A MOLECULAR TRIAD GOVERNING ADULT STEM CELLS ACTIVATION:
CRYSTALLOGRAPHIC STUDIES OF LGR5, R-SPONDIN 1 AND E3 LIGASE ZNRF3

dedicated to my parents

Peng Weng Chuan

Committee:

Prof. dr. Hans Clevers
Prof. dr. Yvonne Jones
Prof. dr. Albert Heck
Prof. dr. Ineke Braakman
Dr. Madelone Maurice
Dr. Wim de Lau

The research described in this thesis was performed at Crystal and Structural Chemistry, Bijvoet Center for Biomolecular Research, Utrecht University, Utrecht, The Netherlands.

ISBN: 978-90-8891-989-3

Cover and layout design by: Weng Chuan Peng

Printed by: Proefschriftmaken.nl, Uitgeverij BOXPress, 's-Hertogenbosch

A MOLECULAR TRIAD GOVERNING ADULT STEM CELLS ACTIVATION:
CRYSTALLOGRAPHIC STUDIES OF LGR5, R-SPONDIN 1 AND E3 LIGASE ZNRF3

EEN MOLECULAIRE TRIADE VERANTWOORDELIJK VOOR ACTIVERING VAN STAMCELLEN IN VOLWASSENEN: EEN KRISTALLOGRAFISCHE STUDIE VAN LGR5, R-SPONDIN 1 EN DE E3 LIGASE ZNRF3

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht op gezag van de rector magnificus, prof.dr. G.J. van der Zwaan, ingevolge het besluit van het college voor promoties in het openbaar te verdedigen op woensdag 12 november 2014 des middags te 2.30 uur

door

Peng Weng Chuan

geboren op 10 november 1982,
te Teluk Intan, Maleisië

Promotor: Prof. dr. P. Gros

This thesis was accomplished with financial support from Nederlandse organisatie voor Wetenschapelijk Onderzoek (NWO) ECHO award and NWO-Spinoza award.

TABLE OF CONTENTS

CHAPTER 1

GENERAL INTRODUCTION

ABSTRACT

Wnt signaling is critical for maintaining intestinal homeostasis by driving intestinal crypt cell proliferation to replace epithelium cells in the villi that are replaced every 3-5 days. With the discovery of Crypt Base Columnar (CBC) cells as intestinal stem cells, the intestinal compartment became an ideal system for studying the molecular basis of Wnt-driven adult stem cell proliferation. In this chapter, the roles of (i) LGR5, a 7-transmembrane G-protein coupled receptor that marks CBC cells, (ii) R-spondin, a potent Wnt agonist and (iii) transmembrane E3 ligases RNF43/ZNRF3 are discussed. Collectively these proteins regulate Wnt-receptor availability on the cell surface, a crucial factor in controlling the growth rate of these adult stem cells.

Wnt signaling drives adult stem cells proliferation in intestinal crypt

Wnt signaling is involved in almost all cell fate decisions during embryonic development. Since a few decades we realize that it also regulates adult tissue homeostasis [1-3]. The Wnt1 gene (Wnt1) was originally discovered by Roel Nusse and Harold Varmus in the 1982 as Int1 gene, a proto-oncogene activated by mouse mammary tumour virus in breast tumors [4]. Nüsselein-Volhard and Wieschaus had earlier in 1980, identified Wingless (Wg), a segment polarity gene in Drosophila [5]. Later it became evident that Wg is the homologue of Int1 in fly. Since then, we refer to them as Wnt1 [6].

The Wnt family of secreted proteins consists of 19 members in humans. All members are about 350 amino-acid residues long and are lipid-modified at two positions [7]. Wnts bind to receptor complexes consisting of Frizzled (FZD) and low density lipoprotein receptor-related protein (LRP) 5 or -6 [8-13]. All 10 FZD receptors are 7-transmembrane (TM) proteins that contain a cysteine-rich ectodomain (CRD). LRP 5 and -6 are single-pass transmembrane proteins that contain four β-propeller ectodomains. The recently described crystal structure of xenopus Wnt8-mouse FZD8 revealed that the palmitoleic acid (one of the lipid modifications) of Wnt8 is inserted into a conserved hydrophobic groove in the Frizzled CRD ectodomain, which contributes to the high binding affinity of Wnt8 to Frizzled 8 [14].

Binding of Wnt to FZD and LRP5/6 leads to stabilization of β-catenin, a hallmark of Wnt signaling activation (Fig 1) [2,15]. In the absence of Wnt, cytoplasmic β-catenin molecules are bound to two scaffold proteins Axin and Adenomatous Polyposis Coli (APC) and are phosphorylated by Glycogen Synthase Kinase 3β (GSK3β) and Casein Kinase I (CKI). Phosphorylated β-catenin molecules are targeted for ubiquitination and subsequently degraded by proteasomes. AXIN, APC, GSK3β and CKI reside in a multi-protein complex collectively termed 'destruction complex'. Upon Wnt binding, LRP6 recruits AXIN and Frizzled recruits Dishevelled (Dsh), which leads to inactivation of the destruction complex. As a result, β-catenin molecules accumulate in the cytoplasm and translocate to the nucleus where they bind to HMG-box transcription factors T-cell lymphoid enhance factor (TCF/LEF). This leads to activation of the transcription of Wnt target genes [16,17].

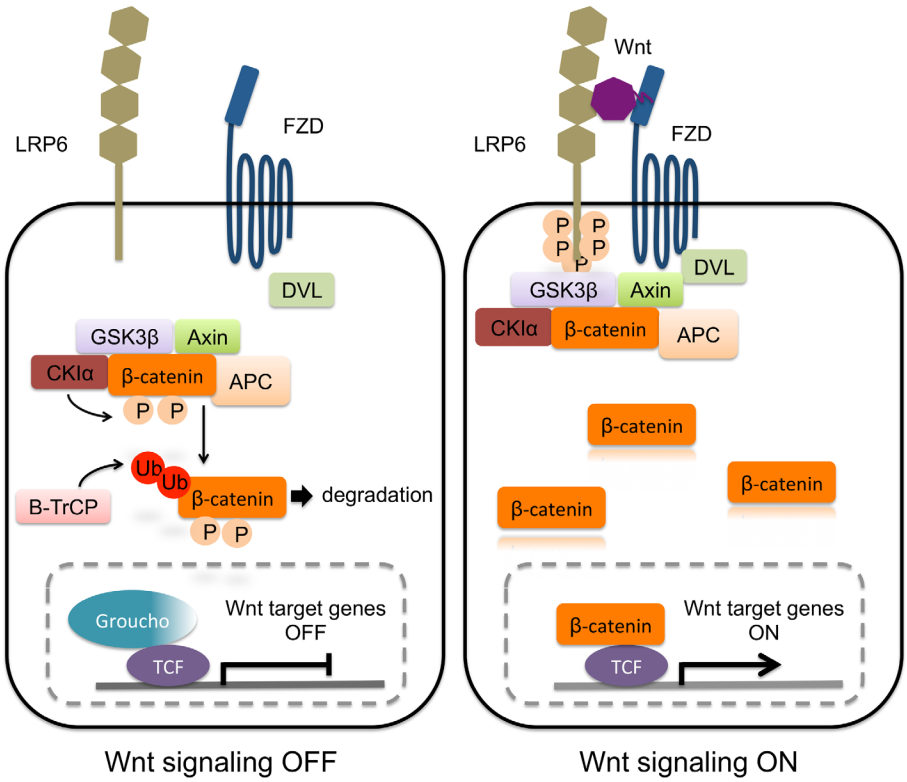


Fig 1. Canonical Wnt signaling pathway
(left) In the absence of Wnt, β-catenin molecules are phosphorylated by 'destruction complex' (composed of GSK3, Axin CKI and APC). Phosphorylated β-catenin molecules are targeted for ubiquitination by B-TrCP and are subsequently degraded. In the absence of β-catenin, Groucho represses Wnt target genes transcription.
(right) In the presence of Wnt, β-catenin molecules escape phosphorylation and degradation. Free β-catenin molecules translocate into the nuclues, bind to transcription factor TCF to activate transcription of Wnt target genes.

Early evidence that Wnt signaling drives renewal of the intestinal epithelium comes from the observation that deletion of TCF4/TCF712 in mice resulted in the absence of the intestinal stem cell compartment [18]. Subsequently, it was shown that overexpression of Wnt inhibitor Dickkopf 1 (Dkk1) [19,20] blocked intestinal crypt proliferation. Similar observations were made with deletion of β-catenin [21,22] or TCF4 [23]. In other tissues, overexpression of Dkk1 eliminates hair follicles and other skin appendage, such as mammary gland [24].

In recent years, the discovery of leucine-rich repeat-containing G-protein coupled receptor (LGR) 5 as adult stem cell marker has helped in uncovering the role of Wnt signaling in Wnt-driven adult stem-cell compartment [25]. LGR5, a Wnt target gene, marks adult stem cells

in intestine and colon, stomach, hair follicle, liver, kidney and mammary gland [26-28]. R-spondins have been identified as a potent Wnt agonist driving proliferation in the intestine. The observation that it is an essential growth factor in intestinal organoid culture is in line with this [29-31]. More recently, transmembrane E3 ligases, Zinc RING Finger (ZNRF) 3 and Ring Finger (RNF) 43, both Wnt target genes, were identified as negative regulators of Wnt signaling [32,33].

The intestinal crypt serves as an excellent system for studying adult stem cells. Much of the molecular basis of Wnt signaling has been investigated for the intestine. Hence, in this chapter, I would like to provide a brief introduction on the architecture of intestinal epithelial compartment, the different stem-cell models, and elaborate on the recent discoveries of LGR5, R-spondin and ZNRF3/RNF43 as key regulators of Wnt signaling in adult stem cells.

Intestinal compartments: Architecture and function of villus and crypt

The small intestine is part of the intestinal tract, which also includes the large intestine or colon, and is responsible for nutrient uptake. The small-intestinal epithelium is characterized by the presence of finger-like protrusions termed villi, which have an extended surface optimized for nutrient uptake [34,35]. The invaginations between villi are called crypts of Lieberkühn (Fig 2). The epithelium is constantly exposed to chemical, biological and physical assaults. Hence, it needs to be constantly renewed. A constant supply of new cells is provided by the intestinal crypts, where intense cell proliferation takes place. Newly formed cells migrate onto the villus, perform their functions for a few days, and enter apoptosis once reaching the tip of the villus. In mouse, the small intestinal epithelium renews itself every 3-5 days, making it one of the most actively renewing tissue in the body [35].

Each villus is surrounded by about 6 to 10 crypts. In the bottom of each crypt, there are about 10 Paneth cells interspersed with about 15 Crypt Base Columnar (CBC) cells [28,36]. These CBC cells comprise the adult stem cells of this system. They divide symmetrically to produce new cells that stochastically adopt a stem cell fate or transit-amplifying (TA) cell fate [37]. The TA compartment forms the remainder of the crypt. TA cells undergo rapid division every 24h [38]. Cells then migrate upward into the villus compartment and are committed to one of the lineages described below [28,34] (Fig 2), i.e.:

- absorptive enterocytes, the most abundant cell type in the villus epithelium, secrete hydrolytic enzyme to breakdown food and absorb the nutrients released.
- goblet cells that secrete mucus for lubrication.
- enteroendocrine cells, which secrete various hormones.
- Paneth cells are long-lived and occupy positions at the base of the crypt in close association with stem cells. They secrete lysozyme and antimicrobial peptides (such as cryptdin and defensins).
- Other rare cells that can reside in both crypt and villus are Cup cells and tuft cells; Peyer's patch associated microfold (M) cells are involved in luminal antigen transportation

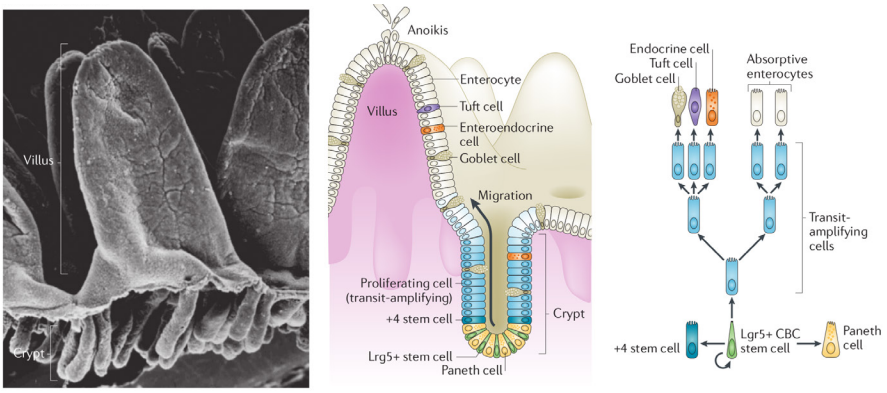


Fig 2. Architechture of the intestinal compartment
(left) Scanning electron micrograph of the intestinal epithelium, with crypt and villus indicated.
(middle) Cartoon representation of crpyt and villus. The crypt is the stem cells 'production factory'. LGR5^{+} Crypt Base Columnar (CBC) stem cells are located at the crypt base, wedged between Paneth cells that serve as the stem-cell niche for CBC cells.
(right) CBC cells divide actively to generate Transit-amplifying (TA) cells. TA cells continue to divide and migrate upwards into the villus and differentiate into various lineages (enterocytes, Tuft cells, enteroendocrine cells, Goblet cells). The location of +4 stem cell proposed by Potten et al. [43] is shown in the middle panel. Reproduced with permission from reference [28].

towards the adaptive immune system.

The identity of the adult stem cells that resides in the crypt had remained elusive for a long time, and two main models are discussed below.

Intestinal crypt: Early stem cell models

Cheng and Leblond were the first to observe CBC cells in the intestinal crypt in 1974 [39] by electron microscopy. Slender, cycling, CBC cells are wedged between Paneth cells and contain a large nucleus with little cytoplasm. Exposure of CBC cells to radioactive ^{3}H thymidine resulted in death of some CBC cells, which were subsequently phagocytosed by neighbouring surviving CBC cells. Initially, radioactive materials were found in the crypt but later in different epithelial lineage cells in the villi, indicating their possible origin from the crypt. Cheng and Leblond proposed 'Unitarian theory of the origin of the four epithelial cell types', suggesting that CBC cells are the intestinal stem cells. Subsequently in 1981, Bjerkness and Cheng proposed the 'stem-cell zone model' [40,41]. In this model, the CBC stem cells reside in 'stem cell permissive environment' in the crypt, at position 1-4 (cell at the bottom of the crypt being position 1). Daughter cells that migrate out of the stem-cell zone enter the 'common origin of differentiation' at position +5. These cells migrate upward into the villus to mature into different lineage of secretory cells, whereas as Paneth cells progenitors move downward to the crypt. Almost two decades later, in 1999, Bjerkness and Cheng used a chemical mutagenesis approach to randomly mark epithelial cells with heritable somatic mutations in a specific

locus. They consistently identified 'marked' CBC cells among these multi-lineage secretory cells, suggesting CBC as the precursor of the differentiated epithelial cells [42].

Alternative stem-cell model: +4 cells model

Potten and colleagues (1977) proposed an alternative stem-cell model termed the '+4 model' [43-45]. The +4 cells are cells located directly above the Paneth cells. They described these cells as label-retaining (LRCs), but yet actively dividing (every 24 hours) and highly radio-sensitive cells. Potten speculated on the discrepancy between 'label-retaining' and 'actively-dividing' characteristics of these cells. Potten proposed that LRCs selectively segregate the newly synthesized DNA from the old DNA template [38,44]. The 'immortal strand hypothesis' had been proposed earlier and suggested that selective segregation of DNA protects stem-cell template DNA from damage due to replication errors [46]. However, this notion has since been found to be inconsistent with recent studies [47-50].

LGR5 marks adult stem cells in intestinal crypt: proof for CBC cells as origin of stem cells

The definitive proof that CBC cells are the adult stem cells comes from in vivo lineage tracing using LGR5 as adult stem cell marker. In the landmark experiment performed in 2007, Barker and colleagues showed that LGR5, a Wnt target gene, is exclusively expressed in the CBC cells in the crypt [25]. The expression of LGR5 in crypt was visualized by LGR5-LacZ and LGR5-EGFP-CreERT2 in a reporter mouse. To perform lineage tracing, mice carrying LGR5-EGFP-ires-Cre-ERT2 were crossbred with conditional reporter line R26R LacZ mice. Low dose of tamoxifen inducible Cre enzyme removed the roadblock for transcription of LacZ. The induction of Cre enzyme introduced LacZ expression as an irreversible genetic marker cell. One day after induction, the lacZ reporter gene activity appeared restricted to a low number of CBC cells. Subsequently, LacZ cells were observed throughout the villus, harboring all epithelial cell types. The lacZ expression persists even after a year of Cre induction. These observations showed that $LGR5^{+}$ CBC cells are the de facto stem cells, which are, multipotent, long-lived and actively dividing.

Subsequently, the same lineage tracing strategy has been used to show that LGR5 marks adult stem cells in colon [25], hair follicles (LGR6) [51], stomach [52], kidney [53], pancreas [54], liver [55] and mammary gland [56].

Paneth cells provide the stem-cell niche

Originally discovered by Schwalbe and Paneth more than a century ago (in 1880s), Paneth cells are long-lived secretory cells that reside at the crypt of Lieberkühn [36]. Paneth cells are involved in innate immunity, protecting us against bacteria attack [57,58]. Hence, Paneth cells contain large secretory granules and secrete lysozyme and antimicrobial peptides, such as defensins and cryptdins. Paneth cells have a lifespan of 3-7 weeks and are the only terminally differentiated cells that escape rapid (weekly) renewal.

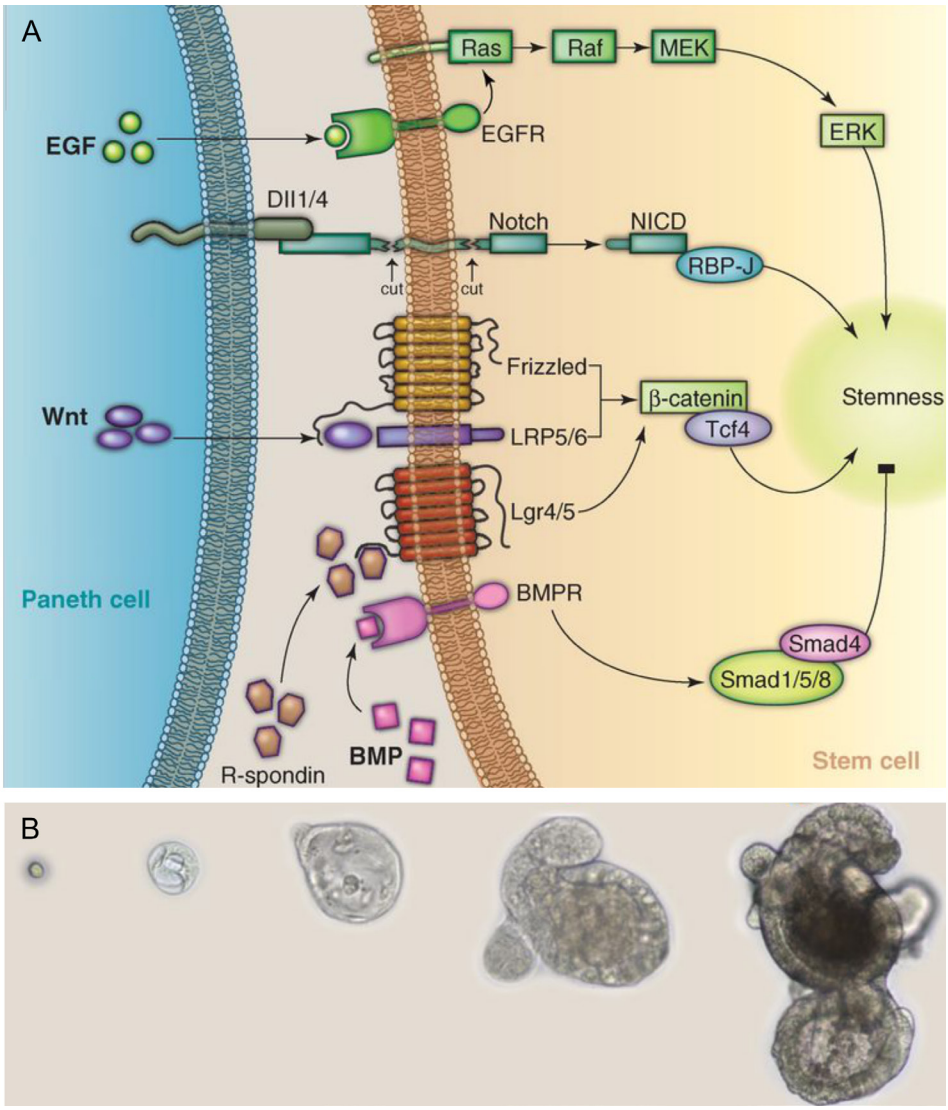


Fig 3. Stem-cell niche and minigut growth
A, Paneth cells provide the stem-cell niche for CBC cells. Activation of Wnt, EGF and Notch signaling pathways and inhibition of BMP signaling pathway are essential for maintaining the stemness of CBC cells.
B, Single LGR5^{+} CBC cell grows into 3D minigut that displays crypt-villus like architecture, shown here in a time-course of two weeks (left to right). Reproduced with permission from reference [60].

However, the most intriguing role of Paneth cells is to provide a stem-cell niche for CBC cells [59,60]. There are about 10 Paneth cells interspersed between 15 CBC cells in each crypt. Loss of Paneth cells resulted in loss of LGR5$^+$ stem cells. Four signaling pathways are involved in maintaining the stemness of CBC cells (Fig 3A), i.e.

I. Wnt for maintaining stem cell fate and drive proliferation. Wnts are secreted by Paneth cells [18,61];
II. Notch signaling prevents stem cells from differentiating into secretory cells. Paneth cells express Dll1$^+$/4$^+$ as surface receptor to activate Notch1/2 receptor on stem cells [62];
III. Paneth cells secrete EGF to activate Ras/Raf/Mek/Erk pathway. Stem cells express EGF receptors [63];
IV. Noggin inhibits BMP signaling in stem cells. Noggin possibly comes from non-epithelial sources [64].

Various studies above have contributed to our understanding of signaling pathways required for maintaining intestinal stem-cell growth. The molecules involved are critical for culturing mini-intestine organoid from single adult stem cells in vitro [31], discussed below.

Single LGR5$^+$ CBC cell generates minigut organoid in vitro

In another landmark experiment, Sato and colleagues (2009) demonstrated that single LGR5$^+$-CBC cells can be grown into self-organizing 3D mini-intestine organoid that displayed crypt-villus architecture [31,60]. Single LGR5$^+$-CBC cells from mice were isolated by FACS and embedded in Matrigel, a laminin- and collagen rich matrix that substitute basal lamina, and supplemented with a cocktail of 'stem-cell factors', i.e. Wnt3a, EGF, Noggin and R-spondin for growth. In less than 2 weeks, the crypt cells spontaneously generated into organoid, in which a crypt-like structure points outward and a villus-like structure pointed inside the lumen (Fig 3B). R-spondin, a Wnt agonist and a growth factor, is known to induce proliferation in Wnt-driven cells. However, the exact mechanism by which R-spondin potentiate Wnt signaling was unknown at that time and will be discussed further below. Another important observation is that LGR5$^+$-CBC-Paneth cells doublets have much higher efficiency to generate organoids compared to LGR5$^+$ CBC cells singlets, further showing the dependence of CBC cells on Paneth cells to provide the stem-cell niche.

The ability of CBC cells to generate organoids in vivo again demonstrates that LGR5$^+$ CBC cells are the multipotent adult stem cells in intestinal crypt. Since then, adult stem cells from other tissues such as stomach, liver and pancreas have also been cultured successfully into organoids in vitro. With the easy availability of organoids for genetic manipulation, such as by infection with lentivirus, Cre-mediated deletion of in vivo floxed genes, or CRISPR/Cas9-mediated DNA editing [65,66], they serve as a promising system for studying human disease.

R-spondin, a stem cell growth factor, potentiates Wnt signaling

The R-spondin family of secreted proteins consists of four members, RSPO1–4, ranging in length from 234-273 amino acids. R-spondin has two cysteine rich furin-like domains (Fu1 and Fu2) at the N-terminus, a thrombospondin type I repeat (TSP-I) and a stretch of polybasic tail at the C-terminus (Fig 4). Kazanskaya and colleague first identified R-spondin as activator of Wnt/β-catenin signaling using the TOPFlash assay by screening cDNAs from a xenopus eye library in HEK 293T cells [29]. R-spondin, in the presence of Wnts, induced stabilization of β-catenin in HEK 293T cells, a hallmark of canonical Wnt signaling. Similar effects were observed with RSPO 1/3/4. They noted that R-spondin alone is unable to stabilize β-catenin. It requires the simultaneous presence of Wnt. The two Fu domains are both required for Wnt signal enhancement, but also sufficient. R-spondins were shown to be essential during various stages of xenopus and mouse embryonic development.

Several important observations were made in this study [29]: RSPO2 functions at extracellular, receptor-ligand level, and its activity is enhanced by presence of Wnts, LRP5/6, and FZD, but not by intracellular components such as Dishevelled (Dsh) or β-catenin. In addition, R-spondin activity is repressed by Dkk1 (LRP6 inhibitor), XDD1 (dominant-negative form of DSH) and GSK-3β. They postulated that R-spondin may function by directly activating the Wnt receptor or by de-activating a Wnt antagonist. Further, no interaction was detected with FZD1-8, LRP5/6 and Dkk1-3, suggesting interaction of R-spondin with a novel receptor.

Using a transgenic mouse model, Kim and colleagues showed that RSPO1 enhanced proliferation of intestinal crypt cells [30], through stabilization of β-catenin in the crypt. RSPO1 expression in knock-in mice led to abdominal distention, characterized by increase in diameter, length and weight of the small intestine. Similar effects were observed in small intestine and colon in mice injected with recombinant RSPO1 protein.

Both studies showed that R-spondin is an important component of the Wnt signaling pathway, which is required during embryonic development in xenopus and mice, and enhanced intestinal proliferation in adult mice. In addition, R-spondin is indispensable for growing mini-intestine organoid in vitro [29-31]. Given the importance of R-spondin, many groups have attempted to search for the receptor that binds R-spondin. It has been proposed that R-spondins bind to FZD [67], LRP6 [68] or Kremen [69] but all these reports were later found to be inaccurate [70].

LGR5 and homologues associate with R-spondin to mediate Wnt signaling

In 2011–2012, four groups independently showed that R-spondins bind to orphan receptors LGR5 and its homologues LGR4 and LGR6 [71-74], using various approaches such as proteomic/mass spectrometry, genome wide siRNA screen and 'rationale candidate ligand' screen. LGR5 is a 7-TM G-protein coupled Receptor (GPCR) with a leucine-rich repeat (LRR) ectodomain (Fig 4).

I. de Lau and colleagues employed a proteomic approach to identify R-spondin receptors [71]. By tandem affinity purification and mass spectrometry analysis, they observed that LGR

4/5 interact with LRP 5/6-FZD Wnt receptor complex [71]. All 4 R-spondins potentiate Wnt signaling, in combination with Wnt3a in TOPFlash assay in HEK 293T cells. Removal of LGR4 by siRNA abrogates signaling mediated by R-spondin but not by Wnt3a, and can be rescued by expression of LGR 4–6. In intestinal organoid, deletion of LGR4 and -5 or withdrawals of R-spondins led to organoid death, which can be rescued by overexpression of Wnt3a or by presence of exogenously provided GSK3 inhibitor. R-spondin 1–4 were co-immunoprecipitated with LGR 4–6. Antibody targeting the N-terminal part of LGR5 effectively blocked R-spondins binding, suggesting that R-spondin may bind directly to the N-terminal part of the LRRs. Surface-plasmon resonance binding studies confirmed a direct physical interaction between RSPO1 and the LGR5 ectodomain with a K_D of 3.1 nM, indicating a high-affinity receptor-ligand binding. Surprisingly, R-spondin-LGR-mediated signaling is not coupled to G-protein activity. Mutation of ERG motif, required for interaction with G-protein, did not affect signaling. de Lau et al. concluded that LGR5 and homologues associate with Wnt receptor complexes and amplified Wnt signaling by binding R-spondins.

II. By candidate ligand approach, Carmon et al. screened for various ligands that bind LGR4–6 and stimulate activity in a GPCR assays, e.g. cAMP production, Ca^{2+} immobilization and b-arrestin translocation [72]. However, they failed to identify any positive hit. R-spondin was then selected as candidate based on strong proliferative effect on intestinal epithelium. By fluorescent microscopy, colocalization of R-spondin and LGR5 in HEK 293T cells was observed and also confirmed by FACS analysis. R-spondins potentiate Wnt signaling in HEK 293T cells overexpressing LGR5, which can be abolished by siRNA targeting LGR 4/5. The binding affinities of RSPO 1–4 to LGR5 were estimated to be in the nanomolar concentration range. They tested extensively the ability of R-spondins to activate GPCR assay signaling but observed that LGR5 signaling is not coupled to hetero-trimetric G proteins or β-arrestin, and observed no cAMP production or Ca^{2+} mobilization.

III. Glinka and colleagues used a genome-wide siRNA screen to identify the R-spondin receptor [74]. RSPO3 and RSPO3-Wnt3a were used to stimulate Wnt signaling in TOPFlash assay in HEK 293T cells. A pool of siRNA's targeting 18,500 genes were screened in HEK 293T cells to determine hits that abrogated activity mediated by RSPO3 but not Wnt1. LGR5 was the only hit identified through this procedure. LGR5-RSPO3 signaling can be inhibited by LRP6 antagonist Dkk1, dominant-negative Casein kinase 1γ, GSK3β and Axin. The interaction between LGR4/5 and RSPO3 was determined at K_D of 2-3 nm.

One interesting observation is that LGR5-RSPO1 complex was quickly internalized in Clathrin-mediated manner, which is required for R-spondin mediated signaling. In contrast, Wnt internalization is caveolin-dependent. In addition, they showed that LGR4/5 is required for planar cell polarity (PCP) signaling in xenopus embryo.

IV. Using a similar approach as Glinka et al., Ruffner and colleagues employed a genome-wide siRNA screen to search for R-spondins receptor [73]. They identified LGR4 as the only hit for RSPO1 in HEK 293T cells, in contrast to LGR5 in the study mentioned above [74]. RSPO1

mediated signaling is abrogated by loss of LGR4, and can be rescued by LGR5/6. They also noted that RSPO1 always colocalize with LGR5 in intracellular vesicles, indicating that LGR4-RSPO1 are internalized together. In addition, R-spondin mediated signaling does not appear to be dependent on G-proteins G_i, G_s or G_q.

All four studies identified LGR 4/5 as high affinity receptor for R-spondin 1–4. A follow-up study also confirmed LGR6 as R-spondins receptor [75]. However, questions remain on how LGR5-R-spondins potentiate Wnt signaling.

Transmembrane E3 ligases ZNRF3 and RNF43 negatively regulate Wnt signaling

In 2012, Koo et al. identified novel transmembrane E3 ligases, RING finger 43 (RNF43) and Zinc RING finger 3 (ZNRF3), which negatively regulate the Wnt receptor complex [32]. RNF43/ZNRF3 are single-pass transmembrane E3 ligase, with a protease-associated (PA) ectodomain and a RING domain on the cytoplasmic side (Fig 4). By gene expression profiling of $LGR5^+$ cells, they observed up-regulation of RNF43 and ZNRF3 genes, amongst other stem-cells specifc genes such as Olfm4, Ascl2 and TROY.

The roles of RNF43/ZNRF3 in Wnt signaling and intestinal crypt proliferation were investigated by using knockout (KO) mice. Homozygous mutants did not show any phenotype. However, when both RNF43/ZNRF3 were removed, intestinal crypts expanded markedly and invaded the villus. The hyper-proliferative compartment displayed high level of β-catenin and was increased in Wnt target genes expression, suggestive of aberrant Wnt signaling in the absence of E3 ligases. Intestinal organoid generated from double KO mice grew faster than wild type, even in the absence of R-spondins. Nonetheless, organoid growth was dependent on the Wnt ligand and can be blocked by the Porcupine (Porc) inhibitor that impaired Wnt secretion. RNF43 expression in intestinal organoids inhibited growth immediately. In addition, RNF43 mutations were identified in two cancer cell lines, RNF43 expression abrogated Wnt3a mediated signaling activity [32].

In HEK 293T cells, RNF43 expression abolished the effect of Wnt3a in Wnt-reporter TOPFlash assay, whereas RING mutants did not show any effect. Overexpression of RNF43 greatly reduced membrane levels of FZD1/3 and LRP5. Besides, FZDs were observed to be rapidly cointernalized with RNF43, suggesting that FZD is most likely a substrate for RNF43. A FZD variant (lacking all internal Lys required for ubiquitination) proved resistant to the RNF43 mediated internalization, indicating a direct role for E3 ligase activity. Besides, lysosomal inhibition with bafilomycin A1 counteracts RNF43 mediated degradation [32].

RNF43/ZNRF3 role as tumour suppressor genes has been documented in various studies. Down-regulation of ZNRF3 in gastric adenocarcinoma correlates with poor tissue differentiation. Overexpression of ZNRF3 suppressed proliferation in gastric cancer cells by down-regulating Wnt signaling [76]. Inactivating mutations of RNF43/ZNRF3 have been found in mucinous cancer in ovary [77], pancreatic cancer cells [78-80], endometrioid carcinoma of uterus [81]

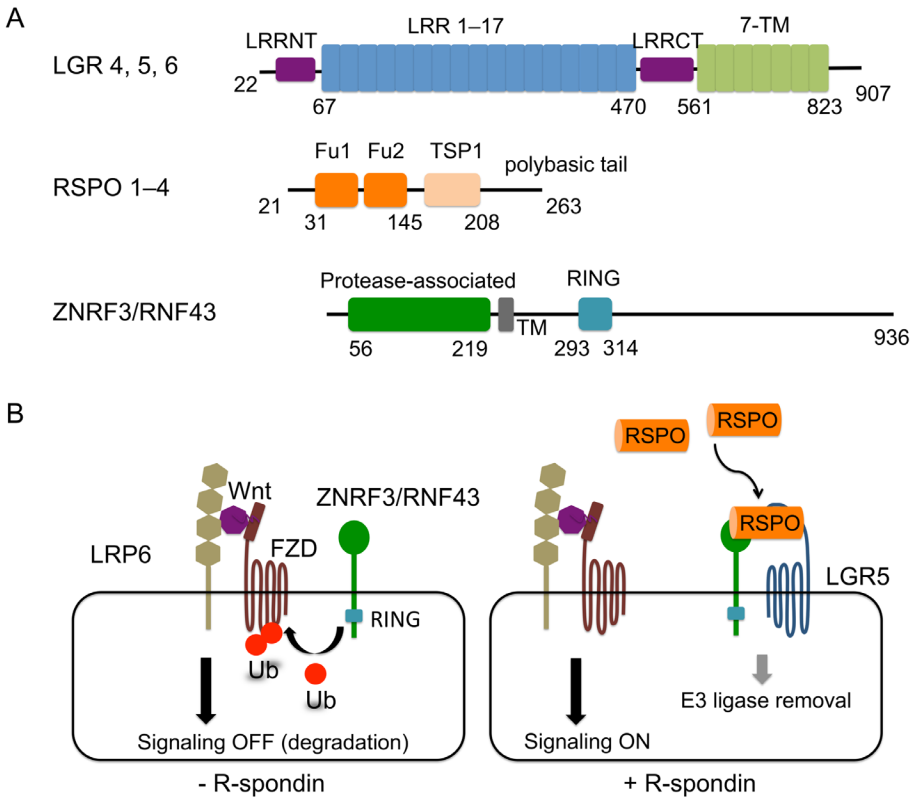


Fig 4. Domain organization of LGR5, R-spondin and E3 ligase ZNRF3/RNF43 and the signaling mechanism
A, Schematic representation of LGR4–6, R-spondin 1–4 and E3 ligases ZNRF3 and RNF43. Residue numbers for each domain are indicated for LGR5, RSPO1 and ZNRF3 (Uniprot sequences). Signal peptides are not shown here.
B, Regulation of Wnt receptor complex by E3 ligases ZNRF3/RNF43. (left) In the absence of R-spondin, LRP6 and FZD receptors are targeted for ubiquitination and degradation by ZNRF3, decreasing Wnt signaling amplitude. (right) R-spondin induces formation of LGR5-RSPO1-ZNRF3 ternary complex and promotes membrane clearance of ZNRF3 from the cell surface, leading to accumulation of Wnt receptor complex and increasing Wnt signaling amplitude.

and liver fluke-associated cholangiocarcinoma [82].

More recently, Moffat et al. reported the identification of E3 ligase, PLR-1, that served as a functional homolog of RNF43/ZNRF3 in C. elegans [83]. PLR-1 controls anterior-posterior neuronal activity in C. elegans. PLR-1 blocks Wnt signaling by reducing surface level of FZD, similar to that of RNF43/ZNRF3. The PA domain of PLR-1 is required for interaction with FZD cysteine-rich domain (CRD); and the RING domain is required for E3 ligase activity. FZD is ubiquitinated at a single lysine present in the second intracellular loop. In addition, PLR-1

regulates surface levels of CAM1/ROR and LIN18/RYK, both Wnt receptors, in a FZD-dependent manner.

R-spondins, the missing link between LGR5 and ZNRF3/RNF43

Hao and colleagues independently identified ZNRF3 by screening for genes that are correlated with Axin2 expression - a Wnt target gene [33]. Based on various genetic and biochemical assays in HEK 293T cells, the group showed that ZNRF3 served as antagonist of Wnt signaling by ubiquitinating LRP6/FZD receptor complex. Physical interaction between FZD and ZNRF3 were detected by immuno-precipitation, confirming FZD as the target for substrate recognition. Besides, antibody targeting of ZNRF3 enhanced surface levels of FZD. In addition, by using KO organisms, they showed that ZNRF3 is critical for embryonic development (dorsal axis patterning and anterior-posterior patterning) in zebrafish, xenopus and mice embryos. Hence, ZNRF3 regulates both Wnt/β-catenin and PCP signaling.

