.R., Nair, S.K., Jedrychowski, M.P., You, J.S., Martinez-Redondo, V., Gygi, S.P., Ruas, J.L., Hornberger, T.A.*, et al.* (2014). G protein-coupled receptor 56 regulates mechanical overload-induced muscle hypertrophy. Proceedings of the National Academy of Sciences of the United States of America *111*, 15756-15761.

Wilde, C., Fischer, L., Lede, V., Kirchberger, J., Rothemund, S., Schoneberg, T., and Liebscher, I. (2016). The constitutive activity of the adhesion GPCR GPR114/ADGRG5 is mediated by its tethered agonist. FASEB journal : official publication of the Federation of American Societies for Experimental Biology *30*, 666-673.

Williams, M.E., de Wit, J., and Ghosh, A. (2010). Molecular mechanisms of synaptic specificity in developing neural circuits. Neuron *68*, 9-18.

Winkler, J., and Promel, S. (2016). The adhesion GPCR latrophilin - a novel signaling cascade in oriented cell division and anterior-posterior polarity. Worm *5*, e1170274.

Woelfle, R., D'Aquila, A.L., Pavlović, T., Husić, M., and Lovejoy, D.A. (2015). Ancient interaction between the teneurin C-terminal associated peptides (TCAP) and latrophilin ligand-receptor coupling: a role in behavior. Frontiers in Neuroscience *9*.

Wojcik, J., Hantschel, O., Grebien, F., Kaupe, I., Bennett, K.L., Barkinge, J., Jones, R.B., Koide, A., Superti-Furga, G., and Koide, S. (2010). A potent and highly specific FN3 monobody inhibitor of the Abl SH2 domain. Nature structural & molecular biology *17*, 519-527.

Wootten, D., Christopoulos, A., and Sexton, P.M. (2013). Emerging paradigms in GPCR allostery: implications for drug discovery. Nat Rev Drug Discov *12*, 630-644.

Wootten, D., Miller, L.J., Koole, C., Christopoulos, A., and Sexton, P.M. (2016). Allostery and Biased Agonism at Class B G Protein-Coupled Receptors. Chemical Reviews.

Wu, M.P., Doyle, J.R., Barry, B., Beauvais, A., Rozkalne, A., Piao, X., Lawlor, M.W., Kopin, A.S., Walsh, C.A., and Gussoni, E. (2013). G-protein coupled receptor 56 promotes myoblast fusion through serum response factor- and nuclear factor of activated T-cell-mediated signalling but is not essential for muscle development in vivo. The FEBS journal *280*, 6097-6113.

Xu, L., Begum, S., Hearn, J.D., and Hynes, R.O. (2006). GPR56, an atypical G protein-coupled receptor, binds tissue transglutaminase, TG2, and inhibits melanoma tumor growth and metastasis. Proceedings of the National Academy of Sciences of the United States of America *103*, 9023-9028.

Yang, L., Chen, G., Mohanty, S., Scott, G., Fazal, F., Rahman, A., Begum, S., Hynes, R.O., and Xu, L. (2011). GPR56 Regulates VEGF production and angiogenesis during melanoma progression. Cancer research *71*, 5558-5568.

Yang, L.Y., Liu, X.F., Yang, Y., Yang, L.L., Liu, K.W., Tang, Y.B., Zhang, M., Tan, M.J., Cheng, S.M., Xu, Y.C*., et al.* (2017). Biochemical features of the adhesion G protein-coupled receptor CD97 related to its auto-proteolysis and HeLa cell attachment activities. Acta pharmacologica Sinica *38*, 56-68.

Yang, M.Y., Hilton, M.B., Seaman, S., Haines, D.C., Nagashima, K., Burks, C.M., Tessarollo, L., Ivanova, P.T., Brown, H.A., Umstead, T.M*., et al.* (2013). Essential regulation of lung surfactant homeostasis by the orphan G protein-coupled receptor GPR116. Cell reports *3*, 1457-1464.

Yang, N.J., and Hinner, M.J. (2015). Getting across the cell membrane: an overview for small molecules, peptides, and proteins. Methods in molecular biology (Clifton, NJ) *1266*, 29-53.