Interestingly, Hao et al. noted that R-spondins increased FZD level on membrane in a manner similar to ZNRF3 inhibition. They observed that RSPO1 bound to ZNRF3 ectodomain also in the absence of LGR4. Overexpression of ZNRF3-ectodomain mutant P103A, blocked RSPO1 mediated FZD stabilization, indicating a direct physical interaction. Further, RSPO1 induced formation of a ternary complex between ZNRF3-RSPO1-LGR4. Strikingly, RSPO1 decreased membrane levels of ZNRF3, which is LGR4 dependent. They demonstrated that induced dimerization of LGR4-ZNRF3 is sufficient to induce membrane clearance of ZNRF3. In addition, the RING domain of ZNRF3 is required for LGR4-ZNRF3 removal from the surface. Hao et al. proposed that R-spondin increases Wnt availability on cell surface by inducing membrane clearance of antagonist ZNRF3, in the presence of LGR4. The proposed role for R-spondin is consistent with previous observations that R-spondin alone does not activate Wnt signaling, but that it shows a synergistic effect with Wnt3a and that LGR5 endocytosis is required for its activity [73]. The proposed mechanism is illustrated in Fig 5.

Summary and Scope of thesis

Wnt signaling is crucial for maintaining intestinal homeostasis [3,18,34]. In the intestinal villi, cells are constantly exposed to biological, chemical and physical assaults. Consequently, these cells undergo rapid apoptosis and are regularly shed into the lumen. Active proliferation takes place in the crypt to supply new cells to the villi. Intestinal compartment is an ideal system for studying adult stem cells, due to the intense proliferation rate and well-studied architecture [34]. Several breakthrough discoveries in recent years, made by studying intestinal crypt, have contributed to our general understanding of adult stem cell proliferation driven by Wnt signaling [28].

LGR5 was initially identified as a Wnt target gene exclusively expressed in CBC cells in the intestinal crypts [25,27,28]. In vivo lineage tracing using LGR5 as marker subsequently proves that CBC cells act as mulitpotent adult stem cells that give rise to all intestinal epithelial cell

types, consistent with the model proposed by Cheng and Leblond [39]. In addition, lineage tracing has helped to pinpoint the identity of adult stem cells in various tissues [26,84], such as in colon [25], hair follicles (LGR6) [51], stomach [52], kidney [53], pancreas [54], liver [55] and mammary gland [56]. LGR5 was later shown to associate with R-spondin to potentiate Wnt signaling [71-74]. R-spondin enhanced intestinal growth [29,30] and is required for mini-intestine culture [31]. How R-spondin binding to LGR5 potentiate Wnt signaling remains unclear, especially since no G-protein activity downstream was detected [71].

Two transmembrane E3 ligases, ZNRF3/RNF43 were recently identified as negative regulators of Wnt signaling [32,33]. ZNRF3 and RNF43 function by ubiquitinating FZD receptors and promoting Wnt receptor complexes degradation. R-spondin was shown to induce membrane clearance of ZNRF3/RNF43 in an LGR4/5-dependent manner [33], which leads to increase of Wnt receptor on the cell surface. LGR5, R-spondin and ZNRF3 interactions at the cell surface determine Wnt receptor availability, which in turn determines the Wnt signaling amplitude, i.e. expression of Wnt target genes required to kickstart adult stem cells proliferation.

The scope of this thesis is to elucidate the molecular interaction between LGR5, R-spondin and E3 ligase ZNRF3 by crystallographic studies. In **chapter 2**, the structural basis of R-spondin recognition by its receptor LGR5, which provide insights into promiscuous binding of R-spondins to LGR4–6, is presented. **Chapter 3** describes the structural data and the mechanism of E3 ligase ZNRF3 inhibition by R-spondin, which offers insights into regulation of Wnt signaling on stem-cells surface. In **chapter 4**, the technical details of expression, crystallization and structure determination of LGR5 and R-spondin is discussed. Production of large quantity (milligrams) of R-spondin and LGR5 proteins is critical for crystallographic studies, biochemical studies and antibody generation. In **chapter 5**, recent developments with regards to signaling roles of LGR5, identification of novel interacting partners and insights obtained from other structural studies, are discussed.

REFERENCES

1. Logan CY, Nusse R (2004) The Wnt signaling pathway in development and disease. Annu Rev Cell Dev Biol 20: 781–810. doi:10.1146/annurev.cellbio.20.010403.113126.
2. Macdonald BT, Tamai K, He X (2009) Wnt/beta-catenin signaling: components, mechanisms, and diseases. Dev Cell 17: 9–26. doi:10.1016/j.devcel.2009.06.016.
3. Clevers H, Nusse R (2012) Wnt/β-catenin signaling and disease. Cell 149: 1192–1205. doi:10.1016/j.cell.2012.05.012.
4. Nusse R, Varmus HE (1982) Many tumors induced by the mouse mammary tumor virus contain a provirus integrated in the same region of the host genome. Cell 31: 99–109.
5. Nüsslein-Volhard C, Wieschaus E (1980) Mutations affecting segment number and polarity in Drosophila. Nature 287: 795–801.
6. Nusse R, Brown A, Papkoff J, Scambler P, Shackleford G, et al. (1991) A new nomenclature for int-1 and related genes: the Wnt gene family. Cell 64: 231.
7. Willert K, Brown JD, Danenberg E, Duncan AW, Weissman IL, et al. (2003) Wnt proteins are

lipid-modified and can act as stem cell growth factors. Nature 423: 448–452. doi:10.1038/nature01611.

8. Bhanot P, Brink M, Samos CH, Hsieh JC, Wang Y, et al. (1996) A new member of the frizzled family from Drosophila functions as a Wingless receptor. Nature 382: 225–230. doi:10.1038/382225a0.
9. Sawa H, Lobel L, Horvitz HR (1996) The Caenorhabditis elegans gene lin-17, which is required for certain asymmetric cell divisions, encodes a putative seven-transmembrane protein similar to the Drosophila frizzled protein. Genes Dev 10: 2189–2197.
10. Yang-Snyder J, Miller JR, Brown JD, Lai CJ, Moon RT (1996) A frizzled homolog functions in a vertebrate Wnt signaling pathway. Curr Biol 6: 1302–1306.
11. Tamai K, Semenov M, Kato Y, Spokony R, Liu C, et al. (2000) LDL-receptor-related proteins in Wnt signal transduction. Nature 407: 530–535. doi:10.1038/35035117.
12. Wehrli M, Dougan ST, Caldwell K, O'Keefe L, Schwartz S, et al. (2000) arrow encodes an LDL-receptor-related protein essential for Wingless signalling. Nature 407: 527–530. doi:10.1038/35035110.
13. Pinson KI, Brennan J, Monkley S, Avery BJ, Skarnes WC (2000) An LDL-receptor-related protein mediates Wnt signalling in mice. Nature 407: 535–538. doi:10.1038/35035124.
14. Janda CY, Waghray D, Levin AM, Thomas C, Garcia KC (2012) Structural basis of Wnt recognition by Frizzled. Science 337: 59–64. doi:10.1126/science.1222879.
15. MacDonald BT, He X (2012) Frizzled and LRP5/6 Receptors for Wnt/ -Catenin Signaling. Cold Spring Harb Perspect Biol 4: a007880–a007880. doi:10.1101/cshperspect.a007880.
16. Behrens J, Kries von JP, Kühl M, Bruhn L, Wedlich D, et al. (1996) Functional interaction of beta-catenin with the transcription factor LEF-1. Nature 382: 638–642. doi:10.1038/382638a0.
17. Molenaar M, van de Wetering M, Oosterwegel M, Peterson-Maduro J, Godsave S, et al. (1996) XTcf-3 transcription factor mediates beta-catenin-induced axis formation in Xenopus embryos. Cell 86: 391–399.
18. Korinek V, Barker N, Moerer P, van Donselaar E, Huls G, et al. (1998) Depletion of epithelial stem-cell compartments in the small intestine of mice lacking Tcf-4. Nat Genet 19: 379–383. doi:10.1038/1270.
19. Pinto D, Gregorieff A, Begthel H, Clevers H (2003) Canonical Wnt signals are essential for homeostasis of the intestinal epithelium. Genes Dev 17: 1709–1713. doi:10.1101/gad.267103.
20. Kuhnert F, Davis CR, Wang H-T, Chu P, Lee M, et al. (2004) Essential requirement for Wnt signaling in proliferation of adult small intestine and colon revealed by adenoviral expression of Dickkopf-1. Proc Natl Acad Sci USA 101: 266–271. doi:10.1073/pnas.2536800100.
21. Ireland H, Kemp R, Houghton C, Howard L, Clarke AR, et al. (2004) Inducible Cre-mediated control of gene expression in the murine gastrointestinal tract: effect of loss of beta-catenin. Gastroenterology 126: 1236–1246.
22. Fevr T, Robine S, Louvard D, Huelsken J (2007) Wnt/beta-catenin is essential for intestinal homeostasis and maintenance of intestinal stem cells. Mol Cell Biol 27: 7551–7559. doi:10.1128/MCB.01034-07.
23. van Es JH, Haegebarth A, Kujala P, Itzkovitz S, Koo B-K, et al. (2012) A critical role for the Wnt effector Tcf4 in adult intestinal homeostatic self-renewal. Mol Cell Biol 32: 1918–1927. doi:10.1128/MCB.06288-11.

24. Andl T, Reddy ST, Gaddapara T, Millar SE (2002) WNT signals are required for the initiation of hair follicle development. Dev Cell 2: 643–653.
25. Barker N, van Es JH, Kuipers J, Kujala P, van den Born M, et al. (2007) Identification of stem cells in small intestine and colon by marker gene Lgr5. Nature 449: 1003–1007. doi:10.1038/nature06196.
26. Barker N, Tan S, Clevers H (2013) Lgr proteins in epithelial stem cell biology. Development 140: 2484–2494. doi:10.1242/dev.083113.
27. Barker N, Clevers H (2010) Leucine-rich repeat-containing G-protein-coupled receptors as markers of adult stem cells. Gastroenterology 138: 1681–1696. doi:10.1053/j.gastro.2010.03.002.
28. Barker N (2014) Adult intestinal stem cells: critical drivers of epithelial homeostasis and regeneration. Nat Rev Mol Cell Biol 15: 19–33. doi:10.1038/nrm3721.
29. Kazanskaya O, Glinka A, del Barco Barrantes I, Stannek P, Niehrs C, et al. (2004) R-Spondin2 is a secreted activator of Wnt/beta-catenin signaling and is required for Xenopus myogenesis. Dev Cell 7: 525–534. doi:10.1016/j.devcel.2004.07.019.
30. Kim K-A, Kakitani M, Zhao J, Oshima T, Tang T, et al. (2005) Mitogenic influence of human R-spondin1 on the intestinal epithelium. Science 309: 1256–1259. doi:10.1126/science.1112521.
31. Sato T, Vries RG, Snippert HJ, van de Wetering M, Barker N, et al. (2009) Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche. Nature 459: 262–265. doi:10.1038/nature07935.
32. Koo B-K, Spit M, Jordens I, Low TY, Stange DE, et al. (2012) Tumour suppressor RNF43 is a stem-cell E3 ligase that induces endocytosis of Wnt receptors. Nature 488: 665–669. doi:10.1038/nature11308.
33. Hao H-X, Xie Y, Zhang Y, Charlat O, Oster E, et al. (2012) ZNRF3 promotes Wnt receptor turnover in an R-spondin-sensitive manner. Nature 485: 195–200. doi:10.1038/nature11019.
34. Clevers H (2013) The intestinal crypt, a prototype stem cell compartment. Cell 154: 274–284. doi:10.1016/j.cell.2013.07.004.
35. Schuijers J, Clevers H (2012) Adult mammalian stem cells: the role of Wnt, Lgr5 and R-spondins. EMBO J 31: 2685–2696. doi:10.1038/emboj.2012.149.
36. Clevers HC, Bevins CL (2013) Paneth cells: maestros of the small intestinal crypts. Annu Rev Physiol 75: 289–311. doi:10.1146/annurev-physiol-030212-183744.
37. Snippert HJ, van der Flier LG, Sato T, van Es JH, van den Born M, et al. (2010) Intestinal crypt homeostasis results from neutral competition between symmetrically dividing Lgr5 stem cells. Cell 143: 134–144. doi:10.1016/j.cell.2010.09.016.
38. Marshman E, Booth C, Potten CS (2002) The intestinal epithelial stem cell. Bioessays 24: 91–98. doi:10.1002/bies.10028.
39. Cheng H, Leblond CP (1974) Origin, differentiation and renewal of the four main epithelial cell types in the mouse small intestine. V. Unitarian Theory of the origin of the four epithelial cell types. Am J Anat 141: 537–561. doi:10.1002/aja.1001410407.
40. Bjerknes M, Cheng H (1981) The stem-cell zone of the small intestinal epithelium. I. Evidence from Paneth cells in the adult mouse. Am J Anat 160: 51–63. doi:10.1002/aja.1001600105.
41. Bjerknes M, Cheng H (1981) The stem-cell zone of the small intestinal epithelium. III. Evidence

from columnar, enteroendocrine, and mucous cells in the adult mouse. Am J Anat 160: 77–91. doi:10.1002/aja.1001600107.

42. Bjerknes M, Cheng H (1999) Clonal analysis of mouse intestinal epithelial progenitors. Gastroenterology 116: 7–14.
43. Potten CS (1977) Extreme sensitivity of some intestinal crypt cells to X and gamma irradiation. Nature 269: 518–521.
44. Potten CS, Hume WJ, Reid P, Cairns J (1978) The segregation of DNA in epithelial stem cells. Cell 15: 899–906.
45. Potten CS, Owen G, Booth D (2002) Intestinal stem cells protect their genome by selective segregation of template DNA strands. J Cell Sci 115: 2381–2388.
46. Cairns J (1975) Mutation selection and the natural history of cancer. Nature 255: 197–200.
47. Escobar M, Nicolas P, Sangar F, Laurent-Chabalier S, Clair P, et al. (2011) Intestinal epithelial stem cells do not protect their genome by asymmetric chromosome segregation. Nat Commun 2: 258. doi:10.1038/ncomms1260.
48. Schepers AG, Vries R, van den Born M, van de Wetering M, Clevers H (2011) Lgr5 intestinal stem cells have high telomerase activity and randomly segregate their chromosomes. EMBO J 30: 1104–1109. doi:10.1038/emboj.2011.26.
49. Steinhauser ML, Bailey AP, Senyo SE, Guillermier C, Perlstein TS, et al. (2012) Multi-isotope imaging mass spectrometry quantifies stem cell division and metabolism. Nature 481: 516–519. doi:10.1038/nature10734.
50. Barker N, van Oudenaarden A, Clevers H (2012) Identifying the stem cell of the intestinal crypt: strategies and pitfalls. Cell Stem Cell 11: 452–460. doi:10.1016/j.stem.2012.09.009.
51. Snippert HJ, Haegebarth A, Kasper M, Jaks V, van Es JH, et al. (2010) Lgr6 marks stem cells in the hair follicle that generate all cell lineages of the skin. Science 327: 1385–1389. doi:10.1126/science.1184733.
52. Barker N, Huch M, Kujala P, van de Wetering M, Snippert HJ, et al. (2010) Lgr5(+ve) stem cells drive self-renewal in the stomach and build long-lived gastric units in vitro. Cell Stem Cell 6: 25–36. doi:10.1016/j.stem.2009.11.013.
53. Barker N, Rookmaaker MB, Kujala P, Ng A, Leushacke M, et al. (2012) Lgr5(+ve) Stem/Progenitor Cells Contribute to Nephron Formation during Kidney Development. Cell Rep 2: 540–552. doi:10.1016/j.celrep.2012.08.018.
54. Huch M, Bonfanti P, Boj SF, Sato T, Loomans CJM, et al. (2013) Unlimited in vitro expansion of adult bi-potent pancreas progenitors through the Lgr5/R-spondin axis. EMBO J 32: 2708–2721. doi:10.1038/emboj.2013.204.
55. Huch M, Dorrell C, Boj SF, van Es JH, Li VSW, et al. (2013) In vitro expansion of single Lgr5+ liver stem cells induced by Wnt-driven regeneration. Nature 494: 247–250. doi:10.1038/nature11826.
56. Plaks V, Brenot A, Lawson DA, Linnemann JR, Van Kappel EC, et al. (2013) Lgr5-expressing cells are sufficient and necessary for postnatal mammary gland organogenesis. Cell Rep 3: 70–78. doi:10.1016/j.celrep.2012.12.017.
57. Wilson CL, Ouellette AJ, Satchell DP, Ayabe T, López-Boado YS, et al. (1999) Regulation of intestinal alpha-defensin activation by the metalloproteinase matrilysin in innate host defense. Science 286: 113–117.

58. Salzman NH, Hung K, Haribhai D, Chu H, Karlsson-Sjöberg J, et al. (2009) Enteric defensins are essential regulators of intestinal microbial ecology. Nat Immunol 11: 76–83. doi:10.1038/ni.1825.
59. Sato T, van Es JH, Snippert HJ, Stange DE, Vries RG, et al. (2011) Paneth cells constitute the niche for Lgr5 stem cells in intestinal crypts. Nature 469: 415–418. doi:10.1038/nature09637.
60. Sato T, Clevers H (2013) Growing self-organizing mini-guts from a single intestinal stem cell: mechanism and applications. Science 340: 1190–1194. doi:10.1126/science.1234852.
61. Farin HF, van Es JH, Clevers H (2012) Redundant sources of Wnt regulate intestinal stem cells and promote formation of Paneth cells. Gastroenterology 143: 1518–1529.e7. doi:10.1053/j.gastro.2012.08.031.
62. Pellegrinet L, Rodilla V, Liu Z, Chen S, Koch U, et al. (2011) Dll1- and dll4-mediated notch signaling are required for homeostasis of intestinal stem cells. Gastroenterology 140: 1230–1240.e1–7. doi:10.1053/j.gastro.2011.01.005.
63. Wong VWY, Stange DE, Page ME, Buczacki S, Wabik A, et al. (2012) Lrig1 controls intestinal stem-cell homeostasis by negative regulation of ErbB signalling. Nat Cell Biol 14: 401–408. doi:10.1038/ncb2464.
64. Haramis A-PG, Begthel H, van den Born M, van Es J, Jonkheer S, et al. (2004) De novo crypt formation and juvenile polyposis on BMP inhibition in mouse intestine. Science 303: 1684–1686. doi:10.1126/science.1093587.
65. Koo B-K, Stange DE, Sato T, Karthaus W, Farin HF, et al. (2012) Controlled gene expression in primary Lgr5 organoid cultures. Nat Methods 9: 81–83. doi:10.1038/nmeth.1802.
66. Schwank G, Koo B-K, Sasselli V, Dekkers JF, Heo I, et al. (2013) Functional repair of CFTR by CRISPR/Cas9 in intestinal stem cell organoids of cystic fibrosis patients. Cell Stem Cell 13: 653–658. doi:10.1016/j.stem.2013.11.002.
67. Nam J-S, Turcotte TJ, Smith PF, Choi S, Yoon JK (2006) Mouse cristin/R-spondin family proteins are novel ligands for the Frizzled 8 and LRP6 receptors and activate beta-catenin-dependent gene expression. J Biol Chem 281: 13247–13257. doi:10.1074/jbc.M508324200.
68. Wei Q, Yokota C, Semenov MV, Doble B, Woodgett J, et al. (2007) R-spondin1 is a high affinity ligand for LRP6 and induces LRP6 phosphorylation and beta-catenin signaling. J Biol Chem 282: 15903–15911. doi:10.1074/jbc.M701927200.
69. Binnerts ME, Kim K-A, Bright JM, Patel SM, Tran K, et al. (2007) R-Spondin1 regulates Wnt signaling by inhibiting internalization of LRP6. Proc Natl Acad Sci USA 104: 14700–14705. doi:10.1073/pnas.0702305104.
70. de Lau W, Peng WC, Gros P, Clevers H (2014) The R-spondin/Lgr5/Rnf43 module: regulator of Wnt signal strength. Genes Dev 28: 305–316. doi:10.1101/gad.235473.113.
71. de Lau W, Barker N, Low TY, Koo B-K, Li VSW, et al. (2011) Lgr5 homologues associate with Wnt receptors and mediate R-spondin signalling. Nature 476: 293–297. doi:10.1038/nature10337.
72. Carmon KS, Gong X, Lin Q, Thomas A, Liu Q (2011) R-spondins function as ligands of the orphan receptors LGR4 and LGR5 to regulate Wnt/beta-catenin signaling. Proc Natl Acad Sci USA 108: 11452–11457. doi:10.1073/pnas.1106083108.
73. Ruffner H, Sprunger J, Charlat O, Leighton-Davies J, Grosshans B, et al. (2012) R-Spondin potentiates Wnt/β-catenin signaling through orphan receptors LGR4 and LGR5. PLoS ONE 7: e40976. doi:10.1371/journal.pone.0040976.

74. Glinka A, Dolde C, Kirsch N, Huang Y-L, Kazanskaya O, et al. (2011) LGR4 and LGR5 are R-spondin receptors mediating Wnt/β-catenin and Wnt/PCP signalling. EMBO Rep 12: 1055–1061. doi:10.1038/embor.2011.175.
75. Gong X, Carmon KS, Lin Q, Thomas A, Yi J, et al. (2012) LGR6 is a high affinity receptor of R-spondins and potentially functions as a tumor suppressor. PLoS ONE 7: e37137. doi:10.1371/journal.pone.0037137.
76. Zhou Y, Lan J, Wang W, Shi Q, Lan Y, et al. (2013) ZNRF3 acts as a tumour suppressor by the Wnt signalling pathway in human gastric adenocarcinoma. J Mol Histol. doi:10.1007/s10735-013-9504-9.
77. Ryland GL, Hunter SM, Doyle MA, Rowley SM, Christie M, et al. (2013) RNF43 is a tumour suppressor gene mutated in mucinous tumours of the ovary. J Pathol 229: 469–476. doi:10.1002/path.4134.
78. Jiang X, Hao H-X, Growney JD, Woolfenden S, Bottiglio C, et al. (2013) Inactivating mutations of RNF43 confer Wnt dependency in pancreatic ductal adenocarcinoma. Proceedings of the National Academy of Sciences. doi:10.1073/pnas.1307218110.
79. Wu J, Jiao Y, Dal Molin M, Maitra A, de Wilde RF, et al. (2011) Whole-exome sequencing of neoplastic cysts of the pancreas reveals recurrent mutations in components of ubiquitin-dependent pathways. Proceedings of the National Academy of Sciences 108: 21188–21193. doi:10.1073/pnas.1118046108.
80. Furukawa T, Kuboki Y, Tanji E, Yoshida S, Hatori T, et al. (2011) Whole-exome sequencing uncovers frequent GNAS mutations in intraductal papillary mucinous neoplasms of the pancreas. Sci Rep 1: 161. doi:10.1038/srep00161.
81. Kinde I, Bettegowda C, Wang Y, Wu J, Agrawal N, et al. (2013) Evaluation of DNA from the Papanicolaou test to detect ovarian and endometrial cancers. Sci Transl Med 5: 167ra4. doi:10.1126/scitranslmed.3004952.
82. Ong CK, Subimerb C, Pairojkul C, Wongkham S, Cutcutache I, et al. (2012) Exome sequencing of liver fluke-associated cholangiocarcinoma. Nat Genet 44: 690–693. doi:10.1038/ng.2273.
83. Moffat LL, Robinson RE, Bakoulis A, Clark SG (2014) The conserved transmembrane RING finger protein PLR-1 downregulates Wnt signaling by reducing Frizzled, Ror and Ryk cell-surface levels in C. elegans. Development 141: 617–628. doi:10.1242/dev.101600.
84. Tan S, Barker N (2014) Epithelial stem cells and intestinal cancer. Semin Cancer Biol. doi:10.1016/j.semcancer.2014.02.005.

CHAPTER 2

Structure of Stem Cell Growth Factor R-spondin 1 in Complex with the Ectodomain of Its Receptor LGR5

Weng Chuan Peng,[1,3] Wim de Lau,[2,3] Federico Forneris,[1] Joke C.M. Granneman,[1] Meritxell Huch,[2] Hans Clevers,[2,*] and Piet Gros[1,*]

Cell Reports 3, 1885–1892, June 27, 2013

[1]Crystal and Structural Chemistry, Bijvoet Center for Biomolecular Research, Department of Chemistry, Faculty of Science, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands

[2]Hubrecht Institute, Royal Netherlands Academy of Arts and Sciences and University Medical Center Utrecht, Uppsalalaan 8, 3584 CT Utrecht, The Netherlands

[3]These authors contributed equally to this work
*Corresponding authors

ABSTRACT

Leucine-rich repeat-containing G protein-coupled receptors 4–6 (LGR4–LGR6) are receptors for R-spondins, potent Wnt agonists that exert profound trophic effects on Wnt-driven stem cells compartments. We present crystal structures of a signaling-competent fragment of R-spondin 1 (Rspo1) at a resolution of 2.0 Å and its complex with the LGR5 ectodomain at a resolution of 3.2 Å. Ecto-LGR5 binds Rspo1 at its concave leucine-rich repeat (LRR) surface, forming a dimeric 2:2 complex. Fully conserved residues on LGR4–LGR6 explain promiscuous binding of R-spondins. A phenylalanine clamp formed by Rspo1 Phe106 and Phe110 pinches Ala190 of LGR5 and is critical for binding. Mutations related to congenital anonychia reduce signaling, but not binding of Rspo1 to LGR5. Furthermore, antibody binding to the extended loop of the C-terminal LRR cap of LGR5 activates signaling in a ligand-independent manner. Thus, our data reveal binding of R-spondins to conserved sites on LGR4–LGR6 and, in analogy to FSHR and related receptors, suggest a direct signaling role for LGR4–LGR6 in addition to its formation of Wnt receptor and coreceptor complexes.

INTRODUCTION

Vertebrate genomes encode four secreted R-spondin proteins (Rspo1–Rspo4), each defined by two N-terminal Furin (Fu) domains and a thrombospondin (Tsp) domain. Functionally, R-spondin proteins act as potent enhancers of Wnt signals (Kazanskaya et al., 2004). Indeed, Rspo1 strongly promotes proliferation of the Wnt-dependent intestinal-crypt stem cell compartment in vivo (Kim et al., 2005) and in vitro (Sato et al., 2009). This activity can be attributed to the two Fu domains, because a Fu1-Fu2 fragment of Rspo1 retained full signaling activity (Kim et al., 2008, Li et al., 2009). R-spondin mutations have been found in two hereditary syndromes in humans. Rspo1 is mutated in a recessive syndrome characterized by XX sex reversal, palmoplantar hyperkeratosis, and squamous cell carcinomas (Schuijers and Clevers, 2012). Mutations in the Rspo4 gene result in congenital anonychia, a severe hypoplasia of fingernails and toenails (Blaydon et al., 2006, Brüchle et al., 2008, Wasif and Ahmad, 2013).

The Wnt target gene leucine-rich repeat-containing G protein-coupled receptor 5 (Lgr5) encodes a serpentine receptor that is exquisitely specific to Wnt-dependent stem cells of a series of adult tissues, including small intestine and colon (Barker et al., 2007), stomach (Barker et al., 2010), hair follicle (Jaks et al., 2008), liver (Huch et al., 2013), kidney (Barker et al., 2012), and mammary gland (Plaks et al., 2013). The LGRs form a small family of seven-transmembrane (7TM) receptors that include the follicle-stimulating hormone, luteinizing hormone, and thyroid-stimulating hormone receptors (FSHR, LHR, and TSHR, also referred to as LGR1–LGR3, respectively) (Hsu et al., 1998). LGR5 (as well as its homologs LGR4 and LGR6) binds R-spondins with high affinity, thus mediating R-spondin input into the canonical Wnt pathway (Carmon et al., 2011, de Lau et al., 2011, Glinka et al., 2011). Indeed, LGR4 and LGR5 proteins physically reside within Frizzled/LRP receptor complexes (de Lau et al., 2011). Whereas deletion of the Lgr5 gene in the intestine has little effect, mutation of Lgr4

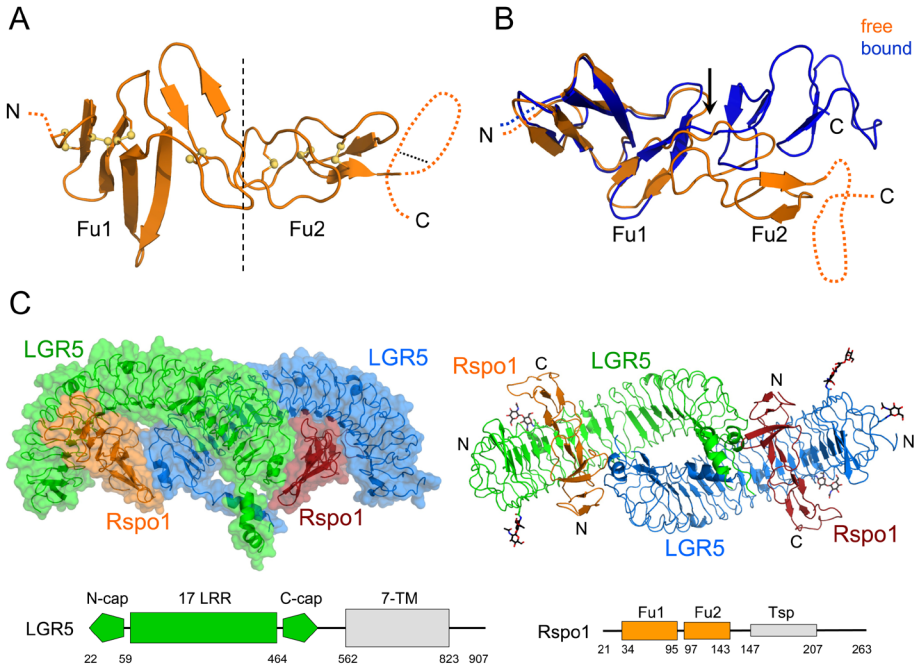


Fig 1. Structure of Rspo1 and the complex of Rspo1 with ecto-LGR5
A, Structure of (unbound) Rspo1-Fu1Fu2 (res. 31-145) at 2-Å resolution. Indicated are disulphide bonds (ball and stick) and disordered residues (dash line).
B, Overlay of bound and unbound Rspo1. Arrow indicates a hinge around which the orientation between the Fu1 and Fu2 domains differ by ~90° between the LGR5-bound and unbound structure of Rspo1.
C, Structure of LGR5 ecto-domain (res. 22-543) in complex with Rspo1 Fu1-Fu2 domains at 3.2-Å resolution in two views, one without and one with surface representation. Domain compositions of LGR5 and Rspo1 are indicated schematically.

(which is expressed by all crypt cells) severely decreases crypt proliferation (Mustata et al., 2011). Double Lgr4 and Lgr5 knockout completely abolishes proliferation (de Lau et al., 2011), implying that R-spondins are major drivers of Wnt-dependent crypt self-renewal.

RESULTS AND DISCUSSIONS

Structures of Free and Bound Rspo1

We sought to address the crystal structure of the Fu1-Fu2 fragment of Rspo1 and its complex with the ligand-binding ectodomain of LGR5. A crystal structure of Rspo1-Fu1Fu2 (residues 31–145) was derived at a resolution of 2.0 Å (Figure 1A). Fu domains are rich in cysteine-knotted β-hairpins. In contrast to earlier mass-spectrometry analysis (Li et al., 2009), Rspo1-Fu1Fu2 displayed disulphide-bond patterns common to other Fu domains (Garrett et al., 1998, Ogiso et al., 2002). Each Fu domain formed a leaflet consisting of three β-hairpins connected

by disulphide bonds (Figure S1). The sets of Fu1 and Fu2 β-hairpins were oriented at ~90° to each other in unbound Rspo1. Next, we determined the structure of this Rspo1 fragment in complex with the LGR5 ectodomain (residues 22–543) at 3.2 Å resolution (Figures 1B and 1C). In the complex, a twist due to a rotation around the longitudinal axis of Rspo1 aligned the two sets of β-hairpins, thus flattening the shape of Rspo1 when bound to LGR5.

Structure of the LGR5-Rspo1 Complex

Rspo1-Fu1Fu2 binds the ecto-LGR5 domain with a K_D of 2–3 nM (de Lau et al., 2011, Glinka et al., 2011). We observed dimeric, 2:2, arrangements of the complex in four crystal structures, determined up to 3.2 Å resolution (Figure 1C and Figure S2). Size-exclusion chromatography indicated 1:1 complexes (Figure S3). In agreement with a physiological existence of LGR dimers on cells, previous mass-spectrometry analysis has revealed interactions between LGR4 and LGR5 in the cell membrane (de Lau et al., 2011). In the absence of the cell membrane and the 7TM region, the interaction between the LGR ectodomains may be lost in solution. Of note, a similar observation was made for FSHR (Fan and Hendrickson, 2005) and TLR5 (Yoon et al., 2012).

The LGR5 ectodomain adopted a typical horseshoe-shaped structure consisting of the 17 leucine-rich repeat (LRR) units (Figure 1C and Figure S4). The N-terminal and C-terminal caps (N- and C-caps) were similar to those of FSHR (Jiang et al., 2012) (Figures S4C and S4D). A long loop in the C-cap contained an α helix but was largely disordered, possibly adopting a stable structure upon interaction with the extracellular loops of the 7TM region. The LRR curve was kinked and twisted by ~20° between LRR10 and LRR11. This kink coincided with a marked sequence variation in LRR9 and LRR10, wherein two bulky phenylalanines occupied the positions of canonical leucines (Figures S4A and S5). Several LRR-containing receptors show curvatures similar to the N-terminal LRR1–LRR10 or the C-terminal LRR11–LRR17 (Figure S6). Due to the kink and twisting of the C-terminal LRR11–LRR17 (yielding an overall twist of ~45°), the protomers bent toward each other in the dimer. Contact points in the dimer bridged LRR10 to LRR17 (Figure S7A) with a large “open” area between the dimer partners, except for an H-bonded interaction between Tyr-361 and its dimeric partner at the center.

LGR5-Rspo1 Binding Interfaces

The ecto-LGR5 and Rspo1-Fu1Fu2 complex observed in the crystals revealed three receptor-ligand contact sites (Figures 2A–2C). Two adjoining sites were formed by the concave surface of the LRR3–LRR9 of LGR5 with Fu2 and Fu1 of Rspo1, burying a total surface area of 870 $Å^2$. A third site (burying 340 $Å^2$) was formed “in trans” between Rspo1-Fu1 and the second copy of LGR5.

In the first contact site, phenylalanine residues 106 and 110 of the Rspo1-Fu2 domain formed a clamp-like arrangement around Ala190 of LGR5 (Figure 2B). Ala190 was surrounded by a hydrophobic rim formed by Cβ atoms and/or side chains of His166, Trp168, Gln189, Val213,

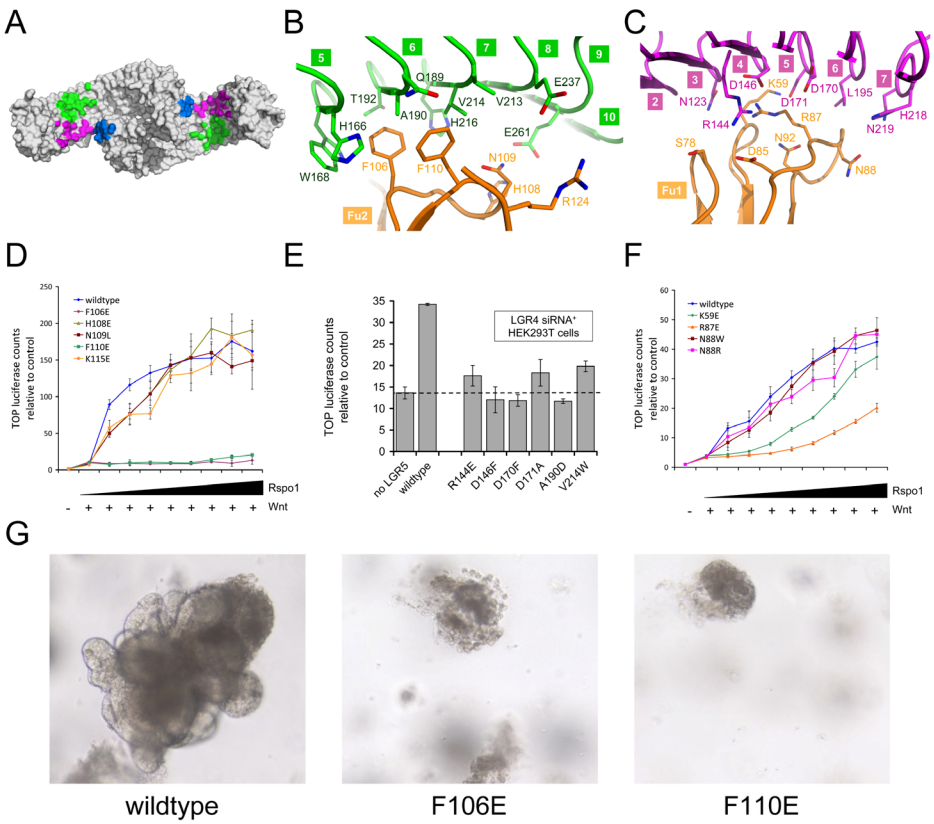


Fig 2. LGR5-Rspo1 interfaces

A, Overview of three binding interfaces between Rspo1 and LGR5. Shown are contact footprints of Rspo1 on LGR5 with contacts of Fu2 indicated in green and for Fu1 in magenta (with 'cis' LGR5) and blue ('trans' LGR5).

B, Interactions between Fu2 of Rspo1 with LGR5;

C, idem between Fu1 of Rspo1 with LGR5 (cis).

D, TOPFlash-assay results for purified Rspo1 mutants from 1.5 to 200 nM (with K115E as additional control).

E, TOPFlash results obtained by transfection-mediated introduction of 0.012 to 15 ng of wildtype and LGR5 variants, mutated on residues interacting with Rspo1. One day prior to introducing LGR5 and reporter plasmids, cells were transfected with LGR4-specific siRNA. The 15 ng results are shown.

F, TOPFlash results for mutations in Rspo1 Fu2 domain (concentrations as in 2D).

G, Small intestinal organoid growth in the presence of wildtype or Rspo1 mutants.

Val214, and His216. On the side of this hydrophobic patch, glutamates 237 and 261 of LGR5 formed H-bonds and salt bridges with His108 and Asn109 from the 106FSHNF110 loop and with Arg-124 from a neighboring loop. All these contact residues of LGR5 are strictly conserved among LGR4–LGR6 (Figure S5). Between the R-spondins, Phe106 and Phe110 are fully conserved, the intervening residues are all hydrophilic or charged, and position 124 is

either an arginine or a lysine (Figure S1B). We generated the Rspo1-Fu1Fu2 mutants F106E and F110E and found that these had no activity in the TOPFlash Wnt reporter assay (Figure 2D). Mutations A190D and V214W of LGR5, disrupting the shallow hydrophobic bowl, showed reduced signaling activity (Figure 2E).

Residues from Rspo1-Fu1 and LGR5 residues from the "lower" part of the concave surface of LRR3–LRR7 formed the second contact site (Figure 2C). This site was predominantly of charged character. Again, the LGR5 residues at the interface were strictly conserved between LGR4–LGR6. These residues were Asn123, Arg144, Asp146, Asp170, Asp171, Leu195, His218, and Asn219. The corresponding Rspo1 residues were Lys59, Ser78, Asp85, Arg87, Asn88, and Asn92. Lys59 and Arg87, at the center of this site, were conserved as lysines or arginines between all four R-spondins. Charge-reversal mutations R87E and K59E showed reduced activity, confirming their role in the interaction (Figure 2F). Mutation of the nonconserved Asn88 (on the side of the interface) had no effect on Rspo1 activity. LGR5 mutants D146F and D170F had lost all signaling activity, whereas mutants R144E and D171A showed reduced activity (Figure 2E). We concluded that Rspo1 utilizes both its Fu1 and Fu2 domains to bind the ectodomain of LGR5 at sites that are fully conserved between LGR4–LGR6. This was in agreement with the observed lack of specificity of each of the R-spondins for LGR4–LGR6 (Carmon et al., 2011, de Lau et al., 2011, Glinka et al., 2011). In vitro "minigut" culture, which depends entirely on functional R-spondin (Sato et al., 2009), confirmed that Rspo1-FuFu2 stimulated organoid growth and two binding-defective Rspo1-Fu1Fu2 F106E and F110E mutants failed to support growth (Figure 2G).

The trans Site and Dimer Interface

In the third site, Rspo1-Fu1 contacted the dimeric partner LGR5 at the last LRR and short α helix of the C-cap (Figure 3A). The interface was formed by LGR5 residues from the C-cap, Gln457, Ser458, Leu459, and Tyr477 and Rspo1 residues Asn51, Leu54, Leu64, Gln71, Asn88, and Met91. However, in several protein copies in the crystal structures, we observed disorder at this interface (Figure S8). The side chain of Rspo1 Gln71 made a H-bond to backbone carbonyl oxygen of the Tyr477 of LGR5 (Figure 3A); both residues are strictly conserved between Rspo1–Rspo4 and LGR4–LGR6, respectively. Gln71 of Rspo1 coincides with the position of one of the four Rspo4 mutations described in patients with congenital anonychia (Blaydon et al., 2006, de Lau et al., 2012). The four related residues are Arg66, Arg70, Gln71, and Gly73, which were located in the second β-hairpin loop of Rspo1-Fu1. Anonychia mutants R66W, Q71R, R70C, and G73R and related mutations in Rspo1-Fu1Fu2 did not affect binding to ecto-LGR5, but showed reduced signaling activity (Figures 3B and 3C). Q71R and G73R point toward the "trans" interface and may disrupt interactions. LGR5 mutations in the trans site showed both reduced (S458R) and enhanced (L459R) TOPFlash activity (Figure S7); these findings could not be corroborated further due to lack of expression on the cell surface of the tested mutants. Next, we mutated LGR5 residues in the dimer interface (Figure S7). Residues Tyr289, Asp290, and His454 were observed at the LGR5-LGR5 dimer interface (Figure S7A). Mutation Y289A/D290A, Y289W/D290A, or H454A did not significantly reduce TOPFlash activity (Figure S7B),

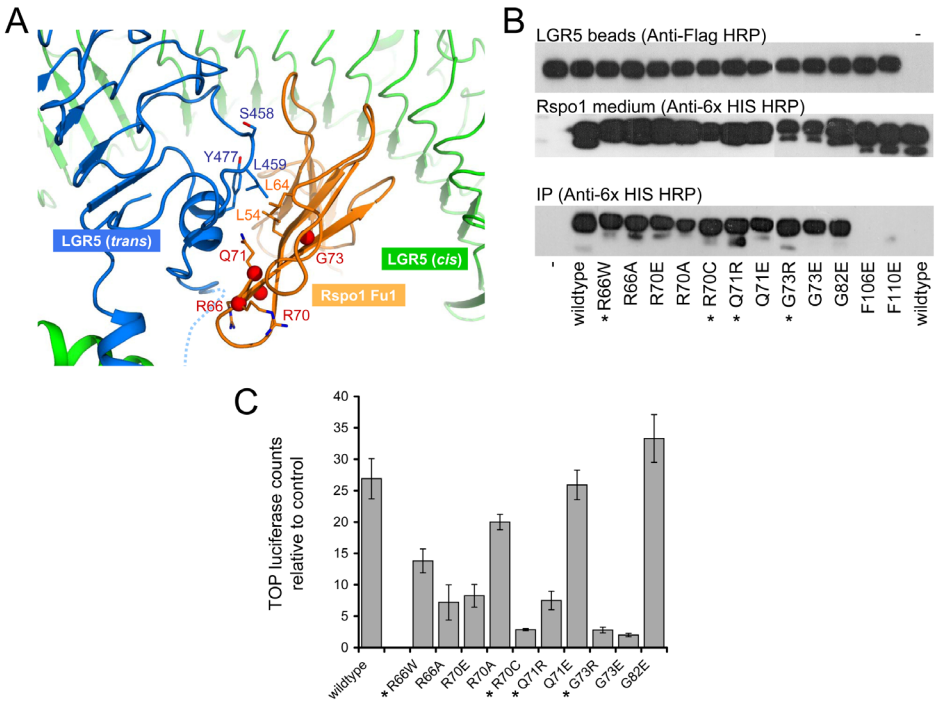


Fig 3. 'Trans' and dimer interactions

A, 'Trans' contact site between Rspo1-Fu1 and the neighbouring C-cap of LGR5 with positions of Anonychia-related (Rspo4) residues indicated by red spheres.

B, Binding assay of Rspo1 Anonychia location-specific mutations, present in conditioned media (middle panel), to immobilized LGR5 ecto-domain (top panel). Bottom panel shows amount of protein captured. Actual patient mutations are indicated (*). G82E is used as positive control and F106E and F110E as negative controls.

C, TOPFlash-assay using conditioned-media derived from wt and Rspo1 mutants.

indicating that the receptor-receptor interface observed in the crystal structure is not critical and that dimer formation may depend on receptor-receptor interactions in the membrane (de Lau et al., 2011).

Activation by Antibody Binding to Disordered C-Cap Loop in the Absence of Rspo1

Ligand-activated signaling is well established for the glycohormone receptors FSHR, TSHR, and LHR (Jeoung et al., 2007, Ryu et al., 1998, Simoni et al., 1997). However, a downstream G-protein has not been identified for LGR5. It is unknown whether LGR5, or its homologs LGR4 and LGR6, is directly involved in transmembrane signaling, or whether the receptor serves to capture its ligand and affect Wnt signaling solely through ternary-complex formation, with LRP5/LRP6 (Wei et al., 2007), Frizzled (Nam et al., 2006), and RNF43/ZNRF3 (Hao et al., 2012), for example. We tested a series of LGR5-specific monoclonal antibodies and observed that

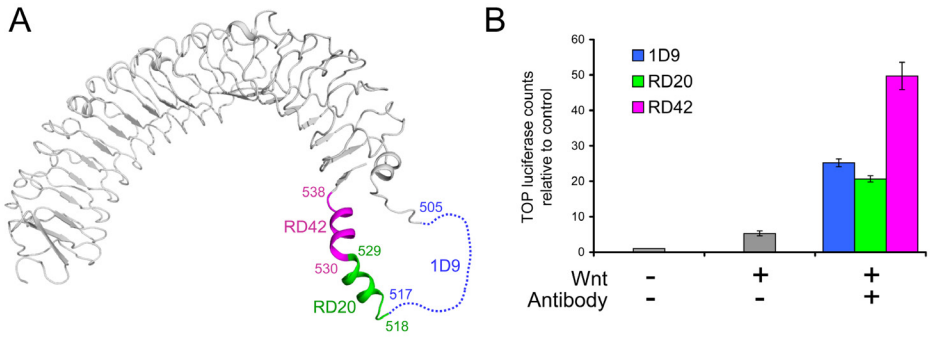


Fig 4. Ligand-independent antibody activation of LGR5
A, Epitopes of antibodies RD20, RD42 and 1D9 mapped onto ecto-LGR5 (the precise location of this protruding region varies among crystal structures, see Fig S2B.
B, TOPFlash-assay for LGR5-specific antibodies RD20, RD42 and 1D9 in the absence of R-spondins.

three of them (1D9, RD20, and RD42) induced TOPFlash activity in human embryonic kidney 293 cells (HEK293 cells) stably expressing LGR5 in the absence of R-spondins (Figures 4A and 4B). The epitopes of these antibodies were mapped onto a flexible region in the ecto-LGR5 C-cap loop (notably, the C-cap was not required for ligand binding as shown for a shortened construct in Figure S3). In contrast, antibody 4D11, which binds to LRR9–LRR11, showed no activity in the TOPFlash assay (data not shown). For the glycohormone receptors, this region (referred to as the hinge region) has been implied to have an autoinhibitory, reverse-agonistic role (Agrawal and Dighe, 2009, Majumdar et al., 2012, Mueller et al., 2010, Vlaeminck-Guillem et al., 2002, Zeng et al., 2001, Zhang et al., 2000). Similarly, antibodies against FSHR (Majumdar et al., 2012) and TSHR (Majumdar and Dighe, 2012) bind between the LRR and 7TM regions and activate signaling, putatively through inducing a conformational change that alleviates the autoinhibitory activity. The TOPFlash activity upon antibody binding may indicate a similar process in LGR5. Moreover, the positions of the anonychia-related mutations provide further indication of interactions beyond the contact sites formed by the ecto-LGR5, with residues Arg66 and Arg70 pointing away from the trans site to where the extended C-cap loop or 7TM is expected (see Figure S9 for a presentation of a hypothetical ectodomain and 7TM region arrangement). To verify a role of the 7TM in signaling, we rigorously substituted residues 562–907 of LGR5 with the corresponding region of an unrelated, single-pass membrane protein, GPA33. Fluorescence-activated cell sorting (FACS) analysis showed that the chimeric receptor is expressed on the cell surface (Figure S10A), and Rspo1-alkaline phosphate staining on the cell surface indicates that the Rspo1 is able to bind to the chimeric receptor (Figure S10B). However, TOPFlash activity reduced to basal levels (Figure S10C), supporting the notion that binding of Rspo1 to the LGR5 ectodomain is not sufficient for signaling.

CONCLUDING REMARKS

The current structural analysis of Rspo1 in complex with the ectodomain of LGR5 provides a framework from which to build insights into the molecular mechanism by which these two

molecules support a wide range of Wnt-dependent stem cell types. R-spondins have been reported to interact with additional membrane receptors such as LRP5/LRP6 (Wei et al., 2007), Frizzled (Nam et al., 2006), RNF43/ZNRF3 (Hao et al., 2012), and Syndecan-4 (Ohkawara et al., 2011). The Wnt agonist Norrin binds LGR4–LGR6 in addition to its interaction with Frizzled4 (Deng et al., 2013). These proteins either compete for binding or, potentially, form complexes. The short R-spondin Fu1Fu2 fragment suffices for full activity in assays in vitro (Kim et al., 2008, Li et al., 2009). Though most of the Fu1-Fu2 surfaces of R-spondins will be in contact with LGR4–LGR6, additional interactions with Fu1 and Fu2 domains may allow the formation of ternary complexes with LRP5/LRP6, Frizzled, or RNF43/ZNRF3, with or without changing the overall arrangement of the R-spondin-LGR4–LGR6 complexes. Furthermore, in full-length R-spondins, the Tsp and C-terminal tail, not studied here, probably point outward from the complex, providing additional potential contact sites close to the membrane surface—for Syndecan-4, for instance (Ohkawara et al., 2011) (Figure S9). Ligand-independent signaling through antibody binding and residues in the C-cap region that are critical for signaling are features reminiscent of signaling in FSHR, TSHR, and LHR (Mizutori et al., 2008, Mueller et al., 2010). Critical interactions beyond ligand binding to the LRR of LGR5 are further supported by the location of the anonychia-related (Rspo4) mutations (Blaydon et al., 2006, de Lau et al., 2012). Taken together, our data indicate that R-spondins bind tightly to the strictly conserved binding sites formed by LRR3–LRR9 of LGR4–LGR6 and that monomeric 1:1 interactions are sufficient for this binding to occur. Full-length LGR4–LGR6 molecules, however, dimerize to establish the trans contact sites that are implied in signaling. Signaling may also occur upon antibody binding to the C-cap loop region independent of binding R-spondins. In analogy to FSHR, TSHR, and LHR (Mueller et al., 2010, Vassart et al., 2004), this binding may induce conformational changes alleviating the putative autoinhibitory, reverse-agonistic activity. Thus, these data suggest a direct role of signaling by the LGR4–LGR6 in addition to ternary-complex formations. Hence, a bewildering complexity exists at the level of the initiation of Wnt signals at the cell surface because of the existence of 19 Wnts, 10 Frizzleds, and 2 LRP Wnt coreceptors. A variety of secreted and membrane-bound Wnt agonists and antagonists add a further level of complexity. The facultative R-spondin-LGR receptor module appears to represent a vertebrate invention for magnifying Wnt signal strength and thus enlarging stem cell compartments. Recent structures have been resolved for Wnt-Frizzled (Janda et al., 2012) and Dkk1-LRP (Ahn et al., 2011, Bourhis et al., 2011, Chen et al., 2011, Cheng et al., 2011). Many additional studies will be required for understanding the inner workings of what must be among the most complex receptor-ligand systems in animal biology.

EXPERIMENTAL PROCEDURES

Ecto-LGR5 (residues 22–543) and Rspo1-Fu1Fu2 (residues 31–145) proteins were produced recombinantly in HEK293 N-acetylglucoaminyltransferase I-deficient (GnTI–) Epstein-Barr virus nuclear antigen I (EBNA) cells, purified to homogeneity, and crystallized with the hanging-drop vapor-diffusion method. Diffraction data were collected at the Swiss Light Source (SLS) in Villigen, Switzerland and at the European Synchrotron Radiation Facility (ESRF) in Grenoble, France. The structures of Rspo-Fu1Fu2 and the ecto-LGR5-Rspo-Fu1Fu2 complex were

determined by experimental phasing and molecular replacement, respectively. Diffraction-data and refinement statistics are provided in Table S1; examples of the electron densities are shown in Figure S11. Monitoring the potential of wild-type and mutated variants of Rspo1-Fu1Fu2 to enhance Wnt signals was done by employing the TCF-dependent TOP luciferase assay in HEK293T cells (Staal et al., 1999). Measuring the ability of mutated versions of the LGR5 to transmit Rspo1-Fu1fu2-driven signaling was performed using the same luciferase assay but with simultaneous small interfering RNA (siRNA)-driven knockdown of endogenous LGR4. The LGR5-binding ability of Rspo1-Fu1Fu2 proteins mutated in the LGR5-interacting domain and Rspo1-Fu1Fu2 mutants representing congenital anonychia mutations were tested in an immunoprecipitation experiment with the ectodomain of LGR5 coated on agarose beads. Additional details are available in the Extended Experimental Procedures.

ACKNOWLEDGEMENTS

We gratefully thank B.T. Burnley, E.G. Huizinga, P.K. Madoori, A.M.M. Schreurs, T. Tenic, H. Teunissen, and M. van de Wetering for discussions and assistance and J. Aho (R&D Systems) for generously providing antibodies. We thank the ESRF and the SLS for the provision of synchrotron radiation facilities and beamline scientists of the SLS, ESRF, and the European Molecular Biology Laboratory for assistance. Financial support was provided by the Netherlands Organization for Scientific Research (NWO-CW grant no. 700.57.010; ZonMW/NGI/Pre-seed grant no. 936.1001), Koningin Wilhelmina Fonds (PF-HUBR 2007-3956), the European Research Council (Advanced Grant no. 233229), and the European Commission's Seventh Framework Programme (FP7/2007-2013) under BioStruct-X (grant no. 283570). W.C.P. performed cloning, expression, purification, crystallization, data collection, structure determination, and refinement. W.d.L. performed cloning; initiated antibody studies, anonychia-related mutations, and chimeric LGR5-GPA33 experiments; and performed signaling assays and immunoprecipitations. M.H. performed intestinal organoid experiments. F.F. performed data collection, structure determination, and refinement and prepared figures. J.C.M.G. performed cloning. W.C.P., W.d.L., F.F., H.C., and P.G. discussed results and wrote the paper. H.C. and P.G. supervised the project.

ACCESSION NUMBERS

The Protein Data Bank (PDB) accession numbers for the coordinates and structure factors of Rspo1 and LGR-Rspo1 reported in this paper are 4BSO, 4BSP, 4BSR, 4BSS, 4BST, and 4BSU.

REFERENCES

Agrawal, G., and Dighe, R.R. (2009). Critical involvement of the hinge region of the follicle-stimulating hormone receptor in the activation of the receptor. J Biol Chem 284, 2636-2647.

Ahn, V.E., Chu, M.L., Choi, H.J., Tran, D., Abo, A., and Weis, W.I. (2011). Structural basis of Wnt signaling inhibition by Dickkopf binding to LRP5/6. Dev Cell 21, 862-873.

Barker, N., Huch, M., Kujala, P., van de Wetering, M., Snippert, H.J., van Es, J.H., Sato, T., Stange, D.E., Begthel, H., van den Born, M., et al. (2010). Lgr5(+ve) stem cells drive self-renewal in the stomach and build long-lived gastric units in vitro. Cell Stem Cell 6, 25-36.

Barker, N., Rookmaaker, M.B., Kujala, P., Ng, A., Leushacke, M., Snippert, H., van de Wetering, M., Tan, S., Van Es, J.H., Huch, M., et al. (2012). Lgr5(+ve) stem/progenitor cells contribute to nephron formation during kidney development. Cell Rep 2, 540-552.

Barker, N., van Es, J.H., Kuipers, J., Kujala, P., van den Born, M., Cozijnsen, M., Haegebarth, A., Korving, J., Begthel, H., Peters, P.J., et al. (2007). Identification of stem cells in small intestine and colon by marker gene Lgr5. Nature 449, 1003-1007.

Blaydon, D.C., Ishii, Y., O'Toole, E.A., Unsworth, H.C., Teh, M.T., Ruschendorf, F., Sinclair, C., Hopsu-Havu, V.K., Tidman, N., Moss, C., et al. (2006). The gene encoding R-spondin 4 (RSPO4), a secreted protein implicated in Wnt signaling, is mutated in inherited anonychia. Nat Genet 38, 1245-1247.

Bourhis, E., Wang, W., Tam, C., Hwang, J., Zhang, Y., Spittler, D., Huang, O.W., Gong, Y., Estevez, A., Zilberleyb, I., et al. (2011). Wnt antagonists bind through a short peptide to the first beta-propeller domain of LRP5/6. Structure 19, 1433-1442.

Bruchle, N.O., Frank, J., Frank, V., Senderek, J., Akar, A., Koc, E., Rigopoulos, D., van Steensel, M., Zerres, K., and Bergmann, C. (2008). RSPO4 is the major gene in autosomal-recessive anonychia and mutations cluster in the furin-like cysteine-rich domains of the Wnt signaling ligand R-spondin 4. J Invest Dermatol 128, 791-796.

Carmon, K.S., Gong, X., Lin, Q., Thomas, A., and Liu, Q. (2011). R-spondins function as ligands of the orphan receptors LGR4 and LGR5 to regulate Wnt/beta-catenin signaling. Proc Natl Acad Sci U S A 108, 11452-11457.

Chen, S., Bubeck, D., MacDonald, B.T., Liang, W.X., Mao, J.H., Malinauskas, T., Llorca, O., Aricescu, A.R., Siebold, C., He, X., et al. (2011). Structural and functional studies of LRP6 ectodomain reveal a platform for Wnt signaling. Dev Cell 21, 848-861.

Cheng, Z., Biechele, T., Wei, Z., Morrone, S., Moon, R.T., Wang, L., and Xu, W. (2011). Crystal structures of the extracellular domain of LRP6 and its complex with DKK1. Nat Struct Mol Biol 18, 1204-1210.

de Lau, W., Barker, N., Low, T.Y., Koo, B.K., Li, V.S., Teunissen, H., Kujala, P., Haegebarth, A., Peters, P.J., van de Wetering, M., et al. (2011). Lgr5 homologues associate with Wnt receptors and mediate R-spondin signalling. Nature 476, 293-297.

de Lau, W.B., Snel, B., and Clevers, H.C. (2012). The R-spondin protein family. Genome Biol 13, 242.

Deng, C., Reddy, P., Cheng, Y., Luo, C.W., Hsiao, C.L., and Hsueh, A.J. (2013). Multi-functional norrin is a ligand for the LGR4 receptor. J Cell Sci 126, 2060-2068.

Fan, Q.R., and Hendrickson, W.A. (2005). Structure of human follicle-stimulating hormone in complex with its receptor. Nature 433, 269-277.

Garrett, T.P., McKern, N.M., Lou, M., Frenkel, M.J., Bentley, J.D., Lovrecz, G.O., Elleman, T.C., Cosgrove, L.J., and Ward, C.W. (1998). Crystal structure of the first three domains of the type-1 insulin-like growth factor receptor. Nature 394, 395-399.

Glinka, A., Dolde, C., Kirsch, N., Huang, Y.L., Kazanskaya, O., Ingelfinger, D., Boutros, M., Cruciat, C.M., and Niehrs, C. (2011). LGR4 and LGR5 are R-spondin receptors mediating Wnt/

beta-catenin and Wnt/PCP signalling. EMBO Rep 12, 1055-1061.

Hao, H.X., Xie, Y., Zhang, Y., Charlat, O., Oster, E., Avello, M., Lei, H., Mickanin, C., Liu, D., Ruffner, H., et al. (2012). ZNRF3 promotes Wnt receptor turnover in an R-spondin-sensitive manner. Nature 485, 195-200.

Hsu, S.Y., Liang, S.G., and Hsueh, A.J. (1998). Characterization of two LGR genes homologous to gonadotropin and thyrotropin receptors with extracellular leucine-rich repeats and a G protein-coupled, seven-transmembrane region. Mol Endocrinol 12, 1830-1845.

Huch, M., Dorrell, C., Boj, S.F., van Es, J.H., Li, V.S., van de Wetering, M., Sato, T., Hamer, K., Sasaki, N., Finegold, M.J., et al. (2013). In vitro expansion of single Lgr5^{+} liver stem cells induced by Wnt-driven regeneration. Nature 494, 247-250.

Jaks, V., Barker, N., Kasper, M., van Es, J.H., Snippert, H.J., Clevers, H., and Toftgard, R. (2008). Lgr5 marks cycling, yet long-lived, hair follicle stem cells. Nat Genet 40, 1291-1299.

Janda, C.Y., Waghray, D., Levin, A.M., Thomas, C., and Garcia, K.C. (2012). Structural basis of Wnt recognition by Frizzled. Science 337, 59-64.

Jeoung, M., Lee, C., Ji, I., and Ji, T.H. (2007). Trans-activation, cis-activation and signal selection of gonadotropin receptors. Mol Cell Endocrinol 260-262, 137-143.

Jiang, X., Liu, H., Chen, X., Chen, P.H., Fischer, D., Sriraman, V., Yu, H.N., Arkinstall, S., and He, X. (2012). Structure of follicle-stimulating hormone in complex with the entire ectodomain of its receptor. Proc Natl Acad Sci U S A 109, 12491-12496.

Kazanskaya, O., Glinka, A., del Barco Barrantes, I., Stannek, P., Niehrs, C., and Wu, W. (2004). R-Spondin2 is a secreted activator of Wnt/beta-catenin signaling and is required for Xenopus myogenesis. Dev Cell 7, 525-534.

Kim, K.A., Kakitani, M., Zhao, J., Oshima, T., Tang, T., Binnerts, M., Liu, Y., Boyle, B., Park, E., Emtage, P., et al. (2005). Mitogenic influence of human R-spondin1 on the intestinal epithelium. Science 309, 1256-1259.

Kim, K.A., Wagle, M., Tran, K., Zhan, X., Dixon, M.A., Liu, S., Gros, D., Korver, W., Yonkovich, S., Tomasevic, N., et al. (2008). R-Spondin family members regulate the Wnt pathway by a common mechanism. Mol Biol Cell 19, 2588-2596.

Li, S.J., Yen, T.Y., Endo, Y., Klauzinska, M., Baljinnyam, B., Macher, B., Callahan, R., and Rubin, J.S. (2009). Loss-of-function point mutations and two-furin domain derivatives provide insights about R-spondin2 structure and function. Cell Signal 21, 916-925.

Majumdar, R., and Dighe, R.R. (2012). The hinge region of human thyroid-stimulating hormone (TSH) receptor operates as a tunable switch between hormone binding and receptor activation. PLoS One 7, e40291.

Majumdar, R., Railkar, R., and Dighe, R.R. (2012). The antibodies against the computationally designed mimic of the glycoprotein hormone receptor transmembrane domain provide insights into receptor activation and suppress the constitutively activated receptor mutants. J Biol Chem 287, 34514-34532.

Mizutori, Y., Chen, C.R., McLachlan, S.M., and Rapoport, B. (2008). The thyrotropin receptor hinge region is not simply a scaffold for the leucine-rich domain but contributes to ligand binding and signal transduction. Mol Endocrinol 22, 1171-1182.

Mueller, S., Jaeschke, H., Gunther, R., and Paschke, R. (2010). The hinge region: an important receptor component for GPHR function. Trends Endocrinol Metab 21, 111-122.

Mustata, R.C., Van Loy, T., Lefort, A., Libert, F., Strollo, S., Vassart, G., and Garcia, M.I. (2011). Lgr4 is required for Paneth cell differentiation and maintenance of intestinal stem cells ex vivo. EMBO Rep 12, 558-564.

Nam, J.S., Turcotte, T.J., Smith, P.F., Choi, S., and Yoon, J.K. (2006). Mouse cristin/R-spondin family proteins are novel ligands for the Frizzled 8 and LRP6 receptors and activate beta-catenin-dependent gene expression. J Biol Chem 281, 13247-13257.

Ogiso, H., Ishitani, R., Nureki, O., Fukai, S., Yamanaka, M., Kim, J.H., Saito, K., Sakamoto, A., Inoue, M., Shirouzu, M., et al. (2002). Crystal structure of the complex of human epidermal growth factor and receptor extracellular domains. Cell 110, 775-787.

Ohkawara, B., Glinka, A., and Niehrs, C. (2011). Rspo3 binds syndecan 4 and induces Wnt/PCP signaling via clathrin-mediated endocytosis to promote morphogenesis. Dev Cell 20, 303-314.

Plaks, V., Brenot, A., Lawson, D.A., Linnemann, J.R., Van Kappel, E.C., Wong, K.C., de Sauvage, F., Klein, O.D., and Werb, Z. (2013). Lgr5-expressing cells are sufficient and necessary for postnatal mammary gland organogenesis. Cell Rep 3, 70-78.

Ryu, K.S., Gilchrist, R.L., Koo, Y.B., Ji, I., and Ji, T.H. (1998). Gene, interaction, signal generation, signal divergence and signal transduction of the LH/CG receptor. Int J Gynaecol Obstet 60 Suppl 1, S9-20.

Sato, T., Vries, R.G., Snippert, H.J., van de Wetering, M., Barker, N., Stange, D.E., van Es, J.H., Abo, A., Kujala, P., Peters, P.J., et al. (2009). Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche. Nature 459, 262-265.

Schuijers, J., and Clevers, H. (2012). Adult mammalian stem cells: the role of Wnt, Lgr5 and R-spondins. EMBO J 31, 2685-2696.

Simoni, M., Gromoll, J., and Nieschlag, E. (1997). The follicle-stimulating hormone receptor: biochemistry, molecular biology, physiology, and pathophysiology. Endocr Rev 18, 739-773.

Staal, F.J., Burgering, B.M., van de Wetering, M., and Clevers, H.C. (1999). Tcf-1-mediated transcription in T lymphocytes: differential role for glycogen synthase kinase-3 in fibroblasts and T cells. Int Immunol 11, 317-323.

Vassart, G., Pardo, L., and Costagliola, S. (2004). A molecular dissection of the glycoprotein hormone receptors. Trends Biochem Sci 29, 119-126.

Vlaeminck-Guillem, V., Ho, S.C., Rodien, P., Vassart, G., and Costagliola, S. (2002). Activation of the cAMP pathway by the TSH receptor involves switching of the ectodomain from a tethered inverse agonist to an agonist. Mol Endocrinol 16, 736-746.

Wasif, N., and Ahmad, W. (2013). A novel nonsense mutation in RSPO4 gene underlies autosomal recessive congenital anonychia in a Pakistani family. Pediatr Dermatol 30, 139-141.

Wei, Q., Yokota, C., Semenov, M.V., Doble, B., Woodgett, J., and He, X. (2007). R-spondin1 is a high affinity ligand for LRP6 and induces LRP6 phosphorylation and beta-catenin signaling. J Biol Chem 282, 15903-15911.

Yoon, S.I., Kurnasov, O., Natarajan, V., Hong, M., Gudkov, A.V., Osterman, A.L., and Wilson, I.A. (2012). Structural basis of TLR5-flagellin recognition and signaling. Science 335, 859-864.

Zeng, H., Phang, T., Song, Y.S., Ji, I., and Ji, T.H. (2001). The role of the hinge region of the luteinizing hormone receptor in hormone interaction and signal generation. J Biol Chem 276,

EXTENDED EXPERIMENTAL PROCEDURES

Protein expression, purification and crystallization. All chemicals were purchased from Sigma-Aldrich (USA) unless specified. Ecto-LGR5 (residue 22-543 or 22-478) and Rspo1-Fu1Fu2 (residue 31-145) constructs were cloned into pUPE vectors (U-Protein Express BV, The Netherlands) carrying hexa-histidine tag. The proteins were produced recombinantly in N-acetylglucoaminyltransferase I deficient (GnTI-) HEK 293 cells that stably expressed Epstein-Barr virus Nuclear Antigen I (EBNA) (Durocher et al., 2000; Reeves et al., 2002). Cultures were harvested 6 days post transfection. Media containing the secreted proteins were concentrated and diafiltrated using Quickstand Hollow fiber system (GE Healthcare). Recombinant proteins were purified using Ni2+-sepharose FF resin, anion exchange MonoQ column and gel filtration Superdex 200 or 75 columns (GE Healthcare, USA). Samples were concentrated to 10-15 mg/ml and crystallized by hanging drop vapour diffusion method at 277 K or 291 K. Crystals of ecto-LGR5-Rspo1-Fu1Fu2 were obtained in several conditions: I, 0.3 M sodium malonate, 0.1 M N-(2-Acetamido)iminodiacetic acid (ADA) pH 6.5, 8% low molecular weight polyglutamic acid (PGA-LM); II, 1.2 M ammonium sulfate, 0.1 M sodium citrate pH 4.5; III, condition same as I; IV, 0.15 M ammonium sulfate, 0.1 M HEPES pH 7.0, 20% polyethylene glycol (PEG) 4000. Rspo1-Fu1Fu2 crystals were obtained in 1.6 M ammonium sulfate, 0.1 M morpholino ethanesulfonic acid (MES) pH 6.5. Crystals were harvested and flash-cooled in the presence of mother liquor supplemented with 20% ethylene glycol.

Analytical gel filtration and molecular weight analysis. Purified samples at 1 mg/ml concentration were injected into a Superdex 200 5/150 GL column (GE Healthcare, USA) connected to an Akta Micro (GE Healthcare, USA) coupled to a MiniDawn TREOS multi-angle laser light scattering (MALLS) detector (Wyatt Technology, USA) and a differential refractive index detector (RID-10A, Shimadzu, Japan). For the ecto-LGR5-Rspo1-Fu1Fu2 complex, a dn/dc value of 0.18 was calculated considering 5 glycosylation sites (4 sites on ecto-LGR5 plus one on Rspo1-Fu1Fu2). This value was used for the molecular weight determination from MALLS analysis using the ASTRA package (Wyatt Technology, USA).

Mutagenesis. Ecto-LGR5 and Rspo1-Fu1Fu2 mutants were generated using QuickChange mutagenesis kit (Stratagene) on TOPO clones inserted in cloning vectors (pCR-TOPO, Invitrogen). After confirmation of successful mutagenesis by DNA sequencing, the mutated genes were subcloned into pUPE vectors for mammalian expression. All mutants were expressed and purified using the same protocol used for their wild-type variants.

Data collection, structure determination and refinement. Diffraction data were collected at Swiss Light Source (SLS Villigen, Switzerland) and at the European Synchrotron Radiation Facility (ESRF Grenoble, France). Data were processed by MOSFLM (Leslie and Powell, 2007), XDS (Kabsch, 2010), EVAL15/SADABS (Schreurs et al., 2010) and AIMLESS (Evans, 2011). Resolution limits were determined by applying a cut-off based on the mean intensity correlation coefficient of half-datasets (CC1/2) approximately of 0.5 (Karplus and Diederichs, 2012). The structure of Rspo1-Fu1Fu2 was determined by experimental phasing. Crystals

were soaked overnight in mother liquor containing 5 mM $HoCl_3$. Initial phases were obtained using SHELX C/D/E packages (Sheldrick, 2008) and HKL2MAP (Pape and Schneider, 2004). Model building was performed by ARP/WARP (Langer et al., 2008) and completed using COOT (Emsley et al., 2010). The structure was refined using PHENIX (Adams et al., 2010) and two holmium ions at special positions were refined anisotropically. For the ecto-LGR5 - Rspo1-Fu1Fu2 complex, the structure was solved by molecular replacement (MR) using PHASER (McCoy et al., 2007). An initial poor MR solution obtained using search models generated from the structures of distant LGR5 homologs decorin (PDB ID 1XKU (Scott et al., 2004) and netrin (PDB ID 3ZYN (Seiradake et al., 2011). This solution was then improved by creating an LGR5 homology model created through manual superposition of separate LRR fragments from different protein structures identified using the LRRML database (Wei et al., 2008) on the initial search models. The structure of Rspo1-Fu1Fu2 was then used as additional search model in PHASER Initial structure refinement was performed by iterative cycles of manual building in COOT, multi-crystal averaging using DM (1994) and jelly-body refinement using REFMAC5 (Winn et al., 2003). Application of ensemble refinement (Burnley et al., 2012) at this stage yielded more interpretable electron density maps. The structures were then further refined using PHENIX (Adams et al., 2010). Non-crystallography symmetry (NCS) was imposed throughout refinement. Electron density maps obtained at the end of the structural refinement are shown in Supplementary Fig. 9. Molprobity (Chen et al., 2010) and PDB-CARE (Lutteke and von der Lieth, 2004) were used for structure validation. Structural analysis was performed using various softwares of the CCP4 suite (1994), EBI PISA (Krissinel and Henrick, 2007) and the DALI server (Holm and Rosenstrom, 2010). Figures were generated with PyMol (Schrodinger, 2010).

TOPFlash-assay. The TOPFlash-assay was essentially performed as described (Staal et al., 1999). HEK293T cells (ATCC) and routinely cultured in DMEM (Gibco) supplemented with 10% fetal bovine serum (FBS). Wnt3a-conditioned medium was prepared as described by Willert et al (Willert et al., 2003). Rspo1-conditioned medium was prepared by polyethylene imine (Polysciences) transfection of 107 HEK 293T cells in 15 cm culture dishes using 25 µg of pcDNA (Invitrogen)-based expression plasmid. After 18 hours medium was exchanged for Advanced DMEM/F12 (Gibco) and medium harvested 5 days later. 3'UTR-specific siRNA-induced knock down of LGR4 in HEK293T cells was performed as described previously (Huch et al., 2013). siRNA sequences specifically applied here were a mix of : gaaaguaaacuguggucaauu and gguaagaaacuccuaauuauu. Non-Targeting Pool siRNA was used as control. Transfection reagent Dharmafect-1 and siRNA sequences were all from Dharmacon/ Thermo Scientific.

Immunoprecipitation. M2 anti-FLAG agarose (Sigma) was used to immobilize Flag-tagged of ecto-LGR5 present in conditioned media derived from transiently transfected HEK 293T cells. Upon washing beads twice with PBS, coated beads were incubated with Rspo-conditioned media and subsequently washed three times with PBS. Bead-attached proteins were eluted with SDS-PAGE sample buffer and run under reducing conditions in a 15% polyacrylamide gel. Blotted proteins were tested for the presence of specific tags using anti-6xHIS/HRP conjugate (Abcam) or a FLAG-specific M2/HRP conjugate (Sigma). For visualization of the HRP enzyme,

ECL (Amersham) was used.

Crypt isolation and cell culture. Crypts were released from murine small intestine by incubation in PBS containing 2 mM EDTA for 30 min at 4 °C in a roller platform. Isolated crypts were counted and pelleted. A total of 200 crypts were mixed with 25 µl of Matrigel (BDBioscience) and plated in 48-well plates. After polymerization of Matrigel, 250 µl of crypt culture medium was added. Culture medium was prepared as described previously, by supplementing Advanced DMEM/F12 (Invitrogen) with penicillin/streptomycin, 10 mM HEPES, Glutamax, N2, B27 (all from Invitrogen), 1 µM N-acetylcysteine (Sigma) and the growth factors (EGF (50 ng/ml, Peprotech), Noggin (100 ng/ml, Peprotech) and 25 nM of Rspo1-Fu1Fu2 mutant protein as stated in the text. The entire medium was changed every 2-3 days for a total period of 7 days. All organoid pictures were taken using a Leica DMIL inverted microscope.

SUPPLEMENTAL REFERENCES

Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., et al. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr 66, 213-221.