Yeo, G., Holste, D., Kreiman, G., and Burge, C.B. (2004). Variation in alternative splicing across human tissues. Genome biology *5*, R74.

Yin, J., Mobarec, J.C., Kolb, P., and Rosenbaum, D.M. (2014). Crystal structure of the human OX orexin receptor bound to the insomnia drug suvorexant. Nature.

Yona, S., Lin, H.-H., Siu, W.O., Gordon, S., and Stacey, M. (2008a). Adhesion-GPCRs: emerging roles for novel receptors. Trends in biochemical sciences *33*, 491-500.

Yona, S., Lin, H.H., Dri, P., Davies, J.Q., Hayhoe, R.P., Lewis, S.M., Heinsbroek, S.E., Brown, K.A., Perretti, M., Hamann, J*., et al.* (2008b). Ligation of the adhesion-GPCR EMR2 regulates human neutrophil function. FASEB journal : official publication of the Federation of American Societies for Experimental Biology *22*, 741-751.

Young, T.R., and Leamey, C.A. (2009). Teneurins: important regulators of neural circuitry. The international journal of biochemistry & cell biology *41*, 990-993.

Yu, X.J., Yang, M.J., Zhou, B., Wang, G.Z., Huang, Y.C., Wu, L.C., Cheng, X., Wen, Z.S., Huang, J.Y., Zhang, Y.D*., et al.* (2015). Characterization of Somatic Mutations in Air Pollution-Related Lung Cancer. EBioMedicine *2*, 583-590.

Yuzwa, S.A., and Miller, F.D. (2017). Deciphering cell-cell communication in the developing mammalian brain. Neurogenesis *4*, e1286425.

Zhang, C., Srinivasan, Y., Arlow, D.H., Fung, J.J., Palmer, D., Zheng, Y., Green, H.F., Pandey, A., Dror, R.O., Shaw, D.E.*, et al.* (2012). High-resolution crystal structure of human protease-activated receptor 1. Nature *492*, 387-392.

Zhang, D.L., Sun, Y.J., Ma, M.L., Wang, Y.J., Lin, H., Li, R.R., Liang, Z.L., Gao, Y., Yang, Z., He, D.F.*, et al.* (2018a). Gq activity- and beta-arrestin-1 scaffolding-mediated ADGRG2/CFTR coupling are required for male fertility. eLife *7*.

Zhang, H., Qiao, A., Yang, D., Yang, L., Dai, A., de Graaf, C., Reedtz-Runge, S., Dharmarajan, V., Zhang, H., Han, G.W.*, et al.* (2017a). Structure of the full-length glucagon class B G-protein-coupled receptor. Nature *546*, 259-264.

Zhang, L., Jensen, R.T., and Maton, P.N. (1990). Characterization of beta-adrenoreceptors on smooth muscle cells from guinea pig stomach. The American journal of physiology *259*, G436-442.

Zhang, R., and Xie, X. (2012). Tools for GPCR drug discovery. Acta pharmacologica Sinica *33*, 372-384.

Zhang, W., Cui, Q., Qu, W., Ding, X., Jiang, D., and Liu, H. (2018b). TRIM58/cg26157385 methylation is associated with eight prognostic genes in lung squamous cell carcinoma. Oncology reports *40*, 206-216.

Zhang, X., Zhao, F., Wu, Y., Yang, J., Han, G.W., Zhao, S., Ishchenko, A., Ye, L., Lin, X., Ding, K.*, et al.* (2017b). Crystal structure of a multi-domain human smoothened receptor in complex with a super stabilizing ligand. Nature communications *8*, 15383.

Zhao, L.H., Yin, Y., Yang, D., Liu, B., Hou, L., Wang, X., Pal, K., Jiang, Y., Feng, Y., Cai, X.*, et al.* (2016). Differential Requirement of the Extracellular Domain in Activation of Class B G Protein-coupled Receptors. The Journal of biological chemistry *291*, 15119-15130.

Zhou, Y., and Nathans, J. (2014). Gpr124 controls CNS angiogenesis and blood-brain barrier integrity by promoting ligand-specific canonical wnt signaling. Developmental cell *31*, 248-256.