Burnley, B.T., Afonine, P.V., Adams, P.D., and Gros, P. (2012). Modelling dynamics in protein crystal structures by ensemble refinement. elife 1, e00311.

Chen, V.B., Arendall, W.B., 3rd, Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., and Richardson, D.C. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr D Biol Crystallogr 66, 12-21.

Collaborative Computational Project, number 4 (1994). The CCP4 suite: programs for protein crystallography. Acta Crystallogr D Biol Crystallogr 50, 760-763.

Durocher, Y., Perret, S., Thibaudeau, E., Gaumond, M.H., Kamen, A., Stocco, R., and Abramovitz, M. (2000). A reporter gene assay for high-throughput screening of G-protein-coupled receptors stably or transiently expressed in HEK293 EBNA cells grown in suspension culture. Anal Biochem 284, 316-326.

Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. Acta Crystallographica Section D 66, 486-501.

Evans, P.R. (2011). An introduction to data reduction: space-group determination, scaling and intensity statistics. Acta Crystallogr D Biol Crystallogr 67, 282-292.

Holm, L., and Rosenstrom, P. (2010). Dali server: conservation mapping in 3D. Nucleic Acids Res 38, W545-549.

Huch, M., Dorrell, C., Boj, S.F., van Es, J.H., Li, V.S., van de Wetering, M., Sato, T., Hamer, K., Sasaki, N., Finegold, M.J., et al. (2013). In vitro expansion of single Lgr5+ liver stem cells induced by Wnt-driven regeneration. Nature 494, 247-250.

Kabsch, W. (2010). Xds. Acta Crystallogr D Biol Crystallogr 66, 125-132.

Karplus, P.A., and Diederichs, K. (2012). Linking crystallographic model and data quality. Science 336, 1030-1033.

Krissinel, E., and Henrick, K. (2007). Inference of macromolecular assemblies from crystalline

state. J Mol Biol 372, 774-797.
Langer, G., Cohen, S.X., Lamzin, V.S., and Perrakis, A. (2008). Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. Nat Protoc 3, 1171-1179.
Leslie, A.G.W., and Powell, H.R. (2007). Processing diffraction data with MOSFLM, Vol 245.
Lutteke, T., and von der Lieth, C.W. (2004). pdb-care (PDB carbohydrate residue check): a program to support annotation of complex carbohydrate structures in PDB files. BMC Bioinformatics 5, 69.
McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. J Appl Crystallogr 40, 658-674.
Pape, T., and Schneider, T.R. (2004). HKL2MAP: a graphical user interface for macromolecular phasing with SHELX programs. Journal of Applied Crystallography 37, 843-844.
Reeves, P.J., Callewaert, N., Contreras, R., and Khorana, H.G. (2002). Structure and function in rhodopsin: high-level expression of rhodopsin with restricted and homogeneous N-glycosylation by a tetracycline-inducible N-acetylglucosaminyltransferase I-negative HEK293S stable mammalian cell line. Proc Natl Acad Sci U S A 99, 13419-13424.
Schreurs, A.M.M., Xian, X., and Kroon-Batenburg, L.M.J. (2010). EVAL15: a diffraction data integration method based on ab initio predicted profiles. Journal of Applied Crystallography 43, 70-82.
Schrodinger, LLC (2010). The PyMOL Molecular Graphics System, Version 1.3r1.
Scott, P.G., McEwan, P.A., Dodd, C.M., Bergmann, E.M., Bishop, P.N., and Bella, J. (2004). Crystal structure of the dimeric protein core of decorin, the archetypal small leucine-rich repeat proteoglycan. Proc Natl Acad Sci U S A 101, 15633-15638.
Seiradake, E., Coles, C.H., Perestenko, P.V., Harlos, K., McIlhinney, R.A., Aricescu, A.R., and Jones, E.Y. (2011). Structural basis for cell surface patterning through NetrinG-NGL interactions. EMBO J 30, 4479-4488.
Sheldrick, G.M. (2008). A short history of SHELX. Acta Crystallogr A 64, 112-122.
Staal, F.J., Burgering, B.M., van de Wetering, M., and Clevers, H.C. (1999). Tcf-1-mediated transcription in T lymphocytes: differential role for glycogen synthase kinase-3 in fibroblasts and T cells. Int Immunol 11, 317-323.
Wei, T., Gong, J., Jamitzky, F., Heckl, W.M., Stark, R.W., and Rossle, S.C. (2008). LRRML: a conformational database and an XML description of leucine-rich repeats (LRRs). BMC Struct Biol 8, 47.
Willert, K., Brown, J.D., Danenberg, E., Duncan, A.W., Weissman, I.L., Reya, T., Yates, J.R., 3rd, and Nusse, R. (2003). Wnt proteins are lipid-modified and can act as stem cell growth factors. Nature 423, 448-452.
Winn, M.D., Murshudov, G.N., and Papiz, M.Z. (2003). Macromolecular TLS refinement in REFMAC at moderate resolutions. Methods Enzymol 374, 300-321.
3451-3458.
Zhang, M., Tong, K.P., Fremont, V., Chen, J., Narayan, P., Puett, D., Weintraub, B.D., and Szkudlinski, M.W. (2000). The extracellular domain suppresses constitutive activity of the transmembrane domain of the human TSH receptor: implications for hormone-receptor interaction and antagonist design. Endocrinology 141, 3514-3517.

SUPPLEMENTAL FIGURES

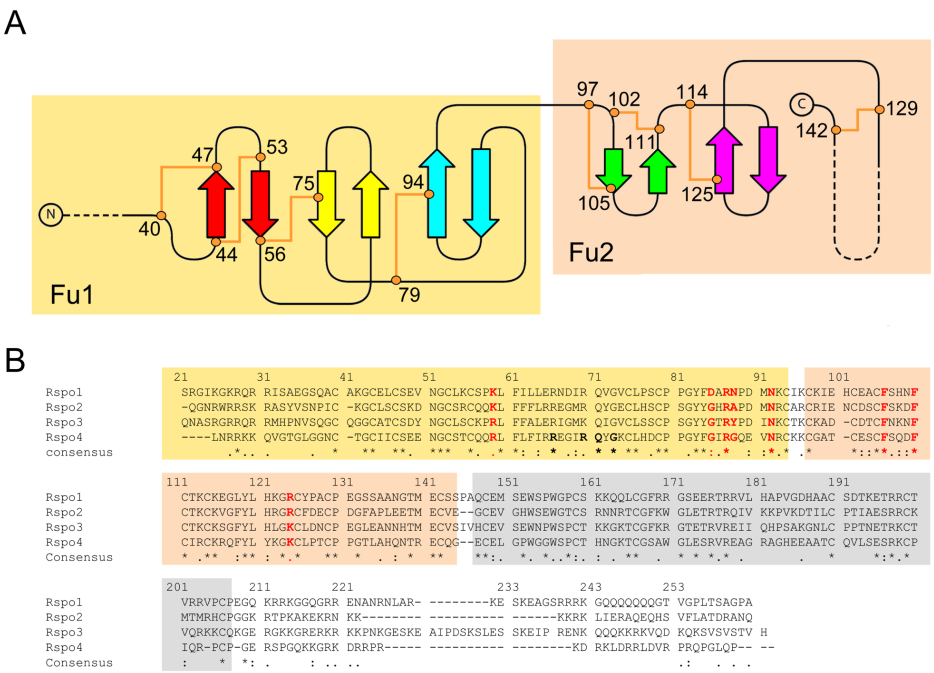


Fig S1. Structural features of Rspo1.
A, Schematic representation of the Rspo-Fu1Fu2 topology with disulphide bonds and β-strands indicated. Separate colours for the β-hairpins reflect the absence of a common β-sheet. Dashed line indicates residues that were not resolved in the electron density. Shaded area indicates domains Fu1 and Fu2.
B, Sequence alignment for Rspo1-4. Sequence numbering shown is according to the Uniprot entry for human Rspo1 Q2MKA7. Residues critical for interaction with ecto-LGR5 are given in bold red, Anonychia-related Rspo4 mutations in bold black. Grey shading indicates the Tsp domain.

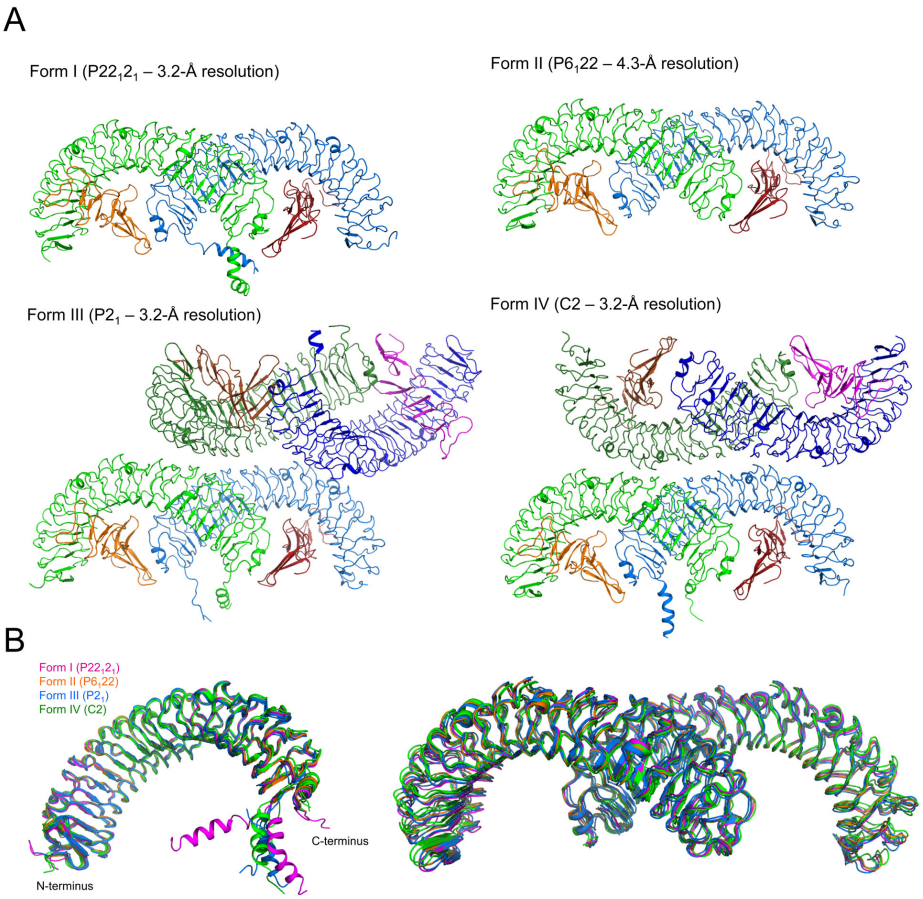


Fig S2. Comparison of four crystal structures of Rspo1-FuFu2 in complex with ecto-LGR5.
A, Cartoon representation of the asymmetric units of the ecto-LGR5 - Rspo1-Fu1Fu2 complexes formed in four different crystals forms I-IV (see Supplementary Table 1).
B, Superposition of the modelled LGR5 structures showing high flexibility in an extended loop of the C-cap region; this region is recognized by the antibodies that activate signaling in absence of the R-spondin ligands. Dimeric arrangements are shown on the right side; dimers superpose with all-atoms RMSDs between 0.6 and 1.2 Å.

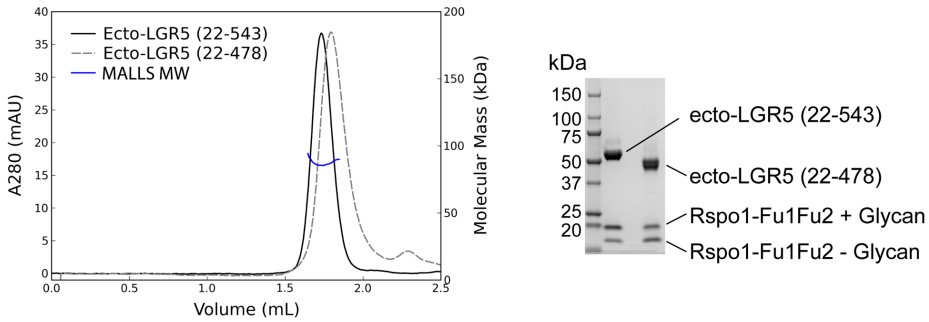


Fig S3. SEC-MALLS analysis of ecto-LGR5 and Rspo1-Fu1Fu2 complexes.
Gel-filtration traces of ecto-LGR5 residues 22-543 (black solid line) and residues 22-478 (grey dashed line) in complex with Rspo1-Fu1Fu2. The molecular weight calculated from MALLS analysis (blue dots) corresponds to a 1:1 heterodimer of 85 kDa. On the right, the SDS-PAGE analysis of elution peaks indicated the presence of both ecto-LGR5 and Rspo1-Fu1Fu2. Rspo1-Fu1Fu2 shows two bands, corresponding to a non-separable mixture of N-linked glycosylated and non-glycosylated samples.

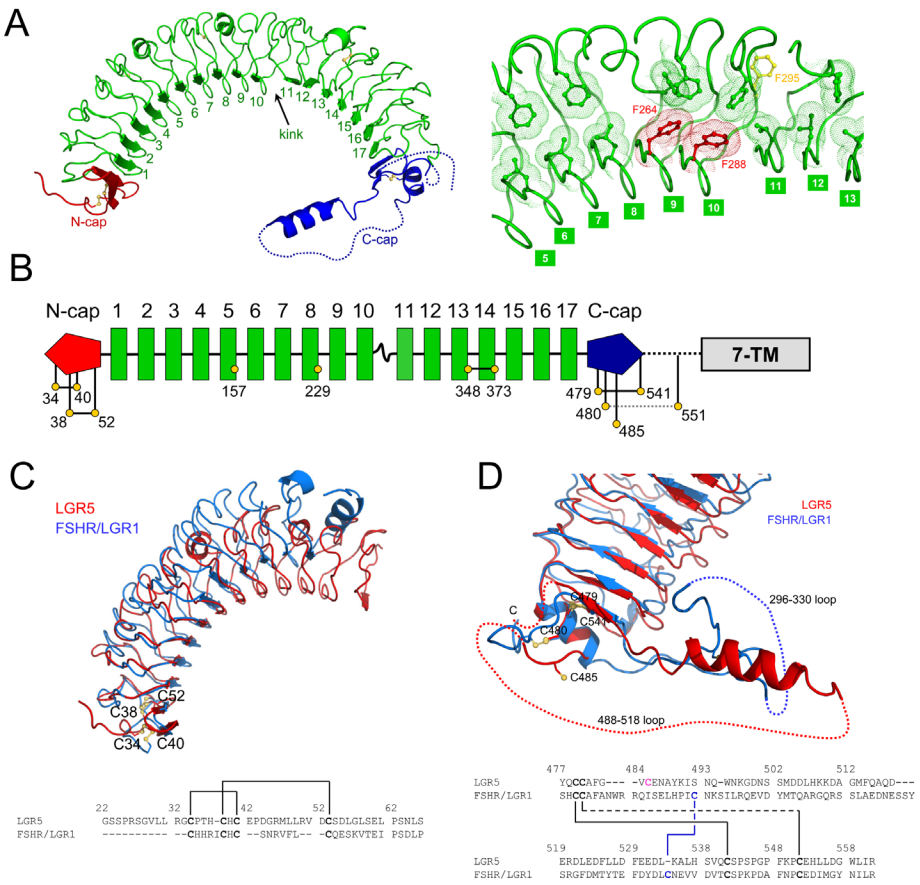


Fig S4. Structural features of ecto-LGR5.

A, (left) structure of the LGR5 ecto-domain indicating the N-terminal cap (red), the 17 LRR units (green) with kink in between LRR10-11 and the C-terminal cap (blue). Disulphide bonds are shown in yellow ball-and-stick representation. (right) Kink observed between LRR10-11. Canonical leucines, or small hydrophobic residues, and phenylanines responsible for the packing inside the LRR are shown in green ball-and-stick and their van der Waals radii are indicated as dotted spheres. Instead of small hydrophobic residues, LGR5 has Phe264 and Phe288, which are indicated in red. Phe264 (of LRR9) is conserved in LGR4, -5 and -6. Phe288 (of LRR10) is conserved only in LGR5 and -6, whereas LGR4 has a canonical leucine at this position. An additional phenylalanine, D, Overlay of the C-cap of FSHR/LGR1 (blue) and LGR5 (red) with missing residues indicated by dashed lines; and, sequence alignment of the C-cap of these proteins with observed disulphide bonds indicated by solid lines. Phe295, present in LRR10 points outwards (indicated in yellow).

B, Schematic representation of domain composition of LGR5 with cysteines, disulphide connections and C-terminal 7-TM region indicated. The Cys480-Cys551 disulphide bond (dashed line) is based on the crystal structure of FSH/LGR120.

C, FSHR/LGR1 (blue) superimposed on the N-cap and first LRR units of LGR5 (red).

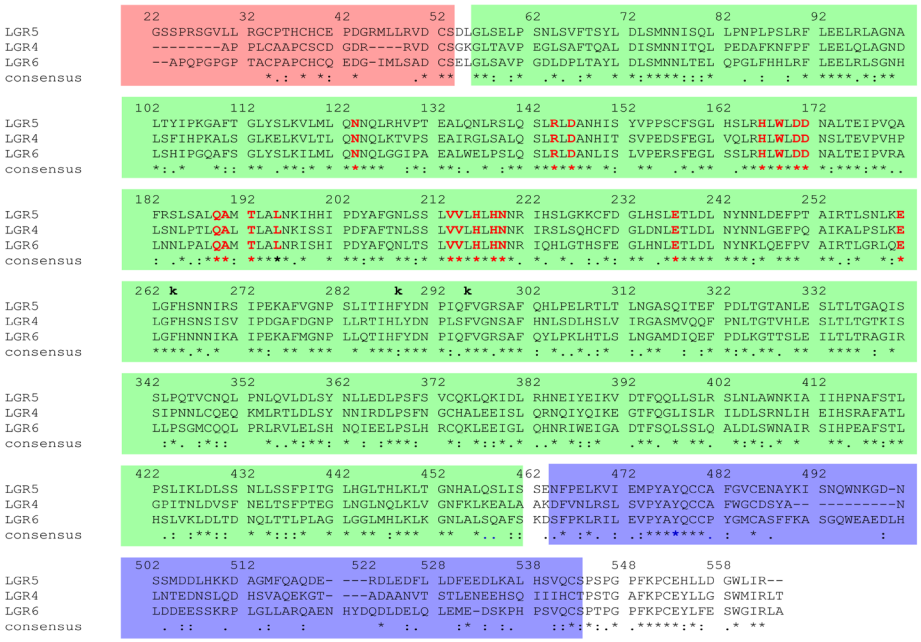


Fig S5. Sequence alignment of LGR4, LGR5 and LGR6 ecto-domains.
Sequence numbering according to the human LGR5 entry O75473 in the Uniprot database. Residues involved in the LGR5-Rspo1 interfaces are highlighted in bold red (cis). Bulky hydrophobic residues implicated in the LRR10-11 kink are indicated with a “k” above the sequence.

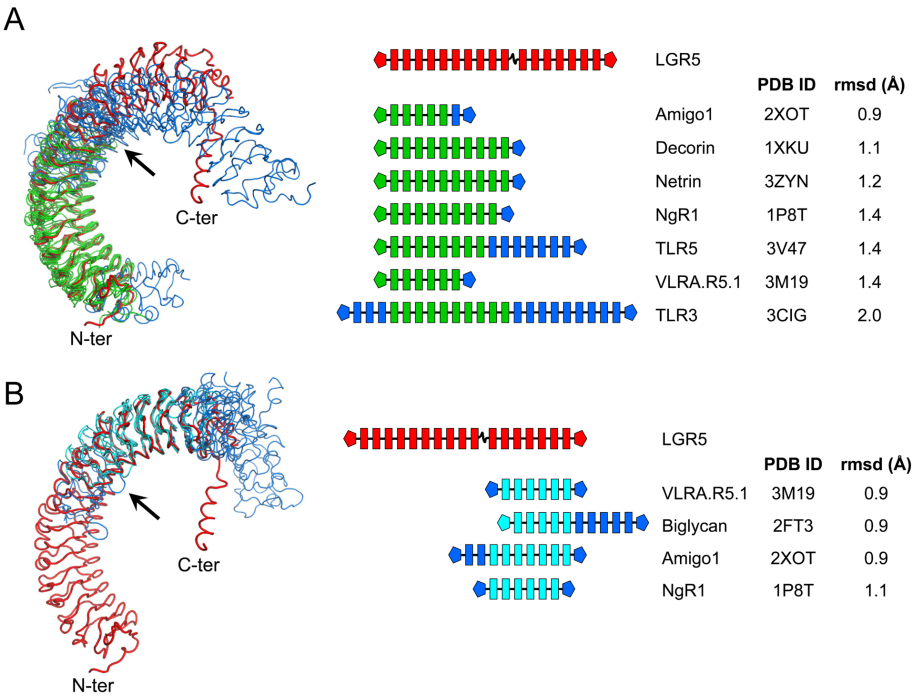


Fig S6. Comparison of LRR-curvatures.
DALI search31 for equivalent structural arrangements of LRR units resulted in two groups of proteins that fit the curvature of either the N-terminal or C-terminal end of the LRRs of LGR5.
A, Overlay of proteins that fit the N-terminal LRR region of LGR5 with schematic representation of the LRR units overlaid and root-mean-square coordinate differences.
B, idem dito for the C-terminal LRR region of LGR5.

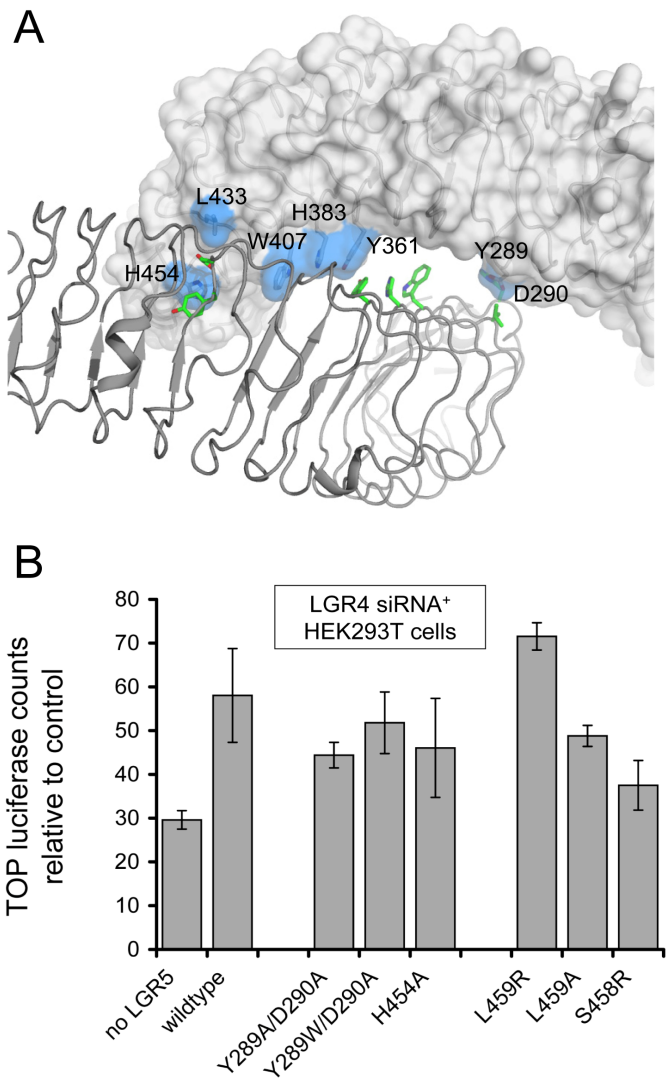


Fig S7. Putative signaling through dimerization.
A, Overview of the LGR5 dimerization interface. Residues involved in the interaction according to PISA analysis are shown as sticks and highlighted with shading on the LGR5 surface.
B, TOPFlash analysis of LGR5 mutants of the LGR5 dimeric interface (Y289A/D290A, Y289W/D290A, H454A) or the LGR5-Rspo1 trans interface (L459R, L459A, S458R). Results obtained by transfection-mediated introduction of 0.012 to 15 ng of wildtype and LGR5 variants (conditions as in Fig 2E). Error bars represent SD (n=3).

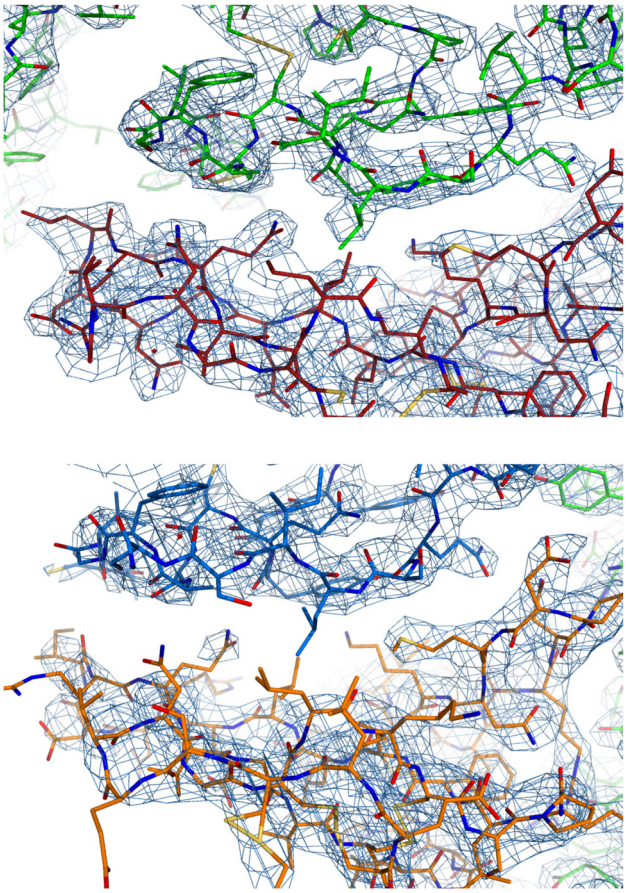


Fig S8. 'Trans' contact site between Rspo1-Fu1 and LGR5 C-cap.
Electron density drawn at 1σ at 'trans' contact sites observed in the two copies in crystal form I; coloured according to Fig. 1c. Contact residues are observed in density in the top panel. The bottom panel shows disordered side chains. This area is less resolved in all other copies in crystal forms (II to IV) indicating further disorder.

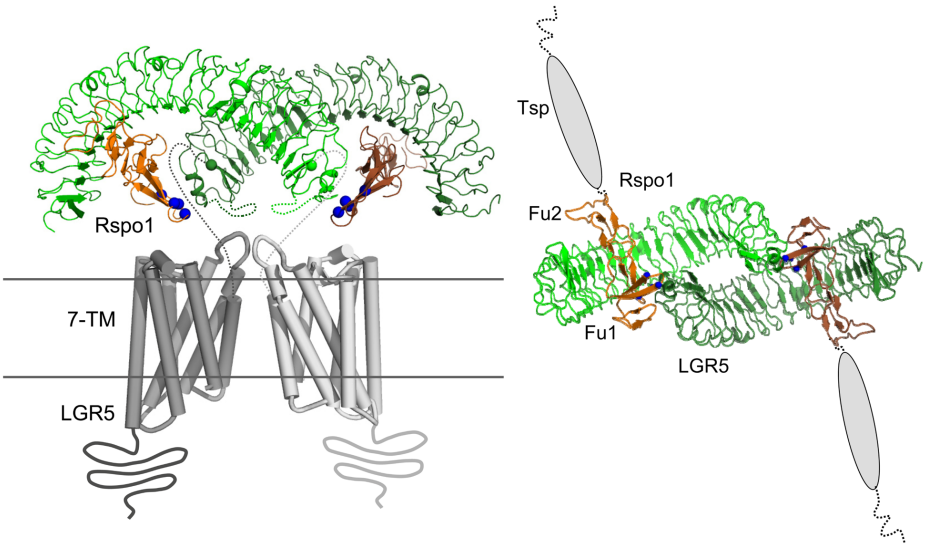


Fig S9. Hypothetical model of full-length Rspo1-4 binding to full-length LGR4-6.
Simplistic model of the overall LGR5 structure in complex with Rspo1. Shown is the observed dimeric crystal structure of ecto-LGR5 in complex with Rspo1-Fu1Fu2 combined with the structure of the 7-TM dimer of the β1 adrenergic G-protein coupled receptor32 (PDB: 4GPO). The dimeric 7-TM's (dark and light grey) are placed at the ecto-LGR5 dimer axis at a distance that may be covered by the missing 18 residues (grey dashed lines) between the two models. The 7-TM region of LGR5 was generated by homology modelling of LGR5 amino-acid sequence using HHPRED33. Based on the location of the termini we hypothetically assigned the left-sided 7-TM (dark grey) to the right-sided ecto-domain and vice versa for the 7-TM on the right side (light grey). Not shown are the long and flexible C-cap loops (indicated by short green dashed lines); these loops are likely to bind to the extracellular loops of the 7-TM's.

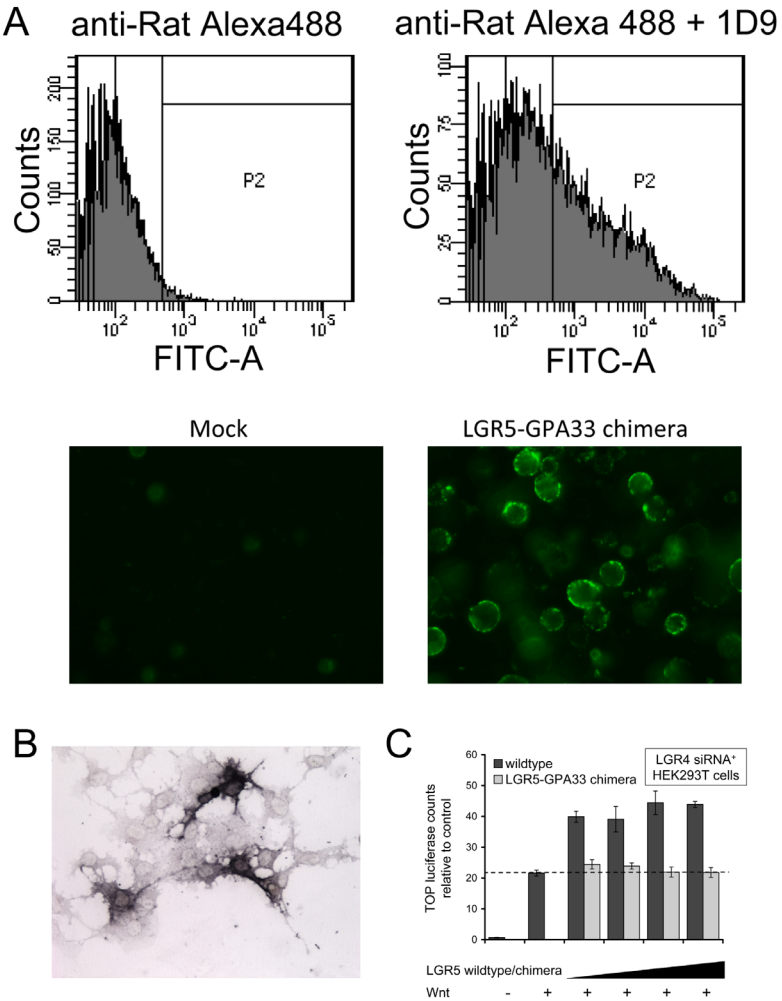


Fig S10. Analysis of the LGR5-GPA33 chimera

A, Cell surface expression for the LGR5-GPA33 chimera. The top panel shows the specific FACS detection of transfected LGR5 chimera in HEK293T cells using a combination of 1D9 and anti-rat antibody conjugated with Alexa488 (right) compared to the same experiment without using 1D9 antibody. The bottom panel shows the strong 1D9-Alexa488 fluorescence signal detected on the surface of HEK293T cells transfected with the chimeric receptor (right) compared to a mock transfection control (left).

B, Incubation of COS cells, transiently transfected with cDNA for LGR5-GPA33, with Rspo1-AP and alkaline phosphatase substrate, shows binding of Rspondin to this chimeric receptor.

C, TOPFlash assay performed using the LGR5-GPA33 chimera shows reduced signaling levels in absence of the LGR5 7-TM region. Error bars represent SD (n=3).

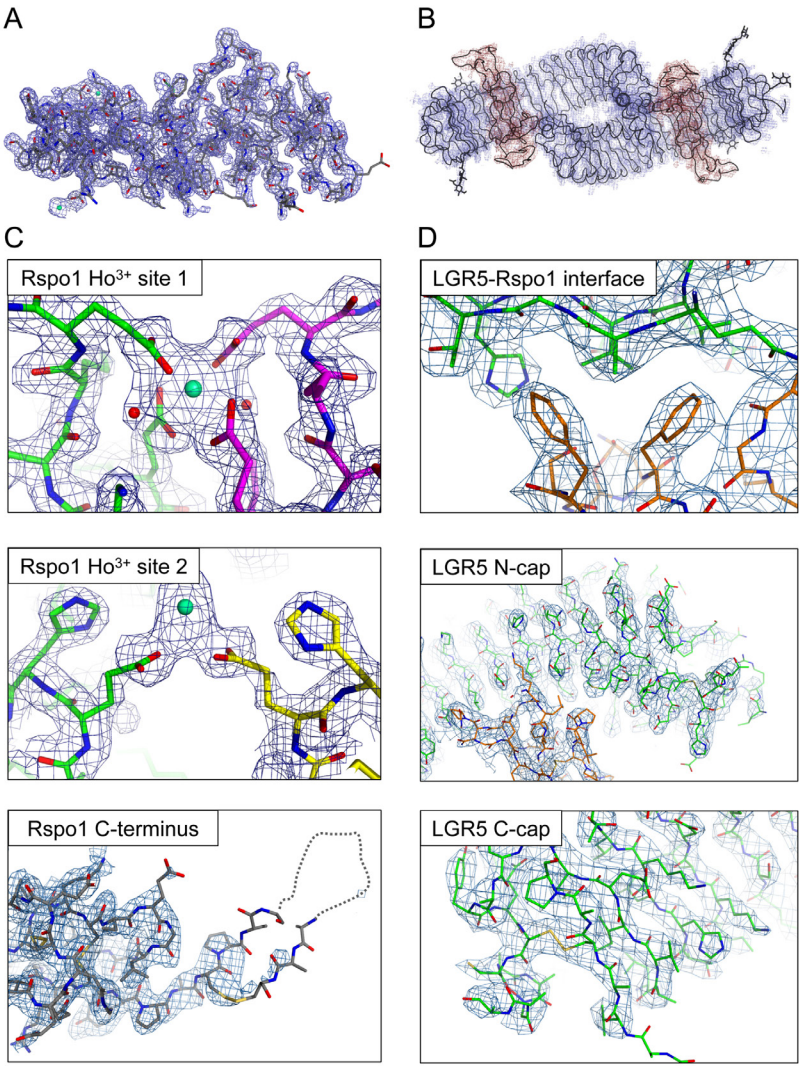


Fig S11, Electron density maps of Rspo1 and Rspo1-LGR5 complex.
Overall 2mFo-DFc electron density maps contoured at 1σ for
A, Rspo1-Fu1Fu2 and
B, The complex of Rspo1-Fu1Fu2 with ecto-LGR5.
C, Zoom-in of electron density for Rspo1 showing the two Holmium sites that stabilize the protein packing and the disorder observed for the C-terminal residues of Rspo1-Fu1Fu2.
D, Zoom-in of the electron density of LGR5 showing the main interaction interface between Rspo1 and ecto-LGR5 and both the N- and C-cap of the LRR region of LGR5.

Table S1, related to Fig 1. Crystallographic statistics for data collection, structure solution and refinement

	LGR5-Rspo Form I	LGR5-Rspo Form II	LGR5-Rspo Form III	LGR5-Rspo Form IV	Rspo1 Native	Rspo1 Ho^{3+} Soak
Data Collection[a]						
X-ray source	SLS X06SA	SLS X06SA	ESRF – ID23 EH1	SLS X06SA	SLS X06SA	ESRF – ID23 EH1
Processing programs	XDS/AIMLESS	XDS/AIMLESS	EVAL15/SADABS	MOSFLM/AIMLESS	XDS/AIMLESS	XDS/AIMLESS
Space group	$P22_12_1$	$P6_122$	$P2_1$	C2	I222	I222
Cell parameters	a = 85.2 Å; α = 90° b = 143.8 Å; β = 90° c = 167.3 Å; γ = 90°	a = 131.1 Å; α = 90.0° b = 131.1 Å; β = 90.0° c = 531.8 Å; γ = 90.0°	a = 120.4 Å; α = 90.0° b = 110.3 Å; β = 109.2° c = 130.6 Å; γ = 90.0°	a = 495.1 Å; α = 90.0° b = 65.2 Å; β = 91.2° c = 104.9 Å; γ = 90.0°	a = 37.2 Å; α = 90° b = 63.9 Å; β = 90° c = 92.5 Å; γ = 90°	a = 35.6 Å; α = 90° b = 64.6 Å; β = 90° c = 92.4 Å; γ = 90°
Wavelength (Å)	1.00	1.00	1.00	1.00	0.88	1.535
Resolution (Å)	47.9 – 3.20 (3.36-3.20)	48.3 – 4.31 (4.72-4.31)	28.9 – 3.20 (3.26-3.20)	97.2 – 3.20 (3.30-3.20)	34.4 – 2.20 (2.28-2.20)	33.2 – 2.00 (2.05-2.00
Unique reflections	34548 (4532)	19429 (4376)	52333 (2632)	52878 (4560)	5873 (560)	7500 (549)
CC1/2[b]	0.99 (0.57)	0.99 (0.68)	0.99 (0.59)	0.93 (0.36)	0.99 (0.84)	0.99 (0.57)
Redundancy	3.7 (3.8)	7.4 (6.3)	2.6 (2.7)	2.8 (2.8)	7.3 (7.5)	10.0 (10.3)
I/σ(I)	7.20 (1.40)	12.4 (2.3)	7.39 (1.21)	7.10 (0.90)	11.1 (2.4)	11.2 (2.2)
Completeness (%)	99.7 (99.9)	99.2 (98.0)	97.4 (98.0)	99.0 (99.1)	100.0 (100.0)	99.7 (100.0)
R_{sym}[c]	0.152 (0.911)	0.087 (0.812)	0.084 (0.768)	0.077 (1.003)	0.105 (0.903)	0.120 (1.199)
R_{sym}[c] (all I^+ and I^-)						0.158 (1.229)
Anomalous Compl.						97.7 (100.0)
Anomalous Mult.						5.2 (5.4)

[a] Values in parentheses are for reflections in the highest resolution shell.

[b] Resolution limits were determined by applying a cut-off based on the mean intensity correlation coefficient of half-datasets (CC1/2) approximately of 0.5 (Karplus and Diederichs, 2012).

[c] $R_{sym} = \Sigma \mid I - \langle I \rangle \mid / \Sigma I$, where I is the observed intensity for a reflection and <I> is the average intensity obtained from multiple observations of symmetry-related reflections. [d] R_{free} values are calculated based on 5% randomly selected reflections.

	LGR5-Rspo Form I	LGR5-Rspo Form II	LGR5-Rspo Form III	LGR5-Rspo Form IV	Rspo1 Native	Rspo1 Ho^{3+} Soak
Structure solution						
Method	MR	MR	MR	MR	MR	SAD
Processing programs	PHASER	PHASER	PHASER	PHASER	PHASER	SHELXC/D/E
Molecules per ASU	2	2	4	4	1	1
Heavy atom sites	-	-	-	-	-	2 (Ho^{3+})
Refinement						
R_{work}/R_{free} [d]	0.231/0.258	0.243/0.268	0.246/0.290	0.236/0.288	0.212/0.250	0.228/0.260
Average B-factors ($Å^2$	67.9	213.4	126.6	133.1	56.4	49.4
Number of atoms:	9318	8833	17912	17955	818	738
Protein	9181	8738	17744	17782	772	712
Ligands (including glycans)	137	95	168	173	26	7
Waters	0	0	0	0	20	19
Structure quality						
Molprobity score	2.74	3.28	3.13	2.84	3.28	2.07
RMS bond lengths (Å)	0.003	0.008	0.005	0.004	0.021	0.005
RMS bond angles (°)	0.90	1.51	1.15	1.23	2.12	0.83
Ramachandran stats						
Favored (%)	89	84	82	88	89	96
allowed (%)	10	13	15	10	7	3
outliers (%)	1	3	3	2	4	1

CHAPTER 3

Structures of Wnt-Antagonist ZNRF3 and Its Complex with R-spondin 1 and Implications for Signaling

Weng Chuan Peng,[1] Wim de Lau,[2] Pramod K Madoori,[1] Federico Forneris,[1] Joke C.M. Granneman,[1] Hans Clevers,[2,*] and Piet Gros[1,*]

PLoS One. 2013; 8(12): e83110

[1]Crystal and Structural Chemistry, Bijvoet Center for Biomolecular Research, Department of Chemistry, Faculty of Science, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands

[2]Hubrecht Institute, Royal Netherlands Academy of Arts and Sciences and University Medical Center Utrecht, Uppsalalaan 8, 3584 CT Utrecht, The Netherlands

ABSTRACT

Zinc RING finger 3 (ZNRF3) and its homolog RING finger 43 (RNF43) antagonize Wnt signaling in adult stem cells by ubiquitinating Frizzled receptors (FZD), which leads to endocytosis of the Wnt receptor. Conversely, binding of ZNRF3/RNF43 to LGR4-6 – R-spondin blocks Frizzled ubiquitination and enhances Wnt signaling. Here, we present crystal structures of the ZNRF3 ectodomain and its complex with R-spondin 1 (RSPO1). ZNRF3 binds RSPO1 and LGR5-RSPO1 with micromolar affinity via RSPO1 furin-like 1 (Fu1) domain. Anonychia-related mutations in RSPO4 support the importance of the observed interface. The ZNRF3-RSPO1 structure resembles that of LGR5-RSPO1-RNF43, though Fu2 of RSPO1 is variably oriented. The ZNRF3-binding site overlaps with trans-interactions observed in 2:2 LGR5-RSPO1 complexes, thus binding of ZNRF3/RNF43 would disrupt such an arrangement. Sequence conservation suggests a single ligand-binding site on ZNRF3, consistent with the proposed competing binding role of ZNRF3/RNF43 in Wnt signaling.

INTRODUCTION

Zinc RING finger 3 (ZNRF3) and its homolog RING finger 43 (RNF43) are trans-membrane E3 ubiquitin ligases that negatively regulate Wnt signaling [1,2]. Mutations in ZNRF3 or RNF43 have been linked to gastric adenocarcinoma [3], pancreatic ductal adenocarcinoma [4], liver fluke-associated cholangiocarcinoma [5] and mucinous ovarian tumors [6]. ZNRF3 and RNF43 contain an extracellular N-terminal protease-associated (PA) domain, a single pass trans-membrane helix and an intracellular C-terminal RING domain with E3 ligase activity [2]. Interaction of ZNRF3 or RNF43 with complexes of frizzled receptors (FZD) and low-density lipoprotein receptor-related protein (LRP) 5/6 leads to Frizzled ubiquitination and endocytosis of the heterodimeric receptors, thereby reducing the capacity of Wnt-driven signal transduction [1,2].

R-spondins 1-4 (RSPO1-4) are stem cell growth factors that bind leucine-rich repeat G-protein coupled receptors 4-6 (LGR4-6) on adult stem cells [7-10], such as in the intestine and colon [11], hair follicles [12], stomach [13], kidney [14], liver [15] and mammary glands [16]. LGR4-6 – R-spondin complexes potentiate Wnt signaling; however, the underlying mechanism is not completely resolved. It was recently reported that LGR4-RSPO1 complex interacts with ZNRF3 and facilitates the removal of ZNRF3 from the membrane, thereby indirectly increasing the number of Wnt receptor/co-receptor complexes on the cell surface [2]. Carmon et al., in contrast, observed that LGR5 forms a supercomplex with FZD-LRP5/6 upon stimulation with R-spondin 1 and Wnt3a and increases the rate of LRP6-FZD receptors internalization and degradation [17]; this model would contradict the role of LGR4/5-RSPO1 in increasing the number of Wnt receptors on the cell surface.

Recent crystal structures [18-21] showed that RSPO1-4 bind LGR4-6 at the concave surface of the extended leucine-rich repeat (LRR) region of the LGR ectodomain. The ‘phenylalanine clamp’ of RSPO furin-like (Fu) 2 domain is critical for binding to the hydrophobic patch on

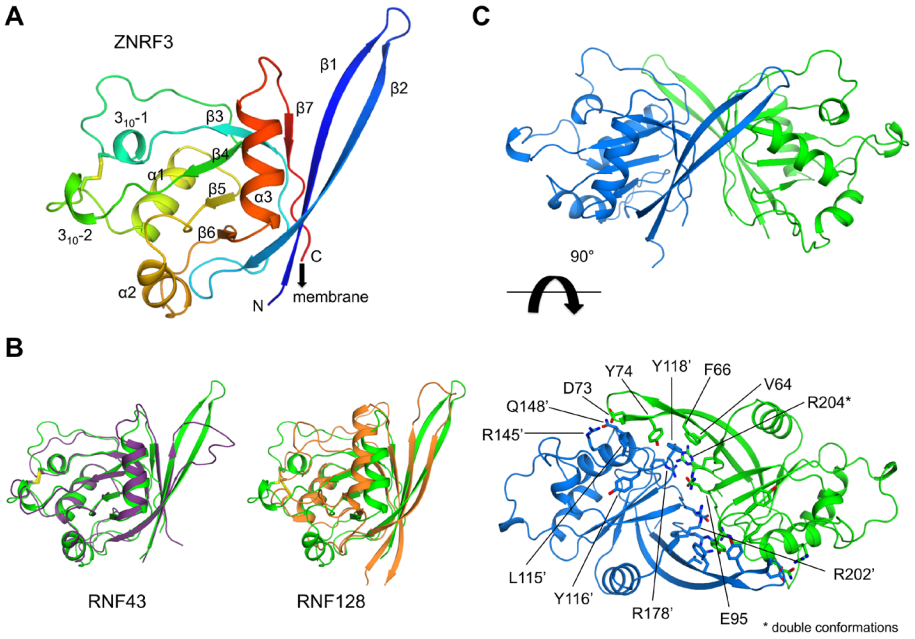


Fig 1. Crystal structure of the ectodomain of ZNRF3.
A, Cα trace of ZNRF3 ectodomain coloured from N- to C-terminus (blue through red) and structural elements indicated. The arrow indicates the connection to the single trans-membrane pass.
B, Overlay of the Cα trace of ZNRF3 (green) with RNF43 (purple) and RNF128 (orange).
C, Two perpendicular views of the dimeric arrangement of ZNRF3 observed in the crystal, with residues at the interface indicated.

LRR3-9. In addition, we observed 2:2 LGR5-RSPO1 complexes in four crystal forms [18]. However, such quaternary arrangement was not observed in LGR4-RSPO1 structure [20,21]. Here, we present crystal structures of the ectodomain of ZNRF3 and its complex with RSPO1. RSPO1 binds ZNRF3 primarily through its Fu1 domain and Fu2 exhibits domain flexibility in the absence of LGR4/5. Mutations in RSPO4 implicated in congenital anonychia [22] correspond to RSPO1 residues that mediate interactions with ZNRF3. Furthermore, superposition of the ZNRF3-RSPO1 with the LGR5-RSPO1 structures shows that ZNRF3 overlaps with the dimeric partner LGR5 in the 2:2 LGR5-RSPO1 complexes. Thus, interaction of ZNRF3 with LGR5-RSPO1 would block or disrupt this quaternary arrangement.

RESULTS AND DISCUSSION

Structure of ZNRF3 protease-associated domain

The ectodomain of ZNRF3 was transiently expressed in HEK293 cells. The protein was purified by immobilized metal ion affinity chromatography and gel-filtration. Size-exclusion

chromatography and multi-angle laser-light scattering indicated that ZNRF3 exists as monomer in solution (data not shown). Purified protein was crystallized and crystals exhibited space group $P2_1$ with cell dimensions a = 35.7 Å, b = 73.5 Å and c = 58.6 Å and β = 97.5°, contained two molecules per asymmetric unit and diffracted to 1.5 Å resolution. Crystallographic data and refinement statistics are given in Table 1; electron density is shown in Figure S1A.

The ZNRF3 ectodomain adopts a typical PA-fold, previously found in e.g. subtilases, transferrin receptors and vacuolar sorting receptors [23,24]. The central core of the molecule is formed by a parallel β-sheet, consisting of strands β3-β4-β5-β6 surrounded by three α-helices and two short 3_{10}-helices (Figure 1A). A disulphide bond Cys107-Cys136, which is present in both ZNRF3 and RNF43 connects the loop regions containing the first and second 3_{10}-helices. The N-terminal and C-terminal residues of the ectodomain form an anti-parallel β-sheet, β2-β1-β7, that packs against helix α3. Consequently, the termini formed by extensions of strands β1 and β7 are close together in space. A linker of approx. 10 residues connects the C-terminus of the ectodomain (Figure 1A) to the trans-membrane helix in the lipid bilayer.

The overall fold of the ZNRF3 ectodomain is similar to the ectodomain of its homolog RNF43, which is in agreement with a sequence identity of 37% between the two ectodomains (Figure 1B and S2). The Cα-positions of ZNRF3 and RNF43 (PDB code 4KNG) [19], can be superimposed on each other with a root-mean-square deviation (rmsd) of 0.75 Å. The largest structural difference between ZNRF3 and RNF43 is observed for the N-terminal strands β1 and β2. In ZNRF3, strands β1 and β2 form an extended β-hairpin 'flap', while RNF43 is three residues shorter and displays a flexible loop in this region. Furthermore, DALI search identified RNF128 (also known as GRAIL) to have a related fold (Figure 1B), despite low sequence identity of 15% and a rmsd of 5.3 Å compared to ZNRF3, suggesting that the ectodomains of the Goliath family E3 ligases (such as RNF13, RNF130, RNF133, RNF148, RNF149, RNF150, RNF167 and RNF204) [25] have related folds for ligand recognition.

In the crystal structure of ZNRF3, we observe two molecules in the asymmetric unit (Figure 1C). The two molecules pack together making an extensive interface burying over ~1,000 $Å^2$ surface area. The extended β1-β2 flaps fold over the other monomer and provide small hydrophobic and aromatic interaction clusters with Val64, Phe66, Gly72 and Tyr74 interacting with Leu115' and Y118' preceding strand β4 and Gly150' preceding β5 (where prime indicates residues from the opposing dimer), on both sides. Furthermore, a charged and H-bonded network is observed, which includes: Asp73-Arg145', Tyr74-Tyr116', Tyr74-Gln148', Glu95-Glu95', Glu95-Arg178' and Glu95-Arg202'. In addition, we observe stacking of guanidinium groups of Arg178-Arg204'. Both C-termini of the dimer point to the same direction, making such arrangement plausible on the membrane. However, at present it is not clear if such arrangement is physiologically relevant.

Structure of ZNRF3-RSPO1: RSPO1 binds to ZNRF3 through Fu1 domain

We co-crystallized ZNRF3 ectodomain with RSPO1 consisting of the Fu1 and Fu2 domains.

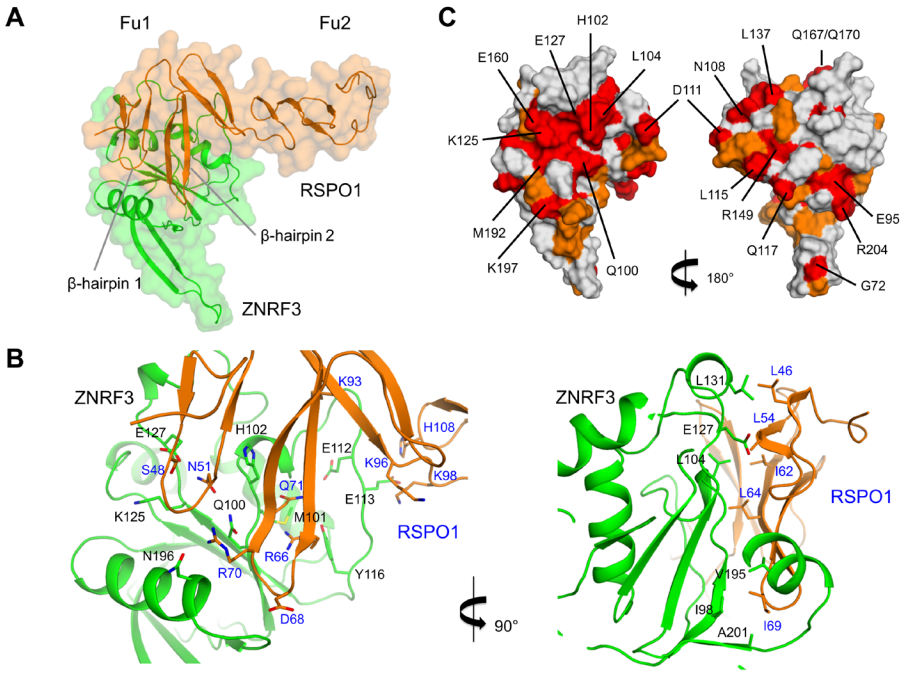


Fig 2. Crystal structure of the ZNRF3-RSPO1 complex.
A, Cα trace with transparent surface representation of ZNRF3 (green) and RSPO1 (orange). ZNRF3 makes contacts to the β-hairpins 1-2 of the Fu1 domain of ZNRF3.
B, Two views of the binding sites with interface residues indicated.
C, Identical (red) and conserved (orange) residues between ZNRF3 and RNF43 are shown in surface representation; two views of ZNRF3 are shown. Identical residues are labeled.

RSPO1 was expressed and purified as described previously [18]. Diffraction data were collected up to 2.8 Å resolution from crystals with space group P1 and cell dimension a = 51.7 Å, b = 80.2 Å, c = 83.0 Å and α = 66.3°, β = 81.4°, γ = 80.7° (see Table 1 for crystallographic data and refinement statistics; and Figure S1B for electron density).

The ZNRF3-RSPO1 complex reveals an extensive interface, burying ~1,200 Å^2, between the two molecules (Figure 2A,B). ZNRF3 forms a large pocket on the side of the molecule opposite to the C-terminus, which therefore likely faces away from the membrane surface. This binding platform is formed by several structural elements of ZNRF3, involving residues from strand β3, loop β3-β4 containing the first 3_{10}-helix, loop β4-α1 and loop α3-β7. Apart from some side-chain rearrangements, no major conformational change is observed in ZNRF3 upon binding to RSPO1.

The interaction site on RSPO1 is formed by the β-hairpins 1 and 2 of the Fu1 domain, which form an extended face that contacts ZNRF3 (Figure 2A). An extensive network of hydrogen

bonds and salt bridges mediates the ZNRF3-RSPO1 interaction. The core of the ZNRF3 interface includes residues Gln100, His102, Lys125 and Glu127, which have identical residues in RNF43, as well as residues Met101, Tyr116 and Asn196 that are not conserved in RNF43. These residues interact with the backbones or side chains of residues Ser48, Asn51, Cys53, Arg66, Arg70 and Gln71 of RSPO1 (Figure 2B), which are identical or strongly conserved (Arg70 is Lys in RSPO3) among RSPO1-4. Upon binding to ZNRF3, the tip of β-hairpin 2 (residues ^{67}NDIR^{70}) becomes ordered, whereas this region was highly flexible in the unbound RSPO1 [18], indicating possible molecular plasticity for binding diverse ligands. The hydrophobic side-chain of Ile69 (with residues Met, Met and Ile at the equivalent position in RSPO2-4, respectively), points into a hydrophobic pocket formed by ZNRF3 and make contacts Ile98 (RNF43: Leu), Val195 (Val) and Ala201 (Ala). Residues Leu46 (RSPO2-4: Ser, Thr, Ile), Leu54 (Ser, Leu, Ser), Ile62 (Phe, Phe, Leu) and Leu64 (Leu, Leu, Ile) lie on one side of β-hairpins 1 and 2 of RSPO1 and their side chains make contacts with Leu104 (RNF43: Leu), Gly105 (Tyr), Glu127 (Glu) and Leu131 (Arg) of ZNRF3 (Figure 2B). However, the residues making hydrophobic interactions are not strictly conserved in other R-spondins (RSPO 2-4). A short stretch of polar and negatively charged residues ^{108}NNNDEED^{114}, on ZNRF3 faces residues Lys93, Lys96, Lys98 and His108 on the hinge region of RSPO1. The charge interactions may contribute to long-range attraction between the molecules. However, large B-factor values for ZNRF3 residues located in this region indicate that these contacts are less well defined in the crystal structure.

A conserved binding platform in ZNRF3/RNF43 and RSPO1-4

The mode of ZNRF3-RSPO1 interaction is consistent with electrostatic interactions observed in RNF43-RSPO1 [19]. The majority of the identical surface-exposed residues cluster on the RSPO1-binding site to form an extended binding platform (Figure 2C and S3A). On the contrary, on the opposite side of the molecule, there are a few scattered identical exposed residues, such as Glu95, Leu115, Gln117 and Arg204 (with neighbouring Arg204 and Glu95 forming a salt bridge); whereas Gly72 and Arg149 expose main-chain atoms only. The presence of one predominant, evolutionary conserved binding platform would indicate that ZNRF3 and RNF43 possibly bind ligands such as Frizzled and RSPO1 at the same or overlapping site.

On RSPO1, the residues involved in binding ZNRF3, i.e. Ser48, Asn51, Arg66, Arg70 and Gln71, are identical among RSPO1-4 (except for Arg70, which is a lysine in RSPO3). Hence, ZNRF3/RNF43 should be able to bind promiscuously to all R-spondins.

Structural flexibility of RSPO1: a hinge region between Fu1 and Fu2 domain

Four copies of ZNRF3-RSPO1 are present in the asymmetric unit (Figure S3B), which are arranged as a dimer of dimers. The dimeric arrangement observed in the structure of ZNRF3 is conserved in the crystal structure of ZNRF3-RSPO1; the two ZNRF3-RSPO1 dimers contact each other sideways through H-bond interactions made by the β1-β2 flaps of ZNRF3. The interactions between ZNRF3 and the Fu1 domain of RSPO1 are identical among the four copies of the complex. Differences, however, are observed with respect to Fu1-Fu2 orientations

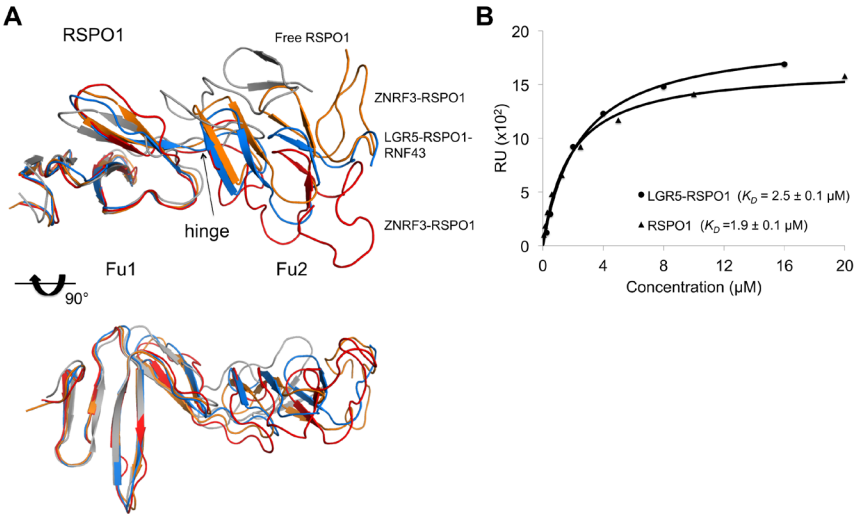


Fig 3. Flexible hinge in RSPO1 and binding of ZNRF3 to RSPO1 and LGR5-RSPO1.
A, Overlay of four representative RSPO1 structures in two orientations with unbound, or 'free', RSPO1 (grey; PDB code 4BSO), RSPO1 in LGR5-RSPO1-RNF43 complex (blue, PDB code 4KNG) and RSPO1 in complex with ZNRF3 (orange and red).
B, Representative SPR dose-response curve used to determine equilibrium binding affinity of LGR5-RSPO1 or RSPO1 to ZNRF3, as described in Material and Methods. Standard deviations are calculated from four experiments.

(Figure 3A). In two copies of RSPO1 (denoted chains F and H) the Fu2 domains are less well packed and display higher B-factors than the other copies (chains E and G; see Figure S3B); these two sets differ by ~20° in Fu1-Fu2 domain orientations. An overlay of RSPO1, RSPO1-ZNRF3 and LGR5-RSPO1-RNF43 structures, reveals a range of Fu1-Fu2 domain orientations with a hinge around residue Lys98. LGR4/5 make interactions with RSPO1 through both Fu1 and Fu2 domains and the LGR4/5-RSPO1 complexes show similar Fu1-Fu2 conformations (PDB codes 4BSR, 4KT1, 4LI2). Binding of RNF43 to LGR5-RSPO1 complex (PDB code 4KNG) does not induce any further conformational change. Apparently, RSPO1 exhibits internal flexibility with a hinge between Fu1 and Fu2; and, this flexibility does not affect ZNRF3 binding, while binding to LGR4/5 straightens the arrangement of the Fu domains.

Binding studies and the role of LGR5 in interactions with ZNRF3

The RSPO1 and ZNRF3 fragments could not be co-purified by size-exclusion chromatography, indicative of weak binding between RSPO1 and the ZNRF3 ectodomain. To determine whether binding of RSPO1 to ZNRF3 is enhanced by the receptor LGR5, we performed surface-plasmon resonance (SPR) binding studies. LGR5-RSPO1 and RSPO1 bind ZNRF3 with KD of 2.5 ± 0.1 μM and 1.9 ± 0.1 μM, respectively (Figure 3B). Previously, Chen et al. determined by isothermal titration calorimetry a KD of 7-10 μM for RSPO1-RNF43 and observed a 10-fold increase

in binding affinity (0.5-1.0 μM) in the presence of LGR5 [19]. Superposition of the ZNRF3-RSPO1 and LGR5-RSPO1-RNF43 complexes (Figure S3C) indicates that no contacts are likely between LGR5 and ZNRF3 in LGR5-RSPO1-ZNRF3 either. This observation is consistent with the observed similar binding affinities of RSPO1 and LGR5-RSPO1 to ZNRF3. The structural data do not explain different affinities for binding to RSPO1 and LGR5/RSPO1 as observed for the RNF43. Under physiological conditions, LGR5 may function to localize R-spondins on the membrane. Only nanomolar concentrations of R-spondin are required for LGR4-6 binding, Wnt signaling activity and stem-cell driven intestinal organoid growth [7,18]. ZNRF3 functions to ubiquitinate Frizzled receptors, and ZNRF3 itself is targeted by R-spondins for removal from surface [2]. The weak binding affinity observed maybe required for balancing these two events, so that ZNRF3 can exist in equilibrium between RSPO1 and FZD. Previously, Hao et al. detected interaction of ZNRF3 and FZD8 by immuno-precipitation [2]. We tested the binding of ZNRF3 to FZD8 cysteine-rich domain by SPR, the domain responsible for binding to Wnt, but did not observe any binding (data not shown). Hence, ZNRF3 possibly binds to FZD8 outside the cysteine-rich domain or additional factors like Wnt are needed to establish ZNRF3-FZD8 binding.

ZNRF3-RSPO1 interface coincides with LGR5-RSPO1 'trans' interfaces

Recently, several structures of LGR4/5-RSPO1 complexes were reported [18-21]. These structures were fully consistent with respect to the primary LGR4/5-RSPO1 binding site and interactions, but the structures differed in quaternary arrangements. Whereas Wang et al. and Xu et al. [20,21] observed a 1:1 LGR4-RSPO1 complex, with possible side-to-side contacts between complex in crystal contact, we observed a 2:2 LGR5-RSPO1 complex in multiple crystal forms, where the LRR11-17 repeats are twisted around each other [18]. In the LGR5-RSPO1 dimeric structures, RSPO1 contacts the second copy of LGR5 via the 'trans' interface. Specifically, the 'trans' interface is formed by Fu1 domain of RSPO1 and C-cap of LGR5. This interface coincides with ZNRF3-RSPO1 interface, also observed in the structure of RNF43 bound to LGR5-RSPO1 (Figure 4). Binding of ZNRF3/RNF43 to LGR5-RSPO1 would therefore disrupt the 2:2 complexes. Indeed in the crystal structure of LGR5-RSPO1-RNF43 an 'open' arrangement is observed, which is possibly dimerized sideways stabilized by a Ni^{2+} ion coordinated by residues His199 and His223 from both LGR5 molecules. Of note, Xu et al. observed a LGR4 dimer in solution [21]. Moreover, evaluation of the crystal structure (PDB code 4LI1) shows that LGR4 forms a related dimeric arrangement in the lattice (reminiscent of the reported 2:2 arrangement for LGR5-RSPO1 complexes).

Congenital anonychia is a mild disorder characterized by the absence of fingernails and toenails for which mutations have been identified in RSPO4 [26-29]. These mutations correspond to residues R66W, R70C, Q71R and G73R in RSPO1. To investigate the effect of mutations on Wnt signaling, we have previously performed Wnt reporter assay (TOPFlash) and observed reduced signaling activity [18]. Gln71 and Gly73 residues are located on the ZNRF3-RSPO1 interface and the Anonychia-related mutations, Q71R and G73R, would affect binding to ZNRF3 due to steric clashes and electrostatic repulsion (Figure S3D). Based on the structural

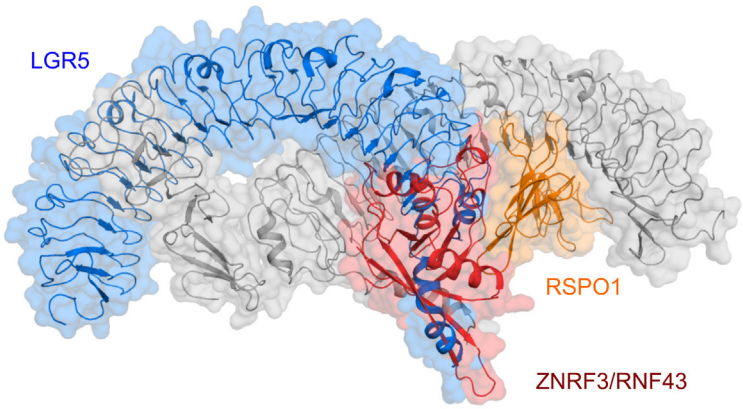


Fig 4. ZNRF3-RSPO1 binding site coincides with LGR5-RSPO1 'trans' site.
Shown in cartoon with transparent surface representation are the crystal structure of the 2:2 LGR5-RSPO1 complex (PDB code 4BSR) and the structure of ZNRF3-RSPO1 superimposed on RSPO1 on the right-hand side (orange), only the Cα trace of ZNRF3 (red) is shown for clarity. The overlapping LGR5 chain is shown in blue. The remaining part of the 2:2 LGR5-RSPO1, i.e. left-hand side RSPO1 and right-hand side LGR5, is shown in grey.

data the effects of R66W and R70C are expected to be less severe, because these residues are located at the periphery of the ZNRF3-RSPO1 interface. Indeed, R66W mutant showed slightly higher activity than Q71R and G73R [18], whereas R70C mutant express minimally as monomeric form in HEK 293 cells (data not shown). As described previously, Q71R and G73R may also affect the 'trans' LGR5-RSPO1 interactions, whereas R66W and R70C lie outside the observed interface and might be accommodated. Thus, based on the structural data both ZNRF3-RSPO1 and 'trans' LGR5-RSPO1 interactions may be affected by Anonychia-related mutations. Moreover, distinguishing between these two types of interactions in functional assays, such as the TOPFlash reporter assay, likely depends critically on the molecular ratio of LGR4-6 and ZNRF3/RNF43 receptors in the membrane.

CONCLUSIONS

Although a dominant role of LGR4-6 in Wnt activation of adult stem cell maintenance and proliferation have become very clear, the specific contributions of LGR4, LGR5 and LGR6 are not yet fully understood. Lgr5 expression is specific to stem cell compartments in various tissues, whereas Lgr4 shows a broader expression pattern [30]. R-spondins have been identified as ligands for LGR4-6, yet signaling does not seem to be coupled to G-proteins [7,9,10]. One recent study reports LGR5 (but not LGR4 or LGR6) activates the $G_{12/13}$-Rho GTPase pathway, but this activity is independent of R-spondins [31]. RNF43 and its homolog ZNRF3 have been identified as E3 ligases [1,2] that ubiquitinate Frizzled receptors for degradation, whereas RSPO1 captures ZNRF3 for removal from membrane, thereby increasing Frizzled expression on the cell surface. Our crystal structures show that ZNRF3 adopts a typical PA

domain, which does not undergo major conformational change upon binding to RSPO1. A dimeric arrangement of ZNRF3 is observed, which is plausible on the membrane, though evidence for a physiological role of such a dimer is currently lacking. The structure of ZNRF3-RSPO1 presented here, and the LGR5-RSPO1-RNF43 structure [19], elucidate the mode of interaction between RNF43/ZNRF3 and RSPO1. These structures provide a framework for studying disease mutations, e.g. those in RSPO4 causing congenital Anonychia. RSPO1 binds to ZNRF3 with weak (micromolar) affinity, in contrast to strong (nanomolar) binding affinity for LGR4-6. LGR5 did not increase the affinity of RSPO1 to ZNRF3; this is in contrast to the 10-fold increased affinity for RNF43 reported for LGR5-RSPO1 versus RSPO1 alone [19]. Overall, LGR4-6 most likely serve as recruitment receptors providing nanomolar-affinity binding sites for R-spondins on the membrane surface. While the strong affinity allows R-spondin to bind to LGR4-6 at low concentration to become effectively associated to the membrane, the weak affinity for ZNRF3-receptor ectodomain possibly allows for regulation of ubiquitination activity on Frizzled receptors, as proposed by Hao et al. [2]. Moreover, ZNRF3 most likely interacts with R-spondins and Frizzled receptors employing the same conserved binding platform.

In an earlier report, LGR4 has been found to interact with LGR5 and physically reside in LRP5/6-FZD complexes on the membrane, by tandem affinity purification and mass spectrometry [7]. In another report, Carmon et al. observed that LGR5 forms a supercomplex with LRP6/FZD5 receptors upon stimulation with RSPO1 [17]. Furthermore, LGR5 increased the endocytosis of LRP5/6 complexes in a dynamin- and clathrin-dependent manner. Crystal structures of 2:2 complexes of LGR5-RSPO1 [18] support the observation of LGR4-LGR5 heterodimers. The structures of ZNRF3-RSPO1 and LGR5-RSPO1-RNF43 [19] show that binding of ZNRF3 or RNF43 would disrupt the 2:2 LGR5-RSPO1 complex. These data would indicate the occurrence of multiple types of receptor complexes with potentially different roles, some of which are mutually exclusive. Further, we observed activation of LGR5 by antibodies, in the absence of R-spondin [18], and Kwon et al. showed $G_{12/13}$-Rho GTPase activation of LGR5 independent of R-spondin, implying a direct signaling role apart from that mediated by ZNRF3/RNF43-R-spondin interactions. In addition, RSPO3/4 interacts with Syndecan 4 to activate Wnt/planar cell polarity signaling [8,32]. Other receptors/ligands are also reported to interact with LGR4-6 and/or R-spondins, such as Norrin [33] and Troy [34]. Interestingly, Norrin binds LGR4-6; however, it only activates LGR4. This activation is unlikely mediated by ZNRF3 or RNF43. Various ligands may function to activate different downstream signaling pathway, spatially and temporally, during development. The multitude of proteins involved in Wnt signaling represent an intricate network essential for diverse activity in developmental biology.

MATERIALS AND METHODS

Protein expression, purification and crystallization

LGR5 ectodomain and RSPO1 furin-like domain were expressed and purified as described previously [18]. Mouse ZNRF3 construct (residue 53-205, Uniprot Q5SSZ7; human ZNRF3 residue number was used in the text and structure, i.e. 56-208, Uniprot Q9ULT6) was cloned into pUPE vector (U-Protein Express BV) carrying hexa-histidine tag. All proteins were produced

recombinantly in HEK 293E cells that stably expressed Epstein-Barr virus Nuclear Antigen I (EBNA) [35,36] provided by Utrecht-Protein Express BV (Utrecht, The Netherlands). Proteins were purified by Ni-NTA and gel filtration. Samples were concentrated to 10-15 mg/ml in buffer 25 mM HEPES pH 8.0, 50 mM NaCl and crystallized by hanging drop vapour diffusion method at 291 K. Crystals of ZNRF3 were obtained in 0.2M ammonium formate pH 6.6 and 20% w/v PEG 3350. Crystals of ZNRF3-RSPO1 were obtained in 0.2M sodium bromide and 20% w/v PEG 3350. Crystals were harvested and flash-cooled in liquid nitrogen in the presence of mother liquor supplemented with 20% ethylene glycol.

Data collection, structure determination and refinement

Diffraction data were collected at Swiss Light Source (SLS Villigen, Switzerland) and at European Synchrotron Radiation Facility (ESRF Grenoble, France). Data were processed by MOSFLM [37], XDS [38] and AIMLESS [39]. Resolution limits were determined by applying a cut-off based on the mean intensity correlation coefficient of half-datasets, $CC_{1/2}$ [40]. The structures of ZNRF3 and ZNRF3-RSPO1 were obtained by molecular replacement [41] using RNF43 (PDB code 4KNG) and RSPO1 (PDB code 4BSO) as search models. Model building for ZNRF3 was performed by ARP/WARP [42] and completed manually using COOT [43]. Structure refinements were performed using PHENIX [44] and REFMAC5 [45]. Molprobity [46] was used for structure validation. Structural analysis was performed using various softwares of the CCP4 suite, EBI PISA [47] and the DALI server [48]. Figures were generated with PyMol [49].

Surface plasmon resonance

Binding studies were performed using IBIS MX96 (IBIS Technologies) according to the protocol described previously [7]. Briefly, ZNRF3 ectodomain or human FZD8 cysteine-rich domain (residue 27-150, Uniprot Q9H461) constructs carrying a C-terminus biotin acceptor peptide (C-BAP) tag were co-expressed with biotin ligase (BirA) in HEK293-E cells to obtain in-vivo biotinylation. Biotinylated ZNRF3 protein was immobilized on a G-streptavidin sensor chip (IBIS Technologies) at different ligand densities. Analytes were flowed on the sensor chip in buffer containing 25 mM HEPES pH 8.0 and 150 mM NaCl at constant temperature of 25 °C. Binding affinities (KD) were calculated by global fitting based on a 1:1 discrete binding mode (SigmaPlot, Systat Software). Standard deviations were calculated from 4 experiments at different ligand density.

ACCESSION NUMBERS

The PDB accession numbers for the coordinates and structure factors of ZNRF3 and ZNRF3-RSPO1 reported in this paper are 4CDJ and 4CDK, respectively.

ACKNOWLEDGMENTS

We gratefully thank Richard Schasfoort and Alexander van der Kooi (IBIS technologies)

for assistance during surface plasmon resonance experiments. We thank the European Synchrotron Radiation Facility (ESRF) and the Swiss Light Source (SLS) for the provision of synchrotron radiation facilities and beamline scientists of the SLS, ESRF and the European Molecular Biology Laboratory for assistance.

AUTHOR CONTRIBUTIONS

Conceived and designed the experiments: PG WCP. Performed the experiments: WCP WDL JCMG PKM FF. Analyzed the data: WCP WDL HC PG. Wrote the manuscript: PG WCP.

REFERENCES

1. Koo B-K, Spit M, Jordens I, Low TY, Stange DE, et al. (2012) Tumour suppressor RNF43 is a stem-cell E3 ligase that induces endocytosis of Wnt receptors. Nature 488: 665–669. doi:10.1038/nature11308.
2. Hao H-X, Xie Y, Zhang Y, Charlat O, Oster E, et al. (2012) ZNRF3 promotes Wnt receptor turnover in an R-spondin-sensitive manner. Nature 485: 195–200. doi:10.1038/nature11019.
3. Zhou Y, Lan J, Wang W, Shi Q, Lan Y, et al. (2013) ZNRF3 acts as a tumour suppressor by the Wnt signalling pathway in human gastric adenocarcinoma. J Mol Histol. doi:10.1007/s10735-013-9504-9.
4. Jiang X, Hao H-X, Growney JD, Woolfenden S, Bottiglio C, et al. (2013) Inactivating mutations of RNF43 confer Wnt dependency in pancreatic ductal adenocarcinoma. Proceedings of the National Academy of Sciences. doi:10.1073/pnas.1307218110.
5. Ong CK, Subimerb C, Pairojkul C, Wongkham S, Cutcutache I, et al. (2012) Exome sequencing of liver fluke-associated cholangiocarcinoma. Nat Genet 44: 690–693. doi:10.1038/ng.2273.
6. Ryland GL, Hunter SM, Doyle MA, Rowley SM, Christie M, et al. (2013) RNF43 is a tumour suppressor gene mutated in mucinous tumours of the ovary. J Pathol 229: 469–476. doi:10.1002/path.4134.
7. de Lau W, Barker N, Low TY, Koo B-K, Li VSW, et al. (2011) Lgr5 homologues associate with Wnt receptors and mediate R-spondin signalling. Nature 476: 293–297. doi:10.1038/nature10337.
8. Glinka A, Dolde C, Kirsch N, Huang Y-L, Kazanskaya O, et al. (2011) LGR4 and LGR5 are R-spondin receptors mediating Wnt/β-catenin and Wnt/PCP signalling. EMBO Rep 12: 1055–1061. doi:10.1038/embor.2011.175.
9. Carmon KS, Gong X, Lin Q, Thomas A, Liu Q (2011) R-spondins function as ligands of the orphan receptors LGR4 and LGR5 to regulate Wnt/beta-catenin signaling. Proc Natl Acad Sci USA 108: 11452–11457. doi:10.1073/pnas.1106083108.
10. Ruffner H, Sprunger J, Charlat O, Leighton-Davies J, Grosshans B, et al. (2012) R-Spondin potentiates Wnt/β-catenin signaling through orphan receptors LGR4 and LGR5. PLoS ONE 7: e40976. doi:10.1371/journal.pone.0040976.
11. Barker N, van Es JH, Kuipers J, Kujala P, van den Born M, et al. (2007) Identification of

stem cells in small intestine and colon by marker gene Lgr5. Nature 449: 1003–1007. doi:10.1038/nature06196.
12. Jaks V, Barker N, Kasper M, van Es JH, Snippert HJ, et al. (2008) Lgr5 marks cycling, yet long-lived, hair follicle stem cells. Nat Genet 40: 1291–1299. doi:10.1038/ng.239.
13. Barker N, Huch M, Kujala P, van de Wetering M, Snippert HJ, et al. (2010) Lgr5(+ve) stem cells drive self-renewal in the stomach and build long-lived gastric units in vitro. Cell Stem Cell 6: 25–36. doi:10.1016/j.stem.2009.11.013.
14. Barker N, Rookmaaker MB, Kujala P, Ng A, Leushacke M, et al. (2012) Lgr5(+ve) Stem/Progenitor Cells Contribute to Nephron Formation during Kidney Development. Cell Rep 2: 540–552. doi:10.1016/j.celrep.2012.08.018.
15. Huch M, Dorrell C, Boj SF, van Es JH, Li VSW, et al. (2013) In vitro expansion of single Lgr5+ liver stem cells induced by Wnt-driven regeneration. Nature 494: 247–250. doi:10.1038/nature11826.
16. Plaks V, Brenot A, Lawson DA, Linnemann JR, Van Kappel EC, et al. (2013) Lgr5-expressing cells are sufficient and necessary for postnatal mammary gland organogenesis. Cell Rep 3: 70–78. doi:10.1016/j.celrep.2012.12.017.
17. Carmon KS, Lin Q, Gong X, Thomas A, Liu Q (2012) LGR5 interacts and cointernalizes with Wnt receptors to modulate Wnt/β-catenin signaling. Mol Cell Biol 32: 2054–2064. doi:10.1128/MCB.00272-12.
18. Peng WC, de Lau W, Forneris F, Granneman JCM, Huch M, et al. (2013) Structure of Stem Cell Growth Factor R-spondin 1 in Complex with the Ectodomain of Its Receptor LGR5. Cell Rep 3: 1885–1892. doi:10.1016/j.celrep.2013.06.009.
19. Chen P-H, Chen X, Lin Z, Fang D, He X (2013) The structural basis of R-spondin recognition by LGR5 and RNF43. Genes Dev 27: 1345–1350. doi:10.1101/gad.219915.113.
20. Wang D, Huang B, Zhang S, Yu X, Wu W, et al. (2013) Structural basis for R-spondin recognition by LGR4/5/6 receptors. Genes Dev 27: 1339–1344. doi:10.1101/gad.219360.113.
21. Xu K, Xu Y, Rajashankar KR, Robev D, Nikolov DB (2013) Crystal Structures of Lgr4 and Its Complex with R-Spondin1. Structure. doi:10.1016/j.str.2013.07.001.
22. de Lau WBM, Snel B, Clevers HC (2012) The R-spondin protein family. Genome Biol 13: 242. doi:10.1186/gb-2012-13-3-242.
23. Mahon P, Bateman A (2000) The PA domain: a protease-associated domain. Protein Sci 9: 1930–1934. doi:10.1110/ps.9.10.1930.
24. Luo X, Hofmann K (2001) The protease-associated domain: a homology domain associated with multiple classes of proteases. Trends Biochem Sci 26: 147–148.
25. Jin X, Cheng H, Chen J, Zhu D (2011) RNF13: an emerging RING finger ubiquitin ligase important in cell proliferation. FEBS J 278: 78–84. doi:10.1111/j.1742-4658.2010.07925.x.
26. Blaydon DC, Ishii Y, O'Toole EA, Unsworth HC, Teh M-T, et al. (2006) The gene encoding R-spondin 4 (RSPO4), a secreted protein implicated in Wnt signaling, is mutated in inherited anonychia. Nat Genet 38: 1245–1247. doi:10.1038/ng1883.
27. Bergmann C, Senderek J, Anhuf D, Thiel CT, Ekici AB, et al. (2006) Mutations in the gene encoding the Wnt-signaling component R-spondin 4 (RSPO4) cause autosomal recessive anonychia. Am J Hum Genet 79: 1105–1109. doi:10.1086/509789.

28. Khan TN, Klar J, Nawaz S, Jameel M, Tariq M, et al. (2012) Novel missense mutation in the RSPO4 gene in congenital hyponychia and evidence for a polymorphic initiation codon (p.M1I). BMC Med Genet 13: 120. doi:10.1186/1471-2350-13-120.
29. Wasif N, Ahmad W (2013) A novel nonsense mutation in RSPO4 gene underlies autosomal recessive congenital anonychia in a Pakistani family. Pediatr Dermatol 30: 139–141. doi:10.1111/j.1525-1470.2011.01587.x.
30. Barker N, Clevers H (2010) Leucine-rich repeat-containing G-protein-coupled receptors as markers of adult stem cells. Gastroenterology 138: 1681–1696. doi:10.1053/j.gastro.2010.03.002.
31. Kwon MS, Park B-O, Kim HM, Kim S (2013) Leucine-rich repeat-containing G-protein coupled receptor 5/GPR49 activates G12/13-Rho GTPase pathway. Mol Cells. doi:10.1007/s10059-013-0173-z.
32. Ohkawara B, Glinka A, Niehrs C (2011) Rspo3 binds syndecan 4 and induces Wnt/PCP signaling via clathrin-mediated endocytosis to promote morphogenesis. Dev Cell 20: 303–314. doi:10.1016/j.devcel.2011.01.006.
33. Deng C, Reddy P, Cheng Y, Luo C-W, Hsiao C-L, et al. (2013) Multi-functional norrin is a ligand for the LGR4 receptor. J Cell Sci. doi:10.1242/jcs.123471.
34. Fafilek B, Krausova M, Vojtechova M, Pospichalova V, Tumova L, et al. (2012) Troy, a Tumor Necrosis Factor Receptor Family Member, Interacts With Lgr5 to Inhibit Wnt Signaling in Intestinal Stem Cells. Gastroenterology. doi:10.1053/j.gastro.2012.10.048.
35. Durocher Y, Perret S, Thibaudeau E, Gaumond MH, Kamen A, et al. (2000) A reporter gene assay for high-throughput screening of G-protein-coupled receptors stably or transiently expressed in HEK293 EBNA cells grown in suspension culture. Anal Biochem 284: 316–326. doi:10.1006/abio.2000.4698.
36. Reeves PJP, Callewaert NN, Contreras RR, Khorana HGH (2002) Structure and function in rhodopsin: high-level expression of rhodopsin with restricted and homogeneous N-glycosylation by a tetracycline-inducible N-acetylglucosaminyltransferase I-negative HEK293S stable mammalian cell line. Proc Natl Acad Sci USA 99: 13419–13424. doi:10.1073/pnas.212519299.
37. Leslie A, Powell HR (2007) Processing diffraction data with MOSFLM. Evolving methods for macromolecular crystallography, NATO Science Series Volume 245, pp 41-51.
38. Kabsch W (2010) XDS. Acta Crystallogr D Biol Crystallogr 66: 125–132. doi:10.1107/S0907444909047337.
39. Evans PR (2011) An introduction to data reduction: space-group determination, scaling and intensity statistics. Acta Crystallogr D Biol Crystallogr 67: 282–292. doi:10.1107/S090744491003982X.
40. Karplus PA, Diederichs K (2012) Linking crystallographic model and data quality. Science 336: 1030–1033. doi:10.1126/science.1218231.
41. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, et al. (2007) Phaser crystallographic software. J Appl Crystallogr 40: 658–674. doi:10.1107/S0021889807021206.
42. Langer G, Cohen SX, Lamzin VS, Perrakis A (2008) Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. Nature protocols.

43. Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. Acta Crystallogr D Biol Crystallogr 66: 486–501. doi:10.1107/S0907444910007493.
44. Adams PD, Afonine PV, Bunkóczi G, Chen VB, Davis IW, et al. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr 66: 213–221. doi:10.1107/S0907444909052925.
45. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, et al. (2011) Overview of the CCP4 suite and current developments. Acta Crystallogr D Biol Crystallogr 67: 235–242. doi:10.1107/S0907444910045749.
46. Chen VB, Arendall WB, Headd JJ, Keedy DA, Immormino RM, et al. (2010) MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr D Biol Crystallogr 66: 12–21. doi:10.1107/S0907444909042073.
47. Krissinel E, Henrick K (2007) Inference of macromolecular assemblies from crystalline state. J Mol Biol 372: 774–797. doi:10.1016/j.jmb.2007.05.022.
48. Holm L, Rosenström P (2010) Dali server: conservation mapping in 3D. Nucleic Acids Res 38: W545–W549. doi:10.1093/nar/gkq366.
49. Schrodinger, LLC (2013). The PyMOL Molecular Graphics System, Version 1.5.0.4.

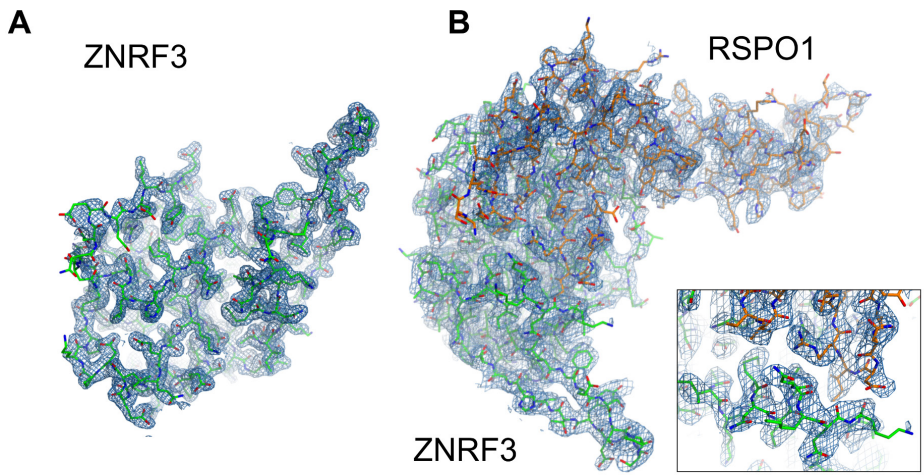


Fig S1. Electron densities of ZNRF3 and the ZNRF3-RSPO1 complex.
A, Electron density (blue), 2mFo-DFc map contoured at 1 σ level, for ZNRF3. The model is shown in green.
B, Electron density for one of the four ZNRF3-RSPO1 complexes in the asymmetric unit with ZNRF3 in green and RSPO1 in orange. The insert shows a zoom-in of the density at the ZNRF3-RSPO1 interface.

A

```
                                      60         70         80         90
hZNRF3          ---- ---------- -----KETAF VEVVLFESSP SGDYTTYTTG LTGRFSRAGA TLSAEG
hRNF43          GFGR TGLVLAAAVE SERSAEQKAI IRVIPLKMDP TGK---LNLT LEGVFAGVAE ITPAEG
                                      47         57         64         74
                                    ::.*: : *: :: .* :*.    .   * * *: ..    .***

                 100        110        120        130        140        150
hZNRF3          EIVQ MHPLGLCNNN DEEDLYEYGW VGVVKLEQPE LDPKPCLTVL GKAKRAVQRG ATAVIF
hRNF43          KLMQ SHPLYLCNAS DDDNL-EPGF ISIVKLESPR RAPRPCLSLA SKARMAGERG ASAVLF
                   84         94        103        113        123        133
                ::::*  *** *** . *:::* * *: :.:****.*    *.***::  .**. * :** *:**:*

                 160        170        180        190        200        210
hZNRF3          DVSE NPEAIDQLNQ GSEDPLKRPV VYVKGADAIK LMNIVNKQKV ARARIQHRPP RQPTEY
hRNF43          DITE DRAAAEQLQQ PLG--LTWPV VLIWGNDAEK LMEFVYKNQK AHVRIELKEP PAWPDY
                 143        153        161        171        181        191
                *::** :  * :**:*      *..** * :.* ** * **::* *::  *..**: . *    .:*

hZNRF3          FDM
hRNF43          ---
```

B

```
                       30         40         50         60         70         80         90         100        110
Rspo1           SRGIKGKRQR RISAEGSQAC AKGCELCSEV NGCLKCSPKL FILLERNDIR QVGVCLPSCP PGYFDARNPD MNKCIKCKIE HCEACFSHNF
Rspo2           -QGNRWRRSK RASYVSNPIC -KGCLSCSKD NGCSRCQQKL FFFLRREGMR QYGECLHSCP SGYYGHRAPD MNRCARCRIE NCDSCFSKDF
Rspo3           QNASRGRRQR RMHPNVSQGC QGGCATCSDY NGCLSCKPRL FFALERIGMK QIGVCLSSCP SGYYGTRYPD INKCTKCKAD -CDTCFNKNF
Rspo4           ----LNRRKK QVGTGLGGNC -TGCIICSEE NGCSTCQQRL FLFIRREGIR QYGKCLHDCP PGYFGIRGQE VNRCKKCGAT -CESCFSQDF
                     .*.. .    .  *   **  **.  ***  *. .* *: : * .:. * * ** .**  .**:.*  : :*.* .*     *::**.::*

                      120        130        140        150        160        170        180        190        200
Rspo1           CTKCKEGLYL HKGRCYPACP EGSSAANGTM ECSSPAQCEM SEWSPWGPCS KKQQLCGFRR GSEERTRRVL HAPVGDHAAC SDTKETRRCT
Rspo2           CTKCKVGFYL HRGRCFDECP DGFAPLEETM ECVE--GCEV GHWSEWGTCS RNNRTCGFKW GLETRTRQIV KKPVKDTILC PTIAESRRCK
Rspo3           CTKCKSGFYL HLGKCLDNCP EGLEANNHTM ECVSIVHCEV SEWNPWSPCT KKGKTCGFKR GTETRVREII QHPSAKGNLC PPTNETRKCT
Rspo4           CIRCKRQFYL YKGKCLPTCP PGTLAHQNTR ECQG--ECEL GPWGGWSPCT HNGKTCGSAW GLESRVREAG RAGHEEAATC QVLSESRKCP
                * .**  :** : *.*   **  *  . : *  **    **:.  *. *..*:. : . **  .  * * *.* .       .   *     *:*.*

                      210        220        229        231        241        251        263
Rspo1           VRRVPCPEGQ KRRKGGQGRR ENANRNLAR- --------KE SKEAGAGSRR RKGQQQQQQQ GTVGPLTSAGPA
Rspo2           MTMRHCPGGK RTPKAKEKRN KK-------- ---------- --------KK RKLIERAQEQ HSVFLATDRANQ
Rspo3           VQRKKCQKGE RGKKGRERKR KKPNKGESKE AIPDSKSLES SKEIPEQREN KQQQKKRKVQ DKQKSVSVSTVH
Rspo4           IQR-PCP-GE RSPGQKKGRK DRRPR----- ---------- --------KD RKLDRRLDVR PRQPGLQP----
                :    *  *: .     : .. ..                           .:   . . .
```

Fig S2. Sequence alignments of ZNRF3 and RNF43 ectodomains and RSPO1-4.
A, Alignment of human ZNRF3 and RNF43. Mouse ZNRF3 differs from its human homolog at three positions (mouse: His77, Met91 and Leu208).
B, Alignment of human RSPO1-4. The shaded areas correspond to the Fu1-Fu2 domains.

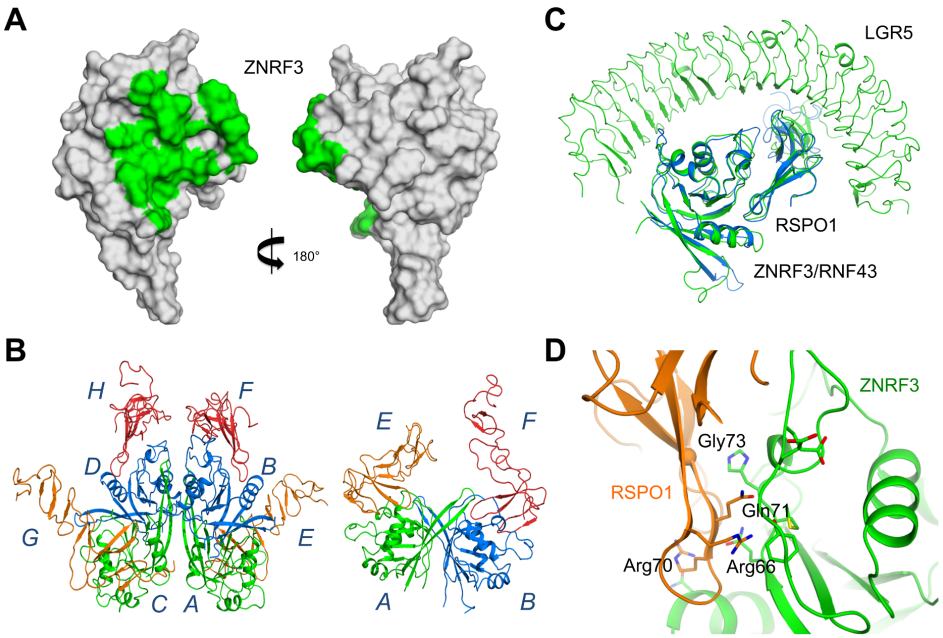


Fig S3. Structural analyses of ZNRF3-RSPO1 complex.
A, Contact area ('footprint') of RSPO1 plotted onto the surface of ZNRF3. ZNRF3 is shown in surface representation with the area in contact with RSPO1 (using a distance criterium of 4.5 Å) highlighted in green. The orientation of the two views is identical as in Fig. 2C.
B, Arrangement of the dimer of dimers of ZNRF3-RSPO1 complexes in the asymmetric unit (left side) and the dimeric arrangement based on the ZNRF3 dimer observed in Fig 1C (right side). ZNRF3 molecules are shown in blue and green, RSPO1 in orange and red; the chain labels are indicated.
C, Superposition of the ZNRF3-RSPO1 structure (blue) onto the structure of the LGR5-RSPO1-RNF43 complex (green; PDB code 4KNG).
D, Zoom-in of the ZNRF3-RSPO1 interface, with RSPO1 shown in orange and ZNRF3 in green, highlighting the four residues related to congenital anonychia mutations in RSPO4: R66W, R70C, Q71R and G73R.

Table 1. Crystallographic statistics for data collection and refinement.

	ZNRF3	ZNRF3-RSPO1
Data Collection[a]		
X-ray source	SLS X06DA	SLS X06SA
Processing programs	XDS/AIMLESS	iMOSFLM/AIMLESS
Space group	$P2_1$	P1
Cell parameters	a = 35.7 Å; α = 90.0° b = 73.5 Å; β = 97.5° c = 58.6 Å; γ = 90.0°	a = 51.7 Å; α = 66.3° b = 80.2 Å; β = 81.4° c = 83.0 Å; γ = 80.7°
Wavelength (Å)	1.00	1.00
Resolution (Å)	45.57 – 1.50 (1.53-1.50)	75.63 – 2.80 (2.95-2.80)
Unique reflections	46197 (4495)	29157 (4105)
$CC1/2$[b]	1.00 (0.94)	0.96 (0.21)
Redundancy	2.7 (2.4)	2.2 (2.3)
$I/\sigma(I)$	24.1 (4.1)	6.2 (3.2)
Completeness (%)	96.6 (94.0)	95.4 (95.0)
R_{sym}[c]	0.019 (0.162)	0.093 (0.391)
Wilson B-factor ($Å^2$)	17.66	54.27
Refinement		
Molecules per ASU	2	4
R_{work}/R_{free}[c]	0.162/0.177	0.218/0.246
Average B-factors ($Å^2$)	28.1	83.8
Number of atoms:	5184	7982
Protein	4861	7927
Ligands	27	0
Waters	296	55
Structure quality		
Molprobity score	1.84	2.05
RMS bond lengths (Å)	0.018	0.005
RMS bond angles (°)	1.83	1.21
Ramachandran favored (%)	98	93
Ramachandran outliers (%)	1	2

[a] Values in parentheses are for reflections in the highest resolution shell.
[b] Resolution limits were determined by applying a cut-off based on the mean intensity correlation coefficient of half-datasets, $CC_{1/2}$ (Karplus and Diederichs, 2012).
[c] $R_{sym} = \Sigma \mid I - \langle I \rangle \mid / \Sigma I$, where I is the observed intensity for a reflection and <I> is the average intensity obtained from multiple observations of symmetry-related reflections. R_{free} values are calculated based on 5% randomly selected reflections.

CHAPTER 4

Strategies for expression, crystallization and structure determination of LGR5 and R-spondin 1

Weng Chuan Peng, Pramod K Madoori, Federico Forneris, Joke C.M. Granneman and Piet Gros

[1]Crystal and Structural Chemistry, Bijvoet Center for Biomolecular Research, Department of Chemistry, Faculty of Science, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands

ABSTRACT

LGR5, a Wnt target gene, is exclusively expressed in the Crypt Base Columnar cells in the intestinal crypts. LGR5 associates with R-spondin, a potent Wnt agonist, to potentiate Wnt signaling. Crystallographic study of LGR5 ectodomain in complex with R-spondin provides a molecular basis of receptor-ligand interaction. In this chapter, we described our strategy for large-scale production of protein for crystallization study. By co-expression of receptor-ligand complex (consisting of LGR5 ectodomain and R-spondin Furin-like domain) and plasmid titration, we obtained milligrams of soluble proteins for structural studies. In addition, we present our strategy for structure determination of LGR5 leucine-rich repeat domain, using non-homologous models in Molecular Replacement searches.

INTRODUCTION

LGR4 and LGR5 are leucine-rich repeat (LRR)-containing G-protein coupled receptors (GPCR) identified by Hsueh's lab in 1998 [1]. LGR4/5 are GPCRs with large ectodomains and belong to the class A family (subfamily A10) of GPCRs. Other receptors in the same family include the well-characterized glycohormone protein follicle-stimulating hormone receptor (FSHR), luteinizing hormone receptor (LHR) and thyroid-stimulating hormone receptor (TSHR) [2,3]. These receptors are also known as LGR 1-3, respectively. FSHR binds follitropin (FSH) and LHR binds lutropin (LH) and choriogonadotropin (CG), essential for gonad development; TSHR binds TSH, essential for thyroid development. FSH, LH and TSH are secreted by anterior pituitary gland, whereas CG is secreted by the placenta. Together, these hormones regulate sexual development and reproductive function [2,3]. However, the function of LGR4/5 was not known at the time of their discovery. Hsu et al. observed that LGR4 is expressed in tissues such as placenta, ovary, testis, and adrenal, spinal cord and thyroid [1]. On the other hand, LGR5 showed more restricted expression, with the highest expression in skeletal muscle, and in tissues such as placenta and spinal cord [1]. However, their ligands were not identified and LGR4/5 remained orphan receptors for many years.

In 2007, Barker et al. identified that LGR5, a Wnt target gene, is exclusively expressed in the Crypt Base Columnar (CBC) cells in the intestinal crypt [4,5]. By lineage tracing they observed that LGR5-expressing CBC cells gave rise to all other differentiated cell types, e.g. Paneth cells, goblet cells and enterocytes. This study identified CBC cells as the adult stem cells in the intestine, supporting early observation made in the 1970s by Cheng and Leblond [6]. LGR5 serves as the adult stem-cell marker and allows tracking of CBC cells proliferation in the crypt and their migration into the villus. Subsequently, Sato et al. demonstrated that single LGR5-expressing CBC cell can be cultured in vitro to form minigut, which displayed crypt-villus like architecture [7,8]. Gradually, lineage tracing strategy was used to identify adult stem cells in colon [4], hair follicles (LGR6) [9], stomach [10], kidney [11], pancreas [12], liver [13] and mammary gland [14].

R-spondin, a small secreted ligand, was identified as a Wnt agonist (activator) by Kazanskaya et

al. in 2004 [15]. R-spondin induced stabilization of β-catenin in HEK 293T cells, in the presence of Wnt and its coreceptors LRP6 and Frizzled. Kim et al. later showed that R-spondin stimulates growth of the intestinal tissue, providing a physiological evidence for R-spondin role in Wnt signaling [16]. Since then, many studies have attempted to identify the receptor that binds R-spondin. LRP6 and Kremen have been proposed [17,18] but the actual receptor was not identified till recently. In 2011, de Lau et al. and other groups independently showed that LGR5 and homologues (LGR4/6) associate with R-spondin to potentiate Wnt signaling [19-22], bridging two seemingly unrelated proteins into one pathway. However, LGR5 signaling does not seem to be dependent on G-proteins and signaling mechanism remained unclear [22,23]. Nonetheless, mass spectrometry analysis indicated that LGR5 interacts with LRP6-Fzd-Wnt signaling complex to potentiate Wnt signaling [23].

We aimed to decipher the molecular details of the LGR5-R-spondin interaction, and to understand how LGR5-R-spondin interaction potentiates Wnt signaling. Knowledge of this would allow us to modulate the activity of LGR5-R-spondin signaling pathway, which is often implicated in intestinal and colorectal cancers [24-28]. However, the productions of LRR proteins are not always straightforward, due to their intrinsic propensity to aggregate, and the yields of expression are usually low and insufficient for crystallization studies [29-31]. In this chapter, we discuss the various strategies for expression, purification and crystallization of recombinant LGR5 and R-spondin. In addition, we present here our strategy for structure determination of LGR5, using non-homologous LRR proteins. Although the expression and structure determination strategies are shown here for LGR5, the approach can be applied to other LRR proteins [33] in general.

RESULTS AND DISCUSSIONS

Domain organization and construct design

LGR5 has a signal peptide, a LRR ectodomain, a 7-transmembrane GPCR and a short cytoplasmic tail (Fig 1) [1]. The LRR ectodomain is characterized by N-terminal cysteine cap (LRRNT) and C-terminal cysteine cap (LRRCT, also called hinge region), which connect the LRRs to the 7-TM domain. 16 LRRs (residues 67-446) were annotated in Uniprot and predicted by SMART [32]. Each LRR has 22-24 residues. The consensus sequence given by SMART is shown in Fig 2. LRRs contain a highly conserved sequence LxxLxLxxNxL characterized by presence of Leu residues (or other small hydrophobic residues) at specific positions of the repeat (Fig 2), where the side chains face inner hydrophobic core of the LRRs, while the Asn residue forms hydrogen bond with neighbouring repeat [33,34]. Upon manual inspection, we identified that residues 446-470 (26 residues), though they do not align perfectly with the rest of LRRs, match the consensus LRR sequence, and hence constitute the 17th LRR (Fig 2). In the LRRNT, there are four Cys residues, predicted to form two disulfide bonds, as observed in FSHR and TSHR (Fig 1B). In LRRCT, there are 5 Cys residues (residues 479, 480, 485, 541 and 551). Two disulfide bonds are possible and a free Cys maybe present.

The structures of FSHR and TSHR (LGR1 and 3, respectively) are composed of 9 LRRs that

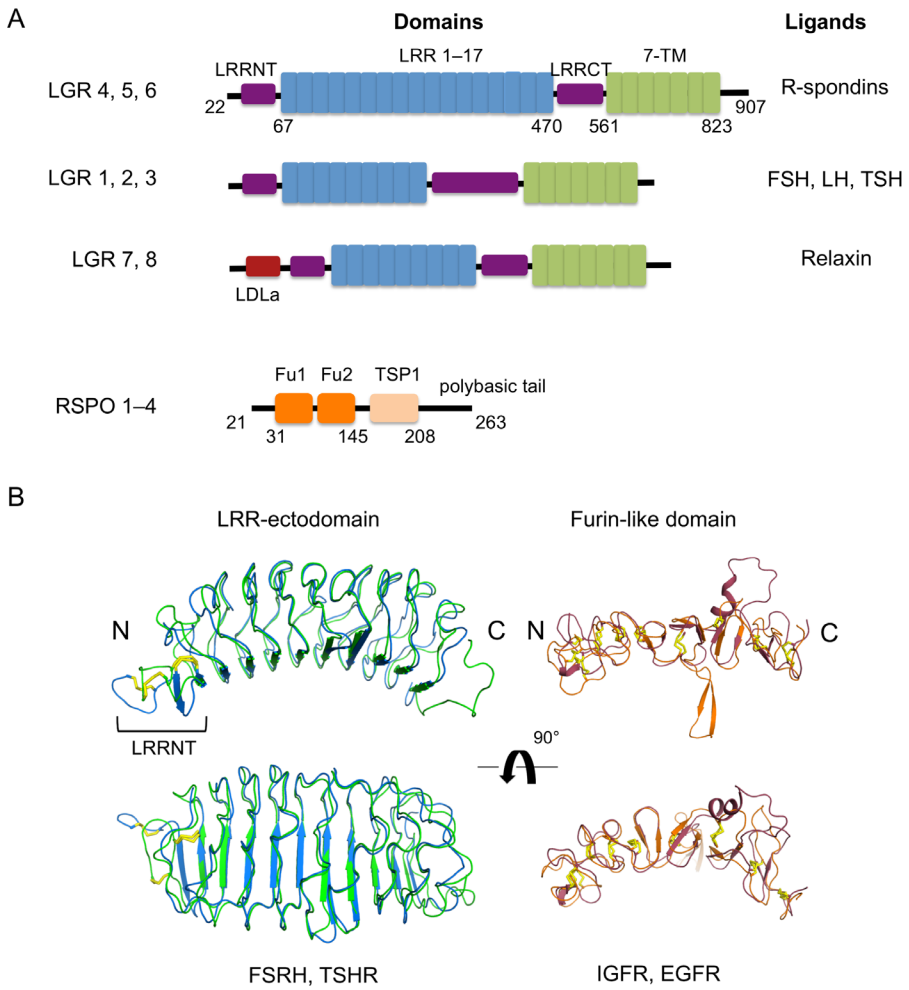


Fig 1. Domains organization and structures of the leucine-rich repeat G-protein coupled Receptor (LGR) and R-spondin families.
A, Schematic representation of the three different classes of LGR receptors and R-spondin.
B, Cartoon representation of LRR ectodomains of FSHR (PDB 1XWD, green) and TSHR (PDB 2XWT, marine) superposed; and furin-like domain of IGFR (PDB 1IGR, orange) and EGFR (PDB 1IVO, red) superposed. Dilsulfide bonds in the LRRNT are indicated in yellow. LRRCT is not present in the crystal structures.

superimpose with a root-mean-square deviation (RMSD) on Cα atoms of 1.0 Å (Fig 1B). Both proteins were crystallized in the absence of LRRCT, presumably to enhance chances of crystallization by omitting the putative disordered region [30,35]. The LRRCT of LGR5 was predicted to be disordered according to PSIPRED [36] analysis (data not shown), so we designed

```
   sLcpLsLssNpLssLPsssFpsLs (consensus: 50%)
   pLppLpLspNplpplstshFpsLs (consensus: 65%)
   pLp.L.LptNtlptls.thhttL. (consensus: 80%)
```

	sLcpLsLssNpLssLPsssFpsLs		sLcpLsLssNpLssLPsssFpsLs
1.	FTSYLDLSMNNISQLLPNPLPSLR	9.	NLKELGFHSNNIRSIPEKAFVGNP
2.	FLEELRLAGNALTYIPKGAFTG	10.	SLITIHFYDNPIQFVGRSAFQH
3.	SLKVLMLQNNQLRHVPTEALQN	11.	ELRTLTLNGASQITEFPDLTGTA
4.	SLQSLRLDANHISYVPPSCFSG	12.	NLESLTLTGAQISSLPQTVCNQ
5.	SLRHLWLDDNALTEIPVQAFRS	13.	NLQVLDLSYNLLEDLPSFSVCQ
6.	ALQAMTLALNKIHHIPDYAFGN	14.	KLQKIDLRHNEIYEIKVDTFQQ
7.	SLVVLHLHNNRIHSLGKKCFDG	15.	SLRSLNLAWNKIAIIHPNAFST
8.	SLETLDLNYNNLDEFPTAIRTLS	16.	SLIKLDLSSNLLSSFPITGLHGLT
		17.	GLTHLKLTGNHALQSLISSENFPELKV

```
s: small ( A,C,D,G,N,P,S,T,V)    c:charged (D,E,H,K,R)
p: polar (C,D,H,E,H,K,N,Q,R,S,T) t:tiny (A,G,S)
h: hydrophobic (A,C,F,G,H,I,K,L,M,R,T,V,W,Y)
.: any amino acids
```

LRRNT:

```
22       31
GSSPRSGVLLRGCPTHCHCEPDGRMLLRVDCSDLGLSELPSNLSV
```

LRR17+LRRCT:

```
           455     463      472   478
GLTHLKLTGNHALQSLISSENFPELKVIEMPYAYQCCAFGVCENAYKISNQ
                                                543
WNKGDNSSMDDLHKKDAGMFQAQDERDLEDFLLDFEEDLKALHSVQCSPSP
       553
GPFKPCEHLLDGWLI
```

Fig 2. Sequence analysis and constructs design for LGR5
Consensus sequence of LRRs are shown and 17 LRRs are aligned to the consensus sequence found in 65% of LRRs. Residues located on Leu, Asn, Phe, Pro positions in the consensus sequence are shown in red.
LRRNT and LRRCT sequences. Underscored residue on LRRNT indicate first residue in all expression contructs, and underscored residues on LRRCT indicate last residue in different expression constructs.

construct with and without LRRCT. Since the disulfide bond assignment is ambiguous in this region, we designed constructs with five Cys (residue 22-553) or four Cys (residue 22-553) in the LRRCT. In addition, a series of constructs omitting the LRRCT were made, with truncation at different residues on the last LRR (Fig 2), i.e. residues 22-478, 22-472, 22-463 and 22-455 (Uniprot O75473 sequence is used to denote residue numbers in constructs throughout the text).

R-spondin is a family of secreted proteins that has four members, R-spondin 1–4 [37,38].

R-spondin has a signal peptide, two furin (Fu) domains (~100 residues), a thrombospondin type I (TSP-1) domain (~60 residues) and a polybasic tail (~60 residues). The furin domain and TSP-1 domain are separated by a short linker of ~10 residues. These domains are conserved in R-spondin 1–4 (which show ~60% sequence identity). The polybasic regions of the four R-spondins have varying lengths [37]. There are two additional isoforms of R-spondin; one isoform lacks the signal peptide yielding presumably an intracellular variant [39] and the other isoform lacks the TSP-1 domain [40]. The functions and cellular localization of these two isoforms are not clear. Furin domains are characterized by short β-strands and loops held by a series of disulfide bond (Fig 1B) [41,42]. Furin domains are commonly found in receptor tyrosine kinase, e.g. insulin-like growth factor receptor (IGFR) and epidermal growth factor (EGFR). In EGFR, the Fu domain mediates dimerization of the receptor [42]. TSP-1 domain was originally described in thrombospondin 1 protein [43], but is also found in other proteins, e.g. the complement components C6, C7 and complement-regulator properdin [43,44]. The TSP-1 domain are characterized by presence of three anti-parallel strand held by three disulfide bonds and a mannosylated Trp [43,44]. Two of the three disulfide bonds cap the N- and C-termini. For the expression of R-spondin for structural studies, we have designed three constructs: full length R-spondin 1 (Uniprot: Q2MKA7, RSPO-FL, residue 21-263), Fu1-Fu2-TSP domains (RSPO-TSP, residue 31-145) and the two Fu domains, Fu1-Fu2 (RSPO-FF, residue 31-145).

Transient expression in HEK 293 E and ES cells

For structural studies, recombinant proteins are expressed in HEK 293 cells, due to the presence of post-translational modification and disulfide bond formation [45,46]. The HEK 293E cell line that stably expresses Epstein-Barr viral antigen I (EBNA-I) is routinely used in our lab. The EBNA-I drives episomal replication of oriP containing plasmid [47]. In addition, the N-acetylglucosaminyl transferase I (GnT-I)-deficient HEK 293 cell line (termed HEK 293ES) was used to recombinant proteins with uniform $Man_5GlcNAc_2$ Asn-linked glycans [48]. All constructs used N-terminal human cystatin-S signal peptide (Uniprot P01036) for secretion in HEK 293 cells.

I. Screening of different expression constructs for LGR5 and RSPO1 in HEK 293E cells

For expression of LGR5 ectodomain, HEK 293E was used for initial screening. All LGR5 expression constructs carried an N-terminal His_6-tag, followed by a Tobacco Etch Virus (TEV) protease site and were purified by Ni-NTA beads. All constructs, except construct 22-455, showed expression at comparable level, with a major band at ~60 kDa (Fig 3A). In non-reducing gel, we did not observe any disulfide-linked aggregates, even in the constructs 22-553, which has 5 Cys residues in the LRRCT.

For R-spondin expression, all three constructs: RSPO1-FL (MW: 27 kDa), RSPO1-TSP (20 kDa) and RSPO1-FF (12.5 kDa) expressed with major bands observed at 40 kDa, 27 kDa and 20 kDa, respectively (Fig 3A). All constructs carried a C-terminus His_6 tag and were purified by Ni-

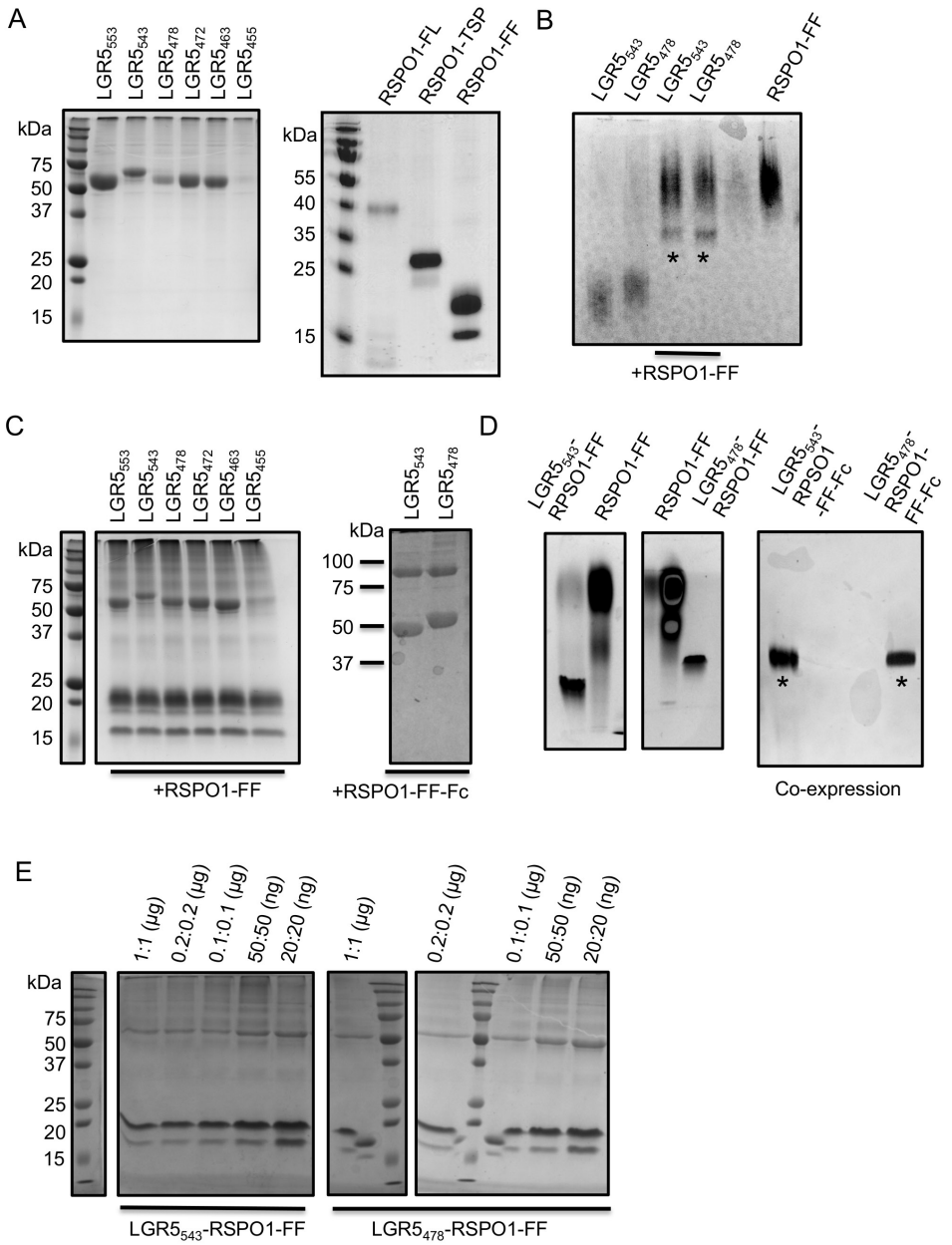
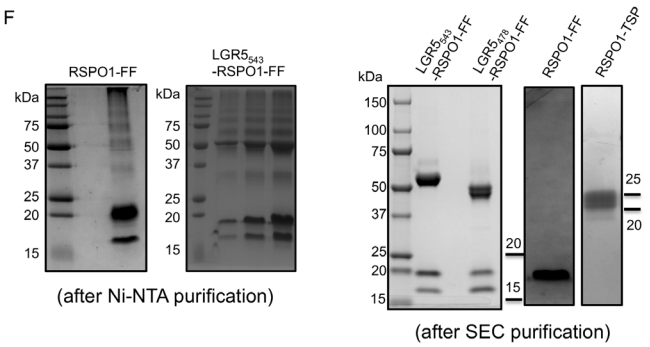


Fig 3. SDS-PAGE and native-PAGE analysis for the expression of LGR5 and RSPO1.
A, Expression of LGR5 ectodomains and RSPO1 constructs.
B, Complex formation of LGR5 and RSPO1.
C, Co-expression of LGR5 and RSPO1.
D, LGR5-RSPO1 complexes formation.
E, Plasmid titration for co-expression of LGR5-RSPO1 in HEK 293ES cells. (see next page)



F, Purification of LGR5-RSPO1 complexes and RSPO1.
A, C, E, and F are non-reducing SDS-PAGE; B and C are native-PAGE; all expression tests were carried out in HEK 293E cells, except E and F in HEK 293 ES cells. Fig F (right panel) is adapted from chapter 2, Fig S3.

NTA beads. The proteins migrated higher than their expected molecular mass in non-reducing gel, possibly due to elongated shape of the protein and presence of post-translational modifications. For RSPO1-FF, two bands were observed (at 20 kDa and 15 kDa), which correspond to the glycosylated and non-glycosylated forms, as a glycosylation site is predicted at Asn137. The expression of RSPO1-FF and RSPO1-TSP is higher than RSPO1-FL, and it was previously reported that full length R-spondin associates with membrane in HEK 293T cells and is not secreted into the medium [15]. An expression construct with mutation N137Q did not show any expression (data not shown), suggesting a role of glycosylation in protein folding [49].

II. Co-expression of LGR5-RSPO1 and 'plasmid titration' in HEK 293 ES cells improved yield and stability

To evaluate the behaviour of a protein (e.g. whether proteins are well-folded or aggregated), we used native PAGE, which provides more reliable evaluation than SDS-PAGE. In addition, native-PAGE can be used to detect receptor-ligand interaction. Hence, we used native-PAGE (Phastsystem, GE) to characterize LGR5-RSPO1 complex formation. On native-PAGE (Fig 3B), $LGR5_{543}$, $LGR5_{478}$ and RSPO1-FF can be seen as diffuse (or multiple) bands. By adding RSPO1-FF to $LGR5_{543}$ or $LGR5_{478}$, an additional well-defined band could be observed, which indicate the formation of the LGR5-RSPO1 complex. Excess RSPO1-FF could also be observed on the gel, migrating above the LGR5-RSPO1 complex band.

$LGR5_{543}$ was chosen as initial target for large-scale expression, but showed aggregation and purified protein had a propensity to precipitate at concentration higher than 0.5 mg/ml. We hypothesized that LGR5 ectodomains are prone to aggregation, presumably due to the hydrophobic characteristic nature of LRRs [31]. Hence, we tested co-expression of LGR5

ectodomain and RSPO1 to improve yield and stability of LGR5 proteins. All LGR5 ectodomain constructs (except residues 22-455) could be co-expressed with RSPO1-FF (Fig 3C). Here, it is also noticeable that RSPO1-FF was expressed at higher yield than LGR5 ectodomain, as judged from band intensities on SDS gels. In addition, $LGR5_{543}$ and $LGR5_{478}$ were co-expressed with an RSPO1-FF construct carrying an Fc tag at the C-terminus. The RSPO1-FF-Fc protein could be co-purified by His_6 tag, indicating complex formation in the medium (Fig 3C). This is possible due to the high nanomolar affinity observed for LGR5-RSPO1 interaction [19,20].

For crystallization purposes, we have expressed LGR5-RSPO1 by co-expression in HEK 293ES cells, because recombinant proteins with uniform $Man_5GlcNAc_2$ glycans are often more favourable for crystallization than proteins with large heterogeneous glycans. The yield of LGR5-RSPO1 complex obtained from ES cells expression is lower than from E cells (Fig 3C). To increase the yield of protein, we have performed 'plasmid titration', i.e. varying the amount of DNA to achieve the optimal DNA concentration for the highest yield of soluble well-folded protein, and to reduce protein aggregation. Plasmid titration has been previously used to obtain higher yield of intracellular proteins that are prone to aggregation, such as Nod-like receptor [31]. Various dilution factors were tested, at LGR5:RSPO1 ratio of 1:1 µg, 0.2:0.2 µg, 0.1:0.1 µg, 50:50 ng and 20:20 ng. The total DNA amount was kept constant at 2 µg with dummy DNA (empty pCR4 TOPO vector, Invitrogen). At the lowest amount of DNA (20 ng), highest yield of $LGR5_{543}$ and $LGR5_{478}$ were obtained (Fig 3E). We did not test lower amount of DNA. At all concentrations, RSPO1-FF was expressed at higher level than LGR5. Subsequently the amount of LGR5:RSPO1-FF DNA was reduced to 20:4 ng for large-scale expression. The $LGR5_{543}$-RSPO1-FF and $LGR5_{478}$-RPSO1-FF complexes purified migrate as well-defined bands in native-PAGE analysis (Fig 3D). The complexes obtained from co-expression are less prone to aggregation during purification and concentration (>15mg/ml) compared to LGR5 ectodomain alone (<1mg/ml) and is stable at room temperature for weeks. It is likely that complex formation of LGR5 and RSPO1 buried a hydrophobic region on LGR5 that would otherwise cause aggregation and precipitation at high protein concentration.

Large scale expression and purification

For large-scale expression, the RPSO1-FF construct was expressed in 2 L of HEK 293E cells with 1 mg of DNA. The culture medium was harvested 6 days post-transfection, concentrated and diafiltrated in binding buffer 25 mM HEPES pH 7.5, 50 mM NaCl (buffer A) for binding to Ni-NTA column. Non-specific proteins binding to Ni-NTA column was washed away with buffer A containing 25 mM imidazole, and bound RSPO1-FF was eluted in buffer A containing 0.3 M imidazole. Fractions containing RSPO1-FF were pooled and treated with PNGase F for removal of glycan. Typically 1-2 µL of enzyme (Roche) was added every day, for a few days until reaction was complete. Deglycosylated protein was evaluated on SDS-PAGE to ensure complete removal of glycan, as sample homogeneity is critical for crystallization. The protein was further purified by size exclusion chromatography (SEC) column Superdex 75 (GE healthcare). Monomeric fractions were effectively separated from aggregated materials (Fig 4). These fractions were pooled together and concentrated for crystallization. Initially, proteins were concentrated

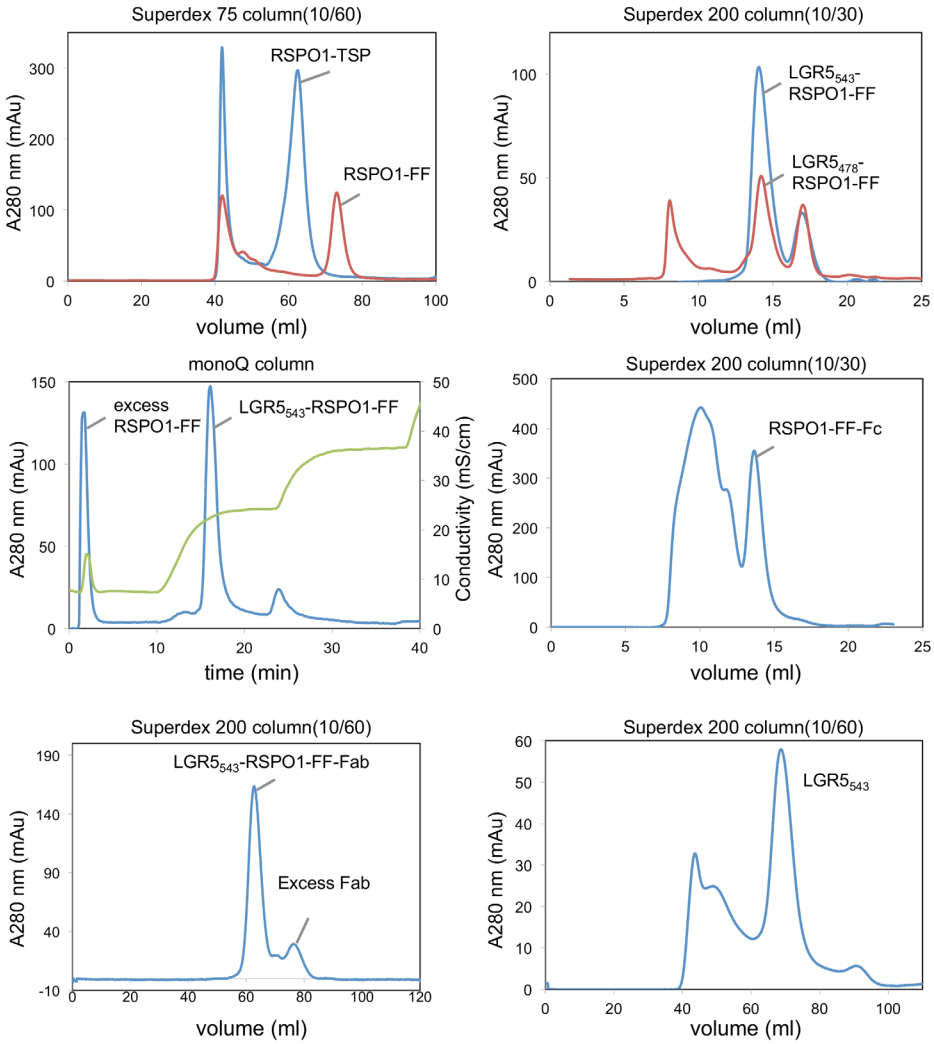


Fig 4. Purification of LGR5 and RSPO1. SEC and IEX chromatogram for large scale purification of LGR5-RSPO1 and other complexes are shown here.

using concentrators with MW cut off of 5 kDa (Sartorius) with polyethersulfone membrane, but often this procedure resulted in low recovery of protein (~20%). This is most likely due to adsorption of the protein to the membrane. Subsequently, concentrators with cellulose based Hydrosart membrane (Sartorius) were used and almost no loss of protein observed after concentration. RSPO1-TSP were harvested and purified following similar protocol as described for RSPO1-FF, except that in size exclusion chromatography a higher NaCl concentration (200 mM) was required to ensure high recovery from the SEC column (Fig 4).

$LGR5_{543}$-RSPO1-FF and $LGR5_{478}$-RSPO1-FF complexes were expressed in HEK293 ES cells. Typically, 10 µg of LGR5 plasmid and 2 µg of RSPO1-FF of plasmid, and 1 mg of dummy DNA were used for transfection of 2 L culture. Both complexes were purified by Ni-NTA column. Initially, $LGR5_{543}$-RSPO1-FF was incubated with $EndoH_F$ (NEB) for glycan removal, and with TEV protease for His_6 tag cleavage. Removal of glycan and His_6-tag resulted in massive precipitation. Subsequently, these steps were skipped and complexes were purified directly by SEC column Superdex 200 using a buffer with 25 mM HEPES pH 8.0 and 150 mM NaCl (Fig 4). We noticed that using a buffer at pH 7.5 (or 6.5 and 5.5), or at lower NaCl concentration (50 mM) often resulted in low recovery of protein from the SEC column.

However, LGR5-RSPO1 complexes purified from size exclusion chromatography usually contained trace amounts of unbound RSPO1, as judged from analytical gel filtration (not shown). Hence, $LGR5_{543}$-RSPO1-FF samples were purified by ion-exchange column (IEX) after the Ni-NTA column elution (Fig 4). Samples were injected into MonoQ (GE healthcare) in a buffer of 25 mM HEPES pH 8.0 and 50 mM NaCl. RSPO1-FF did not bind to column and was eluted in the flow through, whereas $LGR5_{543}$-RSPO1-FF complex was eluted at 200–250 mM NaCl. This purification step effectively separated the LGR5 complex from excess ligands. SDS-PAGE indicated > 95% sample purity. Samples were concentrated in elution buffer to 10–15 mg/ml for crystallization.

The expression yields for all the recombinant proteins were in the range of 5–10 mg/L HEK 293 (E/ES) culture after Ni-NTA elution, and after final step of purification (SEC or IEX), about 3-5 mg (~50%) of proteins were recovered. Purified recombinants proteins were > 90% pure as indicated by SDS-PAGE (Fig 3F). Chromatograms for purification of various complexes are shown in Figure 4.

Oligomerization studies and functional assay

Multiple Angle Laser Light Scattering (MALS) analysis indicated that LGR5 and RSPO1-FF formed a complex of 85 kDa, consistent with a 1:1 monomeric complex (buffer: 25 mM HEPES pH 8.0, 150 mM NaCl) (Fig 5A, left panel). In the absence of R-spondin, $LGR5_{543}$ is a monomer, as shown in the SEC chromatogram (Fig 5A, right panel) Since the TSP domains in properdin has been shown to oligomerize [50], we tested if RSPO-TSP or full-length R-spondin induce oligomerization of LGR5-RSPO1 complexes. SEC analysis indicated that these complexes remain 1:1 monomeric complexes (Fig 5A, right panel).

Next, we tested the activity of purified recombinant R-spondin proteins by Wnt reporter TOPFlash assay in HEK 293T cells. Both RSPO-TSP and RSPO-FF (glycosylated and deglycosylated forms) showed comparable levels of activity (Fig 5B), indicating that the Fu domains alone are sufficient for binding and mediating activity of LGR5, consistent with previous report [15]. Hence, the LGR5-RSPO1 complex that we used for crystallization is relevant for studying signaling mechanism.

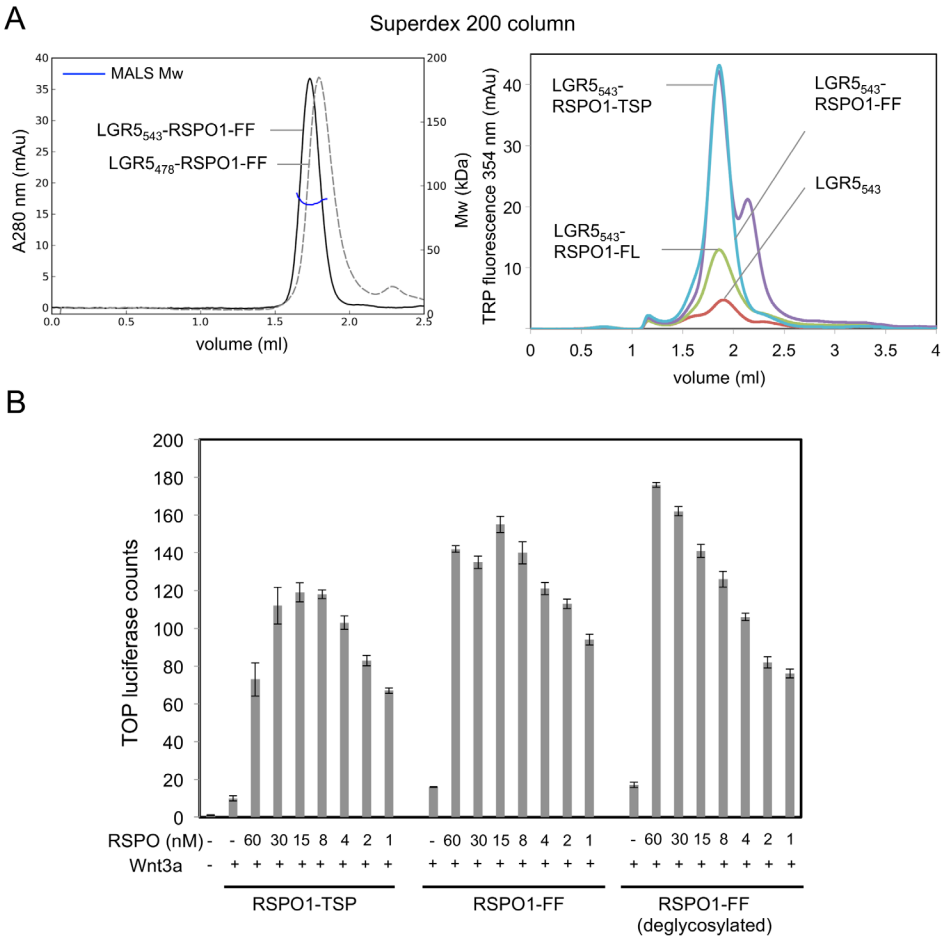


Fig 5. Oligomerization analysis of LGR5-RSPO1 complexes and functional analysis of RSPO1.
A, SEC and MALS analysis of various LGR5-RSPO1 complexes.
B, Wnt reporter assay for various RSPO1 constructs, performed in HEK 293T cells. Fig 5A (left panel) is adapted from chapter 2, Fig S3.

Screening and optimization of crystals

For crystallization screening, RSPO1-FF was initially concentrated to 6-7 mg/ml and 300 conditions were screened. Most drops (> 90%) were clear, indicating that the protein concentration was far from saturation. However, in 1.6 M $(NH_4)_2SO_4$, 0.1 M MES pH 6.5, 10% v/v dioxane (denoted condition A) many small and highly birefringent, hexogonal crystals were obtained (Fig 6A). These crystals were tested at ESRF ID23-1 beam and consistently diffracted to 2.0-2.5 Å resolution. In another condition, 0.2 M NaCl, 0.1 M phosphate-citrate pH 4.2, 20% w/v PEG 8000, needle clusters were observed (Fig 6B). These needles originated from a single nucleation spot. Crystals with similar morphology were also observed in another

related condition, 0.2 M LiCl 0.1 M citrate pH 4.0, 20% w/v PEG 6000 (data not shown). These crystals were easily reproduced, and sometimes appeared as many tiny needles scattered throughout the drop, depending on the concentration of precipitant. However, these crystals usually diffracted to ~20 Å, when tested at the synchrotron.

We attempted to produce more crystals for heavy atom soaking for experimental phasing. However, in the same condition we have always observed clear drops, indicating that the protein is far below the concentration required for saturation. Since the protocol for purification of RSPO1-FF was improved, more proteins could be obtained for rescreening crystallization at higher concentration: 12 mg/ml (1000 conditions tested), 20 mg/ml (300 conditions) and finally 40 mg/ml (400 conditions). None of these trials yielded any new crystal form, apart from needle clusters observed earlier.

After several failed attempts to reproduce the hexagonal crystals, we reasoned that the lack of globular structure, and presence of many loops in R-spondin, may have precluded proper crystal packing (as observed in needle clusters). Hence we designed a new construct: RSPO1-FF with Fc protein at the C-terminus. The rationale of this approach is that Fc may enhance protein-protein contacts required for crystal packing. The hinge region of the Fc protein was omitted to reduce the flexibility of the fusion protein. The fusion protein was expressed at high yield (> 10 mg/L from HEK 293E cells) and purified with SEC (Fig 4), concentrated to 10 mg/ml for crystallization, but did not yield any crystals. We observed that, although in most screens many drops (> 90%) were clear, in screens containing PEG as precipitant (e.g. PACT and PEGION screens), > 50% of the drops showed precipitation. Hence, we added a small amount of PEG 3350 to the original condition A, and observed crystal formation when 1-2% w/v PEG 3350 was added. It is likely that in the original condition the mother liquor was contaminated by traces of PEG.

For crystallization of RSPO-TSP, screening was performed at an initial concentration of 7.3 mg/ml (700 conditions). Since most of the drops (>90%) are clear, screening was performed at higher concentration: 45 mg/ml (300 conditions) and 80 mg/ml (300 conditions). Granular precipitate was observed (data not shown) and optimization did not yield crystals. This construct did not crystallize, presumably due to the flexibility between the Fu and TSP domain that may have hindered crystal-packing interactions.

For $LGR5_{543}$-RSPO1-FF, screening was performed at 15 mg/ml (800 conditions). An initial hit was obtained in 0.8 M $(NH_4)_2SO_4$, 0.1 M citrate pH 4.0. Microcrystalline precipitate was observed throughout the whole drop. By varying the pH from 3.5 to 9.0, and increasing the drop size to 1 µL or 2 µL (original drop size: 0.3 µL), we are able to obtain needle clusters at 0.1 M citrate pH 4.5. Since the formation of needle clusters is often associated with fast nucleation resulting in improper crystal packing, we set up crystallization drops at 4 °C, to reduce the nucleation rate. Here, we obtained isolated single hexagonal crystals in precipitant from 0.8 M-1.2 M $(NH_4)_2SO_4$ and 0.1 M citrate pH 4.5 (Fig 6J). The crystals were easily reproducible, and sometimes grew up to 0.2-0.3 µm in size. These crystals were screened

extensively at synchrotrons but diffraction was typically limited to 10 Å resolution. However, one crystal diffracted to 4.3 Å resolution. From Matthew analysis, this crystal form contained 67% volume solvent content. Protein crystals typically exhibit 27-67% solvent content, with an average of 43% [51]. A high-solvent content of a crystal usually indicates less protein-protein contacts when compared to crystals with low-solvent content, and is often associated with poor diffraction, due to poor crystal packing.

In another condition, 0.15 M $(NH_4)_2SO_4$, 0.1 M MES pH 6.0, 20% w/v PEG 4000, non-amorphous precipitate that appeared like 'crumble' was observed (Fig 6E). Closer inspection revealed that this precipitate consisted of clusters of tiny crystalline plates. We were able to obtain crystals by lowering the precipitant concentration to 8-12% w/v PEG 4000 (Fig 6H). These crystals were typically plate or rod-like with multiple crystals stacked onto each other, making the task of isolating them for diffraction studies difficult. Subsequent streak seeding or microseeding yielded many small plate-like crystals and needle clusters. Setting up crystallization drops at lower temperature (4 °C and 12 °C) did not yield isolated single crystals. Subsequently, the multiple crystals were carefully broken apart and tested at the synchrotron. These crystals diffracted to medium resolution (to ~4 Å resolution), with streaky spots and sometimes multiple spots due to overlapping crystal lattices. We managed to obtain a few datasets at 3-4 Å, with the best dataset integrated up to 3.2 Å resolution.

In addition, we performed rescreening with a novel crystallization screen (Molecular Dimension), which used poly-glutamic acid (PGA) polymer as a precipitant. In condition 0.3 M sodium malonate, 0.1 M cacodylate pH 6.5, 8% w/v PGA we observed solid objects that grew larger over time at the edge of the drop (Fig 6F). However, these objects did not appear to have defined shape, and were not translucent, as one would expect for crystals. However, upon optimization, single crystals were obtained at 4 °C (Fig 6I) and the best crystal diffracted to 3.2 Å resolution. For $LGR5_{478}$-RSPO-FF, non-amorphous precipitates were observed (Fig 6K) that did not yield any crystals after optimization.

In recent years, several labs have shown that Fab fragments and nanobodies can been used to enhance crystallization [52,53]. We have previously produced several antibodies that bind to the LRRCT of LGR5 to activate Wnt signaling (described in chapter 2). Fab from antibody '1D9' was generated by papain digestion and the ternary complex of LGR5-RSPO1-Fab was co-purified by SEC (Fig 4). We set up co-crystallization screens at 10 mg/ml and obtained crystals, which diffracted to 2 Å resolution. Unfortunately, these crystals contained only Fab (data not shown).

Structure determination of LGR5-RSPO1 complex by Molecular Replacement

We obtained a native diffraction dataset at 4.3 Å resolution from a crystal that belonged to space group $P6_122$ (Fig 6J). For structure determination, Molecular Replacement (MR) strategy can be used when homologous structures are available. In general, homologous models with sequence identity >30% and with a RMSD within 2.0 Å (from the target structure) are usually

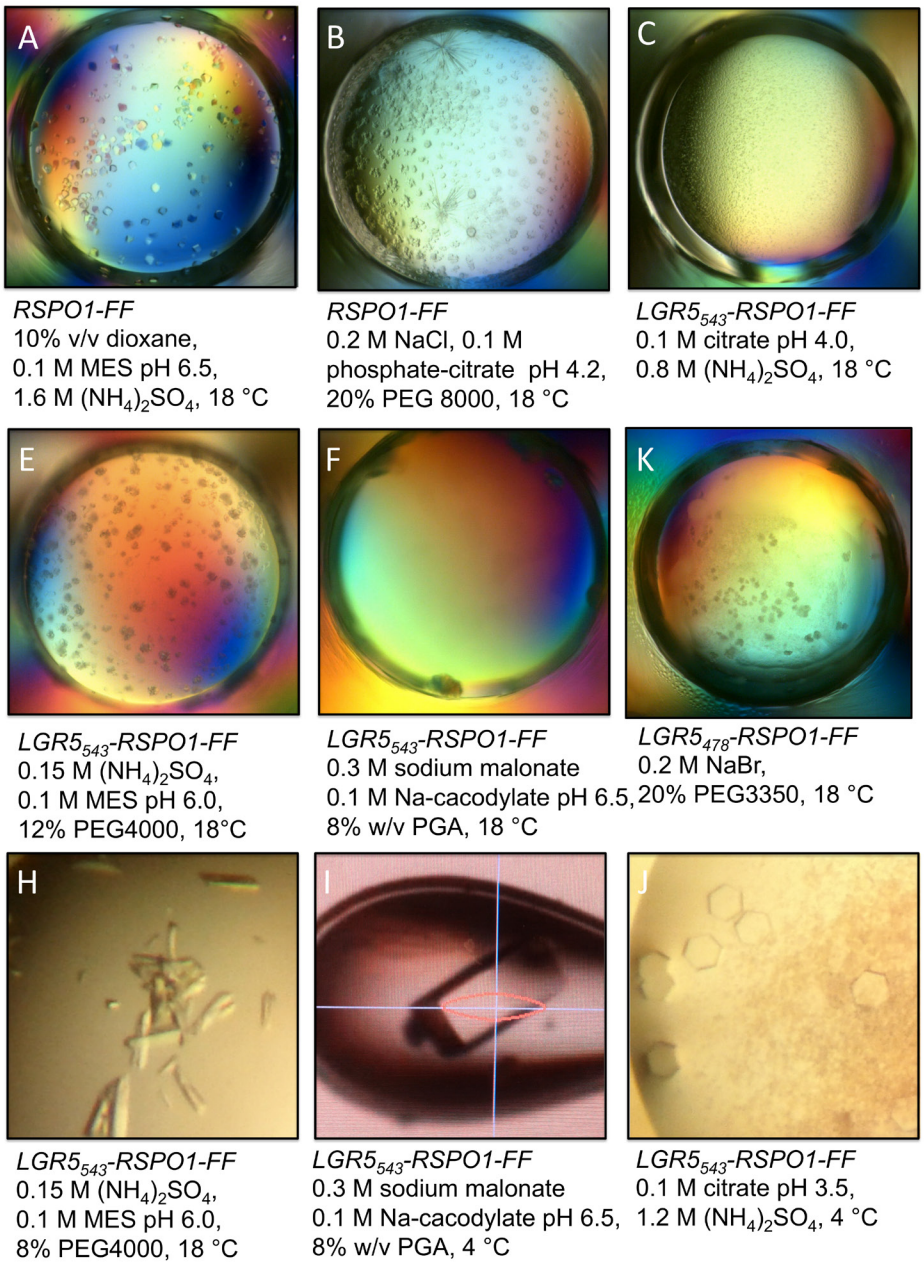


Fig 6. A-J, Crystallization of LGR5-RSPO1 and RSPO1 in conditions as indicated. Crystals grown in conditions H, I, and J were optimized from conditions E, F, and C, respectively.

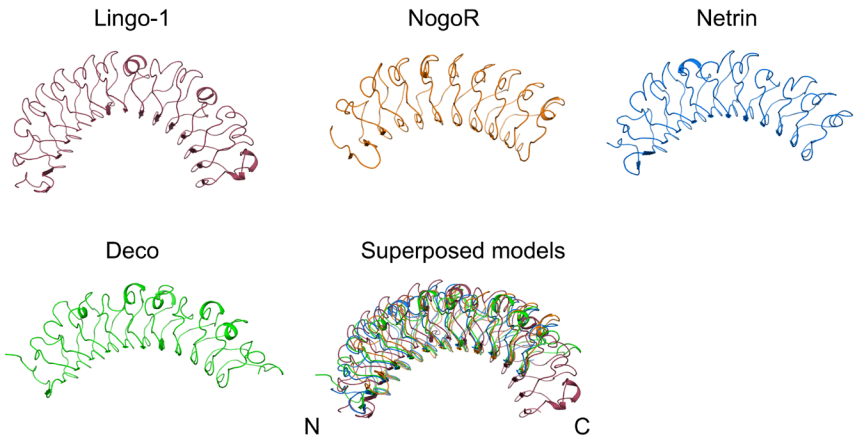
required for successful MR solutions [54,55]. The structures of FSHR and TSHR (LGR5 family class A) are composed of 9 LRRs and show only 10% sequence identity to LGR5 (LGR family class B) with 17 LRRs, suggesting strong sequence divergence and possibly large structural deviation between the two classes of receptors. In general, LRR domains are assembled from tandem repeats of LRR (20-30 residues) that form a curved solenoid structure [33], with different curvatures and twists, resulting in variation of tertiary structures. Hence, we reasoned that non-homologous structures can be used as search models, if they have tertiary structure resembling LGR5. We performed model search using protein-blast software (NCBI) against all structures in PDB, and did not obtain any hits with 17 LRRs. However, we identified some LRR domains that matched the N-terminal half of LGR5 (consisting of 10 to 12 LRRs), with 22-24 residues in each repeat, matching those of LGR5. The search models that we selected based on high sequence identity (>30% of the aligned region) were Lingo-1, Netrin-G ligand 3 (termed 'Netrin' here), Decorin core protein (Deco) and Nogo receptor (NogoR), shown in Fig 7A. These domains contain 9-12 LRRs with 30-35% sequence identity. The LRRCT regions, when present in the search models, were truncated.

We performed MR using the crystallographic CCP4 software PHASER [55]. A correct MR solution is usually a single solution with rotation function z-score (RFZ) higher than 5 and translation function z-score (TFZ) higher than 8 [54]. Log-likelihood gain (LLG) scores are used to rank solutions with higher LLG scores indicating better solutions. A large gap is expected to be present between the correct and wrong solutions [54]. However, these parameters are only indicators of the quality of solutions, and a right solution can only be verified by inspecting electron density after performing many cycles of building and refinement.

The Matthew analysis indicated that 2-4 copies of LGR5 were present in the asymmetric unit (ASU) based on solvent content analysis [51]. Hence, we searched for 2 copies of LGR5. Search with Lingo-1 produced 38 solutions (three highest LLG: 73, 73, 73), NogoR produced 22 solutions (LLG 94, 93, 81), Netrin produced 2 solutions (LLG 155, 153), and Deco produced 3 solutions (LLG 212, 200, 184). These solutions are shown in Fig 8A. In all cases RFZ and TFZ scores are 2.5-3.3 and 4.7-9.2, respectively. At this stage, with many solutions generated it is hard to identify the right solution based on RFZ, TFZ and LLG scores, which showed only small gaps between different solutions. Interestingly, by inspecting Deco and Netrin solutions, we observed an identical 'head-to-head' (N-N) arrangement in all 5 solutions. Superposition of these solutions showed minor shift in positions between different solutions (Fig 8A). Crystal packing of the two Decorin molecules (Fig 8B) showed that there was enough space to accommodate the C-terminal of LGR5, indicating such crystal packing would be plausible. To improve the electron density map and corresponding phase information, we performed density modification with PARROT for Deco [56], the solution with the highest LLG of all. Strikingly, the resulting electron density map revealed additional contiguous density for the unmodelled C-terminal half of LGR5 (Fig 8B), which is a good indication that the MR solutions was indeed correct. From the density it appeared that the C-terminal half of LGR5 extended towards each other. By searching for Deco-dimer, as shown in the rectangular box (Fig 8B), we obtained a single solution with RFZ 9.1, TFZ 32.4 and LLG of 867, which can be refined in

A

Protein	PDB	Domain	Sequence identity of matched region
Lingo-1	2ID5	LRRNT-LRR12	29%
Netrin-G ligand 3 (Netrin)	3ZYN	LRRNT-LRR10	34%
Nogo Receptor (NogoR)	3KJ4	LRRNT-LRR9	35%
Decorin (Deco)	1XKU	LRRNT-LRR10	30%



B

Protein	PDB	Domain	Sequence identity of matched region
Scaffold protein	3RFS	6LRR	34%
Nogo receptor (nogoR)	1P8T	7LRR	30%
Engineered protein (NSGC)	4PSJ	6LRR	35%
Engineered protein (NSGC)	4PQ8	7LRR	33%

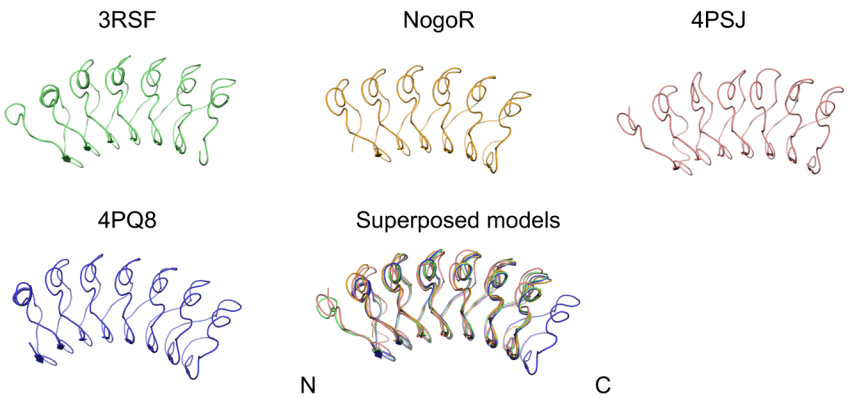


Fig 7. Search models of LGR5 ectodomain. A, N-terminal half and B, C-terminal half.

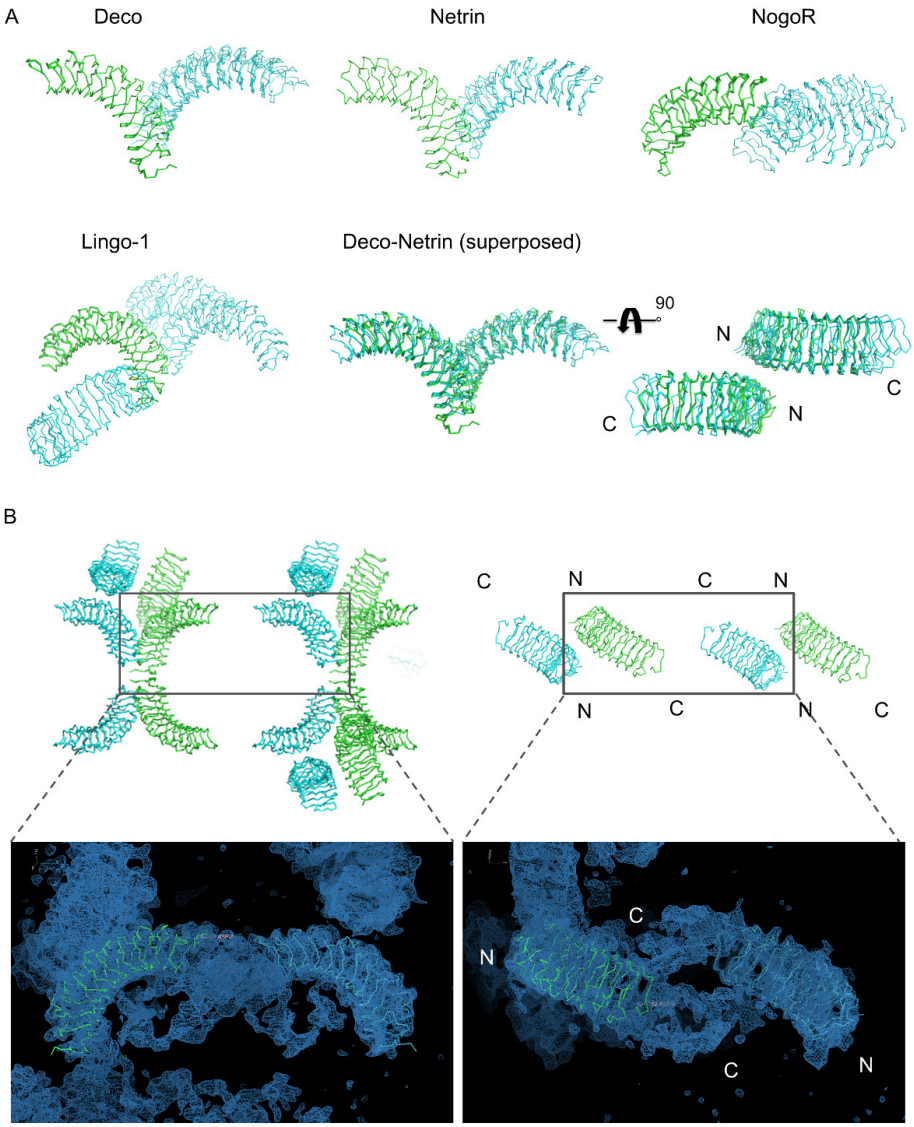


Fig 8. Analysis of MR solutions.
A, MR solutions using various search models.
B, Crystal packing for Deco molecules, the top solution with highest LLG is shown here. Rectangular box indicates asymetric unit (ASU) that contains 2 copies of search model (upper). Density modified map contoured at 1.5 σ level (lower). N and C indicates N- and C-termini, respectively.

REFMAC5 [57] to an R-free of 0.46 (while an R-free of 0.54 indicate random solution).

With the N-terminal half of LGR5 correctly placed, we proceeded to search for the C-terminal half of LGR5 (LRR 11-LRRCT). By protein-blast against residues 280-543, we identified several search models: (i) designed scaffold protein 3RFS, (ii) nogo Receptor, (iii) engineered protein 4PSJ; and, (iv) 4PQ8. 4PSJ and 4PQ8 are the same protein (with 8 LRRs); 4PSJ is truncated to 6 LRRs, while 4PQ8 is truncated to 7 LRRs. The LRRNT, present in these models, were truncated (Fig 7B).

Here we performed MR search for 2 copies of the molecules. The search with 3RFS (6 LRRs) produced a single solution, however with only 1 copy placed (RFZ 1, TFZ 7.6, LLG 121), see Fig 9. Initial search for 2 copies of NogoR produced no solution. Therefore the threshold for clashes was relaxed to 25%, i.e. allowing for 25% of NogoR residues (Cα atoms) to clash with the Deco molecules, which was already placed in the ASU. This strategy produced 3 solutions, each with only 1 copy of NogoR placed. By inspecting the solution, we observed that the 3 solutions (with LLG 126, 104, 102) were placed in the same area and that these solutions yielded plausible connections to the Deco model. The solutions showed minor shift in the positioning of LRRs. The first, second and third solutions have 1, 2 and 3 LRRs overlapped with the last LRRs of Deco molecules, respectively. The search with 4PSJ (6LRRs) produced 5 solutions, with 2 copies placed. The top two solutions were plausible solutions, i.e. they were connected to the Deco molecules, while the third solutions had 1 copy placed away from the Deco molecules, indicating a wrong solution. The search with 4PQ8 (7 LRRs) produced a single solution with 2 copies placed (RFZ 0.9, TFZ 9.9, RFZ 1.2, TFZ 8.8, LLG 225). Both copies of molecules were placed such that they are connected to the Deco models (Fig 9).

Superposition of solutions from 4PQ8 and NogoR showed that they were placed in the same area, with the C-terminal part of the models extended towards each other, creating an 'embracing dimer' (Fig 9). The best solution from Deco-4PQ8 (LLG 225) could be refined to an R-free of 0.42, which improved from R-free of 0.46 previously, indicating that this was likely the right solution. The MR search strategy allowing 25% of clashes (default is 5% of Cα atoms) was crucial to ensure that the right solutions were not rejected during packing test. A hybrid model was generated by combining Deco and 4PQ8, creating a model containing 16 LRRs, with a noticeable kink between LRR10-LRR11 (Fig 9).

Subsequently, we obtained crystals (Fig 6I) that diffracted to 3.2 Å resolution and belonged to space group $P22_12_1$. The hybrid-model was used to search for 2 copies of molecule, and PHASER unambiguously placed the two molecules with a LLG of 910 (1st copy: RFZ 3.3, TFZ 9.9, 2nd copy: RFZ 3.6, TFZ 9.4). The two molecules form an 'embracing dimer', identical to the dimer that we observed in earlier searches. This solution was further refined in REFMAC5 to an R-free of 0.43. The resulting electron density map showed that the 16 LRRs are placed in the correct position (Fig 10). In general, the β-strands on the concave side placed fitted well within the electron density, indicating that 'actual' LGR5 structure did not deviate much from standard LRRs model. The convex side showed more deviations from the density, which

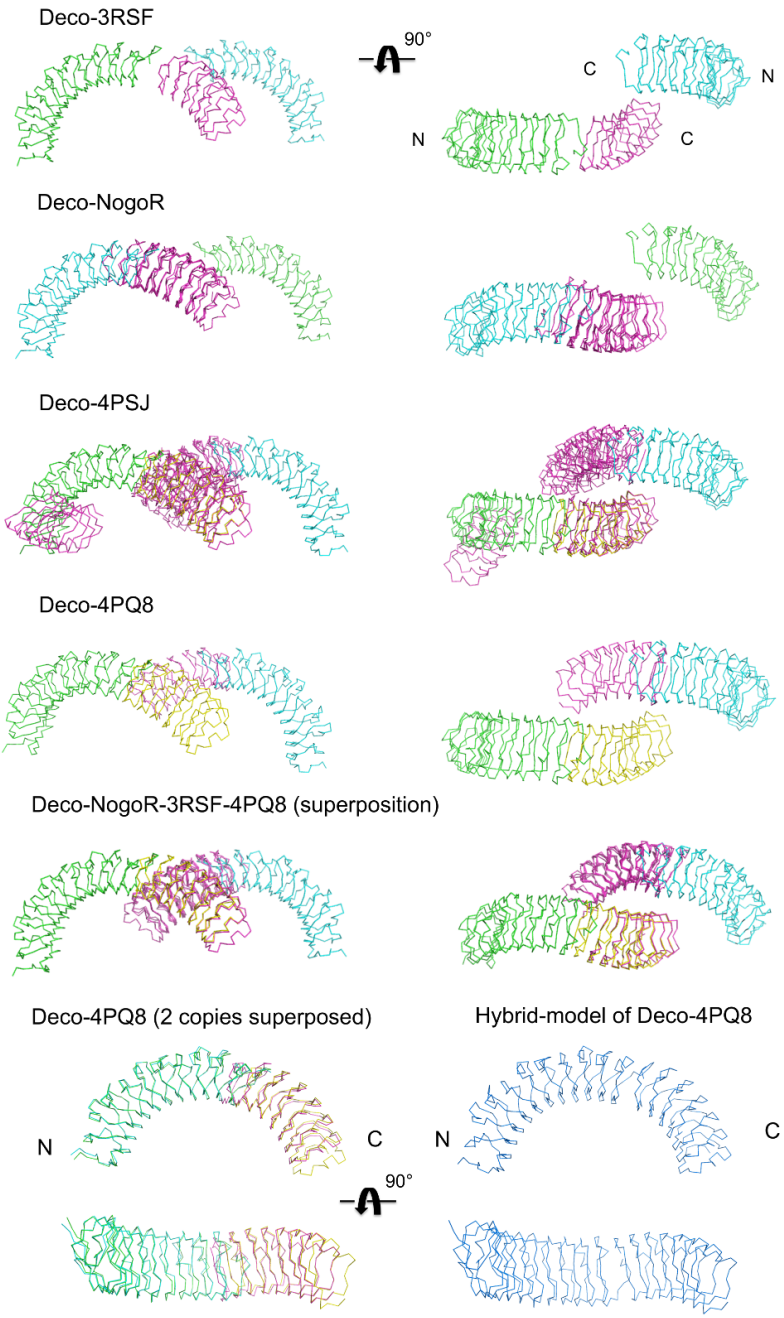


Fig 9. MR solutions for the C-terminal half of LGR5 ectodomain, using various search models. The hybrid-model of Deco-4PQ8 used for subsequent MR is also shown here in the last panel (right).

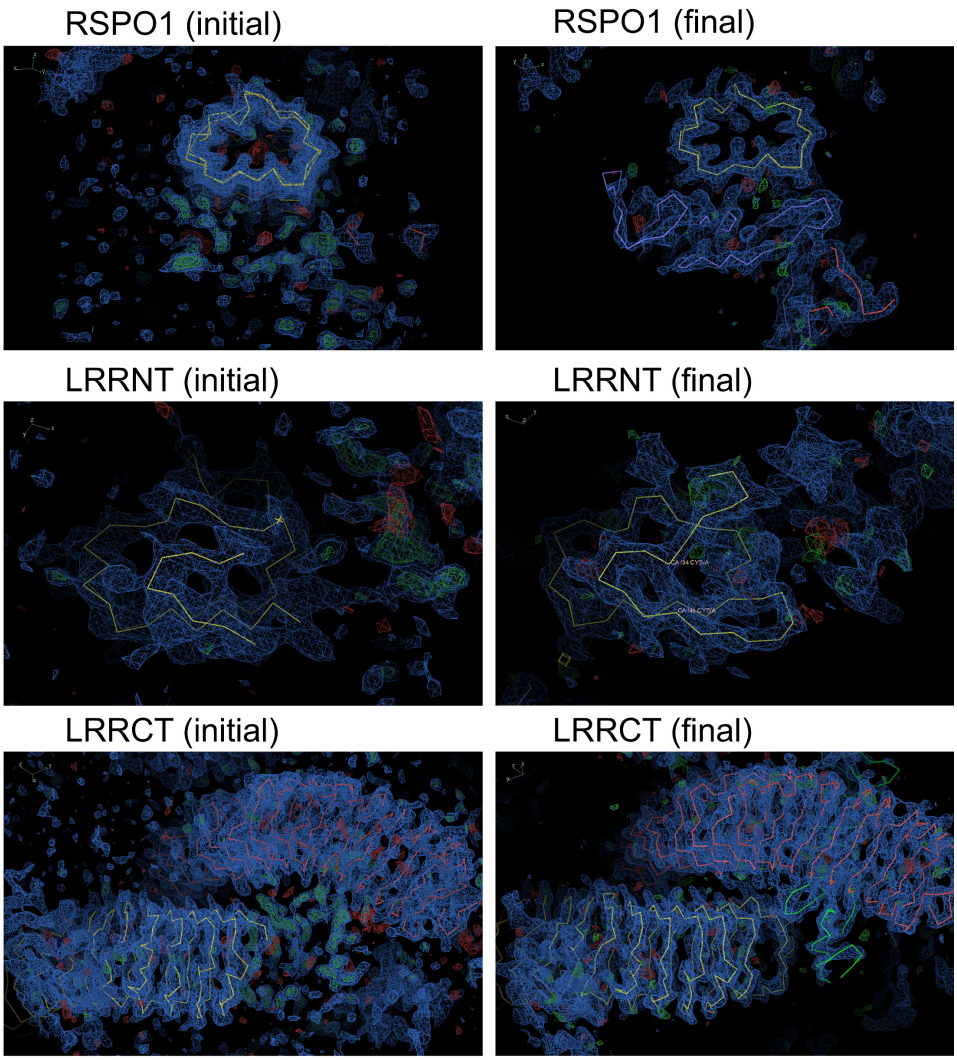


Fig 10. Electron density maps for LGR5-RSPO1 complex. Electron density maps after MR, with Deco-4PQ8 model placed, are shown in the left panel. For comparison, final electron density maps for LGR5-RSPO1, after many cycles of tracing and refinement, are shown in the right panel.
Electron density maps (2mFo-DFc) are shown in blue contoured at 1.2 σ level. mFo-DFc density maps are shown in green and red, contoured at 3.0 σ level. Green density (where density is positive) indicates where model is missing and red density (where density is negative) indicates where model should not be present.

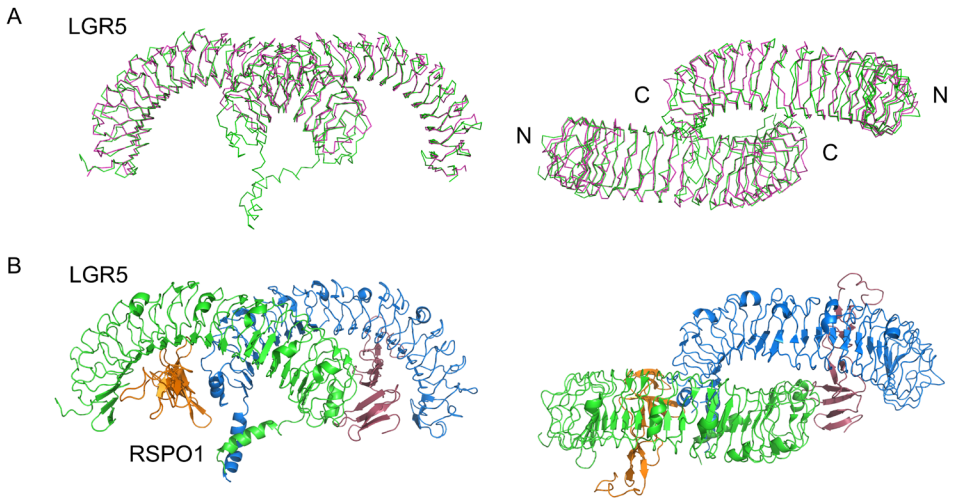


Fig 11. Structure of LGR5-RSPO1 and Deco-4PQ8 model.
A, Superposition of Deco-4PQ8 model and LGR5. Cα traces are shown in ribbon representation.
B, Cartoon representation of LGR5-RSPO1 complex in two orthogonal views.

was expected since they are composed of short disordered loops and the sequences are less conserved. In addition, we could identify extra density that corresponded to LRRNT and LRRCT, e.g. a disulfide bond in the LRRNT could be traced here, which was not present in the search model (Fig 10). At this stage, many cycles of manual tracing (i.e. placing the right amino acids in the electron density) were performed to modify the starting model (Deco-4PQ8) to the actual residues of LGR5. With each cycle of rebuilding and refinement, the model became closer to the actual structure and the resulting electron density map improved. LRRNT and LRRCT were modeled according to the new FSHR structure (PDB 4MQW) [58], which became available during our refinement stage (the initial FSHR structure did not include LRRCT). Although density for R-spondin became visible, manual tracing of R-spondin was difficult due to lack of a homologous model. As can be observed in the final structure, R-spondin does not have a globular structure, but contains many loops and short strands held together by disulfide bonds, which made chain tracing difficult.

We crystallized RSPO1-FF that diffracted to 2.0 Å resolution and belonged to space group I222 (Fig 6A). Initially, we attempted to determine the structure of RSPO1-FF by MR, using a search model comprised of Fu domains, e.g. based on PDB 1IGR and 1EVO. The MR searches, using 1IGR and 1EVO were not successful, possibly due to large structural deviation from the actual structure given the low sequence identity of 10-15% The lack of secondary structure, such as helices, also makes the MR search more difficult.

After extensive trials and unsuccessful attempts to solve the R-spondin structure with MR,

we managed to reproduce R-spondin crystals for experimental phasing experiments. An attempt to determine R-spondin structure by sulfur SAD phasing, which exploits the intrinsic anomalous scattering of native proteins [59,60], was not successful. Subsequently, crystals were soaked overnight in heavy atoms, i.e. 5 mM of holmium chloride ($HoCl_3$), samarium chloride ($SmCl_3$), ytterbium chloride ($YbCl_3$) or gadolinium chloride ($GdCl_3$). Of all the heavy atoms tested, only $HoCl_3$- and $GdCl_3$-soaked crystals displayed anomalous scattering, indicating that the heavy atoms were successfully bound to the protein molecules. The heavy-atom soaked crystal diffracted to ~2.0 Å resolution. The structure determination of R-spondin is described in chapter 2. The R-spondin structure was placed into LGR5 structure by MR [55]. A 90° rotation in Fu2 domain, was observed in the LGR5 bound structure and was manually adjusted to fit into the density (described in chapter 2, Fig 2 and Fig S11). The resulting LGR5-R-spondin was refined to an R-work of 0.23 and R-free of 0.26. An overlay of the Cα traces of Deco-4PQ8 model and LGR5 structure (RMSD 4.5 Å) are shown in Fig 11A and the 2:2 dimeric complexes of LGR5-RSPO1 is shown in Fig 11B. The data collection and refinement statistics are shown in chapter 2, Supplementary Table 1. Structural analysis and biological implications of LGR5-RSPO1 were described in chapter 2.

CONCLUDING REMARKS

The production of protein in large quantity (~miligrams) is often required for structural study and therapeutic development and applications [61]. However, overexpression of recombinant proteins in mammalian system is not always straightforward and may result in aggregation of proteins [31]. Here, we showed that optimization strategy, such as co-expression of LGR5-RSPO1 complex and 'plasmid titration' improved yield, stability and solubility of LGR5. The availability of large amount of LGR5-RSPO1 allowed us to perform extensive crystallization screening, biochemical study and antibody generation. In addition, co-expressions of high-affinity (K_D of nanomolar) complexes of secreted proteins by transient expression in HEK 293 cells are now routinely used in our lab to improve yield and stability of recombinant proteins.

In addition, we present here our method for structure determination of LGR5 with MR, using non-homologous LRR models. We demonstrate that, fragments of LRR proteins with similar twist and curvature, can be used as partial models for successful MR, despite showing low sequence identity (i.e. partial models of ~30% sequence identity matching half of the target structure). This method eliminates the needs for cumbersome heavy atoms soaking experiments required for experimental phasing. The success of LGR5 case study suggests that this approach may be suitable for proteins consisting of structural repeat elements like the LRR proteins.

MATERIAL AND METHODS

Cloning of constructs and transient expressions of proteins in HEK 293 E and ES cells were performed as described previously [31,62]. All protein purification steps were performed on AKTA systems (GE healthcare) or Prominence UFLC system (Shimadzu). All columns were

purchased from GE healthcare. MALS analysis was performed with Minidawn Treos detector (Wyeth). Wnt reporter assays were performed according to protocols described earlier [62,63]. Crystallization screens were purchased from Molecular Dimension or Hampton Research.

ACKNOWLEDGEMENTS

We are grateful to Eric Huizinga and Utrecht-Protein Express (U-PE BV) for helps in constructs design and performing HEK 293 cells expressions; Wim de Lau and Hans Clevers for their collaboration, advice and performing Wnt reporter assays; beamline scientists at ESRF, Grenoble and SLS, Villigen for their assistance during data collection.

REFERENCES

1. Hsu SY, Liang SG, Hsueh AJ (1998) Characterization of two LGR genes homologous to gonadotropin and thyrotropin receptors with extracellular leucine-rich repeats and a G protein-coupled, seven-transmembrane region. Mol Endocrinol 12: 1830–1845.
2. Vassart G, Pardo L, Costagliola S (2004) A molecular dissection of the glycoprotein hormone receptors. Trends Biochem Sci 29: 119–126. doi:10.1016/j.tibs.2004.01.006.
3. Pierce JG, Parsons TF (1981) Glycoprotein hormones: structure and function. Annu Rev Biochem 50: 465–495. doi:10.1146/annurev.bi.50.070181.002341.
4. Barker N, van Es JH, Kuipers J, Kujala P, van den Born M, et al. (2007) Identification of stem cells in small intestine and colon by marker gene Lgr5. Nature 449: 1003–1007. doi:10.1038/nature06196.
5. Barker N (2014) Adult intestinal stem cells: critical drivers of epithelial homeostasis and regeneration. Nat Rev Mol Cell Biol 15: 19–33. doi:10.1038/nrm3721.
6. Cheng H, Leblond CP (1974) Origin, differentiation and renewal of the four main epithelial cell types in the mouse small intestine. V. Unitarian Theory of the origin of the four epithelial cell types. Am J Anat 141: 537–561. doi:10.1002/aja.1001410407.
7. Sato T, Vries RG, Snippert HJ, van de Wetering M, Barker N, et al. (2009) Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche. Nature 459: 262–265. doi:10.1038/nature07935.
8. Sato T, Clevers H (2013) Growing self-organizing mini-guts from a single intestinal stem cell: mechanism and applications. Science 340: 1190–1194. doi:10.1126/science.1234852.
9. Snippert HJ, Haegebarth A, Kasper M, Jaks V, van Es JH, et al. (2010) Lgr6 marks stem cells in the hair follicle that generate all cell lineages of the skin. Science 327: 1385–1389. doi:10.1126/science.1184733.
10. Barker N, Huch M, Kujala P, van de Wetering M, Snippert HJ, et al. (2010) Lgr5(+ve) stem cells drive self-renewal in the stomach and build long-lived gastric units in vitro. Cell Stem Cell 6: 25–36. doi:10.1016/j.stem.2009.11.013.
11. Barker N, Rookmaaker MB, Kujala P, Ng A, Leushacke M, et al. (2012) Lgr5(+ve) Stem/Progenitor Cells Contribute to Nephron Formation during Kidney Development. Cell Rep 2: 540–552. doi:10.1016/j.celrep.2012.08.018.
12. Huch M, Bonfanti P, Boj SF, Sato T, Loomans CJM, et al. (2013) Unlimited in vitro expansion

of adult bi-potent pancreas progenitors through the Lgr5/R-spondin axis. EMBO J 32: 2708–2721. doi:10.1038/emboj.2013.204.
13. Huch M, Dorrell C, Boj SF, van Es JH, Li VSW, et al. (2013) In vitro expansion of single Lgr5+ liver stem cells induced by Wnt-driven regeneration. Nature 494: 247–250. doi:10.1038/nature11826.
14. Plaks V, Brenot A, Lawson DA, Linnemann JR, Van Kappel EC, et al. (2013) Lgr5-expressing cells are sufficient and necessary for postnatal mammary gland organogenesis. Cell Rep 3: 70–78. doi:10.1016/j.celrep.2012.12.017.
15. Kazanskaya O, Glinka A, del Barco Barrantes I, Stannek P, Niehrs C, et al. (2004) R-Spondin2 is a secreted activator of Wnt/beta-catenin signaling and is required for Xenopus myogenesis. Dev Cell 7: 525–534. doi:10.1016/j.devcel.2004.07.019.
16. Kim K-A, Kakitani M, Zhao J, Oshima T, Tang T, et al. (2005) Mitogenic influence of human R-spondin1 on the intestinal epithelium. Science 309: 1256–1259. doi:10.1126/science.1112521.
17. Wei Q, Yokota C, Semenov MV, Doble B, Woodgett J, et al. (2007) R-spondin1 is a high affinity ligand for LRP6 and induces LRP6 phosphorylation and beta-catenin signaling. J Biol Chem 282: 15903–15911. doi:10.1074/jbc.M701927200.
18. Binnerts ME, Kim K-A, Bright JM, Patel SM, Tran K, et al. (2007) R-Spondin1 regulates Wnt signaling by inhibiting internalization of LRP6. Proc Natl Acad Sci USA 104: 14700–14705. doi:10.1073/pnas.0702305104.
19. de Lau W, Barker N, Low TY, Koo B-K, Li VSW, et al. (2011) Lgr5 homologues associate with Wnt receptors and mediate R-spondin signalling. Nature 476: 293–297. doi:10.1038/nature10337.
20. Carmon KS, Gong X, Lin Q, Thomas A, Liu Q (2011) R-spondins function as ligands of the orphan receptors LGR4 and LGR5 to regulate Wnt/beta-catenin signaling. Proc Natl Acad Sci USA 108: 11452–11457. doi:10.1073/pnas.1106083108.
21. Glinka A, Dolde C, Kirsch N, Huang Y-L, Kazanskaya O, et al. (2011) LGR4 and LGR5 are R-spondin receptors mediating Wnt/β-catenin and Wnt/PCP signalling. EMBO Rep 12: 1055–1061. doi:10.1038/embor.2011.175.
22. Ruffner H, Sprunger J, Charlat O, Leighton-Davies J, Grosshans B, et al. (2012) R-Spondin potentiates Wnt/β-catenin signaling through orphan receptors LGR4 and LGR5. PLoS ONE 7: e40976. doi:10.1371/journal.pone.0040976.
23. de Lau W, Peng WC, Gros P, Clevers H (2014) The R-spondin/Lgr5/Rnf43 module: regulator of Wnt signal strength. Genes Dev 28: 305–316. doi:10.1101/gad.235473.113.
24. Kemper K, Prasetyanti PR, de Lau W, Rodermond H, Clevers H, et al. (2012) Monoclonal Antibodies Against Lgr5 Identify Human Colorectal Cancer Stem Cells. Stem Cells. doi:10.1002/stem.1233.
25. Hirsch D, Barker N, McNeil N, Hu Y, Camps J, et al. (2014) LGR5 positivity defines stem-like cells in colorectal cancer. Carcinogenesis 35: 849–858. doi:10.1093/carcin/bgt377.
26. Schepers AG, Snippert HJ, Stange DE, van den Born M, van Es JH, et al. (2012) Lineage tracing reveals Lgr5+ stem cell activity in mouse intestinal adenomas. Science 337: 730–735. doi:10.1126/science.1224676.
27. Barker N, Ridgway RA, van Es JH, van de Wetering M, Begthel H, et al. (2009) Crypt

stem cells as the cells-of-origin of intestinal cancer. Nature 457: 608–611. doi:10.1038/nature07602.
28. Krausova M, Korinek V (2014) Wnt signaling in adult intestinal stem cells and cancer. Cell Signal 26: 570–579. doi:10.1016/j.cellsig.2013.11.032.
29. Hong M, Yoon S-I, Wilson IA (2012) Recombinant expression of TLR5 proteins by ligand supplementation and a leucine-rich repeat hybrid technique. Biochem Biophys Res Commun 427: 119–124. doi:10.1016/j.bbrc.2012.09.021.
30. Fan QR, Hendrickson WA (2005) Structure of human follicle-stimulating hormone in complex with its receptor. Nature 433: 269–277. doi:10.1038/nature03206.
31. Halff EF, Versteeg M, Brondijk THC, Huizinga EG (2014) When less becomes more: Optimization of protein expression in HEK293-EBNA1 cells using plasmid titration - A case study for NLRs. Protein Expr Purif 99C: 27–34. doi:10.1016/j.pep.2014.03.010.
32. Letunic I, Doerks T, Bork P (2011) SMART 7: recent updates to the protein domain annotation resource. Nucleic Acids Res 40: D302–D305. doi:10.1093/nar/gkr931.
33. Bella J, Hindle KL, McEwan PA, Lovell SC (2008) The leucine-rich repeat structure. Cell Mol Life Sci 65: 2307–2333. doi:10.1007/s00018-008-8019-0.
34. Matsushima N, Miyashita H (2012) Leucine-Rich Repeat (LRR) Domains Containing Intervening Motifs in Plants. Biomolecules 2: 288–311. doi:10.3390/biom2020288.
35. Sanders P, Young S, Sanders J, Kabelis K, Baker S, et al. (2011) Crystal structure of the TSH receptor (TSHR) bound to a blocking-type TSHR autoantibody. J Mol Endocrinol 46: 81–99. doi:10.1530/JME-10-0127.
36. Buchan DWA, Minneci F, Nugent TCO, Bryson K, Jones DT (2013) Scalable web services for the PSIPRED Protein Analysis Workbench. Nucleic Acids Res 41: W349–W357. doi:10.1093/nar/gkt381.
37. de Lau WBM, Snel B, Clevers HC (2012) The R-spondin protein family. Genome Biol 13: 242. doi:10.1186/gb-2012-13-3-242.
38. Kim K-A, Wagle M, Tran K, Zhan X, Dixon MA, et al. (2008) R-Spondin family members regulate the Wnt pathway by a common mechanism. Mol Biol Cell 19: 2588–2596. doi:10.1091/mbc.E08-02-0187.
39. Ota T, Suzuki Y, Nishikawa T, Otsuki T, Sugiyama T, et al. (2004) Complete sequencing and characterization of 21,243 full-length human cDNAs. Nat Genet 36: 40–45. doi:10.1038/ng1285.
40. Parma P, Radi O, Vidal V, Chaboissier MC, Dellambra E, et al. (2006) R-spondin1 is essential in sex determination, skin differentiation and malignancy. Nat Genet 38: 1304–1309. doi:10.1038/ng1907.
41. Garrett TP, McKern NM, Lou M, Frenkel MJ, Bentley JD, et al. (1998) Crystal structure of the first three domains of the type-1 insulin-like growth factor receptor. Nature 394: 395–399. doi:10.1038/28668.
42. Ogiso H, Ishitani R, Nureki O, Fukai S, Yamanaka M, et al. (2002) Crystal structure of the complex of human epidermal growth factor and receptor extracellular domains. Cell 110: 775–787.
43. Tan K, Duquette M, Liu J-H, Dong Y, Zhang R, et al. (2002) Crystal structure of the TSP-1 type 1 repeats: a novel layered fold and its biological implication. J Cell Biol 159: 373–382.

doi:10.1083/jcb.200206062.
44. Hadders MA, Bubeck D, Roversi P, Hakobyan S, Forneris F, et al. (2012) Assembly and regulation of the membrane attack complex based on structures of C5b6 and sC5b9. Cell Rep 1: 200–207. doi:10.1016/j.celrep.2012.02.003.
45. Durocher Y, Perret S, Kamen A (2002) High-level and high-throughput recombinant protein production by transient transfection of suspension-growing human 293-EBNA1 cells. Nucleic Acids Res 30: E9.
46. Pham PL, Kamen A, Durocher Y (2006) Large-scale transfection of mammalian cells for the fast production of recombinant protein. Mol Biotechnol 34: 225–237. doi:10.1385/MB:34:2:225.
47. Durocher Y, Perret S, Thibaudeau E, Gaumond MH, Kamen A, et al. (2000) A reporter gene assay for high-throughput screening of G-protein-coupled receptors stably or transiently expressed in HEK293 EBNA cells grown in suspension culture. Anal Biochem 284: 316–326. doi:10.1006/abio.2000.4698.
48. Reeves PJP, Callewaert NN, Contreras RR, Khorana HGH (2002) Structure and function in rhodopsin: high-level expression of rhodopsin with restricted and homogeneous N-glycosylation by a tetracycline-inducible N-acetylglucosaminyltransferase I-negative HEK293S stable mammalian cell line. Proc Natl Acad Sci USA 99: 13419–13424. doi:10.1073/pnas.212519299.
49. Land A, Zonneveld D, Braakman I (2003) Folding of HIV-1 envelope glycoprotein involves extensive isomerization of disulfide bonds and conformation-dependent leader peptide cleavage. FASEB J 17: 1058–1067. doi:10.1096/fj.02-0811com.
50. Alcorlo M, Tortajada A, Rodríguez de Córdoba S, Llorca O (2013) Structural basis for the stabilization of the complement alternative pathway C3 convertase by properdin. Proceedings of the National Academy of Sciences 110: 13504–13509. doi:10.1073/pnas.1309618110.
51. Matthews BW (1968) Solvent content of protein crystals. J Mol Biol 33: 491–497.
52. Griffin L, Lawson A (2011) Antibody fragments as tools in crystallography. Clin Exp Immunol 165: 285–291. doi:10.1111/j.1365-2249.2011.04427.x.
53. Pardon E, Laeremans T, Triest S, Rasmussen SGF, Wohlkönig A, et al. (2014) A general protocol for the generation of Nanobodies for structural biology. Nat Protoc 9: 674–693. doi:10.1038/nprot.2014.039.
54. Oeffner RD, Bunkóczi G, McCoy AJ, Read RJ (2013) Improved estimates of coordinate error for molecular replacement. Acta Crystallogr D Biol Crystallogr 69: 2209–2215. doi:10.1107/S0907444913023512.
55. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, et al. (2007) Phaser crystallographic software. J Appl Crystallogr 40: 658–674. doi:10.1107/S0021889807021206.
56. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, et al. (2011) Overview of the CCP4 suite and current developments. Acta Crystallogr D Biol Crystallogr 67: 235–242. doi:10.1107/S0907444910045749.
57. Murshudov GN, Skubák P, Lebedev AA, Pannu NS, Steiner RA, et al. (2011) REFMAC5 for the refinement of macromolecular crystal structures. Acta Crystallogr D Biol Crystallogr

67: 355–367. doi:10.1107/S0907444911001314.
58. Jiang X, Liu H, Chen X, Chen P-H, Fischer D, et al. (2012) Structure of follicle-stimulating hormone in complex with the entire ectodomain of its receptor. Proc Natl Acad Sci USA 109: 12491–12496. doi:10.1073/pnas.1206643109.
59. Doutch J, Hough MA, Hasnain SS, Strange RW (2012) Challenges of sulfur SAD phasing as a routine method in macromolecular crystallography. J Synchrotron Rad 19: 19–29. doi:10.1107/S0909049511049004.
60. Liu Q, Dahmane T, Zhang Z, Assur Z, Brasch J, et al. (2012) Structures from anomalous diffraction of native biological macromolecules. Science 336: 1033–1037. doi:10.1126/science.1218753.
61. Bandaranayake AD, Almo SC (2014) Recent advances in mammalian protein production. FEBS Lett 588: 253–260. doi:10.1016/j.febslet.2013.11.035.
62. Peng WC, de Lau W, Forneris F, Granneman JCM, Huch M, et al. (2013) Structure of Stem Cell Growth Factor R-spondin 1 in Complex with the Ectodomain of Its Receptor LGR5. Cell Rep 3: 1885–1892. doi:10.1016/j.celrep.2013.06.009.
63. Staal FJ, Burgering BM, van de Wetering M, Clevers HC (1999) Tcf-1-mediated transcription in T lymphocytes: differential role for glycogen synthase kinase-3 in fibroblasts and T cells. Int Immunol 11: 317–323.

4

EXPRESSION, CRYSTALLIZATION AND STRUCTURE DETERMINATION OF LGR5-RSPO1

CHAPTER 5

Summary and Discussion

Weng Chuan Peng

Part of this chapter have been published in Genes and Development (2014) 28: 305-316

Crystal and Structural Chemistry, Bijvoet Center for Biomolecular Research, Department of Chemistry, Faculty of Science, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands

INTRODUCTION

LGR5 is a 7-transmembrane receptor from the 'Leucine-rich repeat-containing G-protein-coupled Receptor' (LGR) family [1]. LGR5 was first identified as one of a set of 80 Wnt target genes, essential for the high self-renewal capacity of the intestinal epithelial layer [2]. Subsequently, it was identified as a unique marker of Crypt Base Columnar (CBC) cells. These cells, residing at the bottom of the intestinal crypt compartment, function as adult stem cells enabling the continuous production of new cells populating the villus compartment. LGR5 marks not only adult stem cells in the intestine and colon, but also in the stomach, skin, hair-follicle, pancreas, kidney and mammary gland [3-9]. While LGR5 expression is restricted to CBC cells in the crypt, LGR4 showed broader expression throughout the crypts [10]. The intestinal epithelial layer, due to its well-characterized architecture and high regenerative capacity, serves as an excellent model system to study adult stem cell biology. Since Wnt signaling is the major driving force of this tissue renewal, it also constitutes an excellent resource for studying the fundamental aspects of this signaling pathway.

LGR5 and homologues LGR4 and LGR6 associate with R-spondin to potentiate Wnt signaling [11-14]. R-spondins are a family of small, secreted proteins consisting of four members (RSPO1–4). R-spondin is a potent Wnt agonist that induces β-catenin stabilization in HEK cells in the presence of a low-level of Wnt, and stimulates intestinal cell proliferation [15,16]. In line with this, R-spondin is an essential growth factor in minigut organoid cultures [17,18]. However, the mechanism LGR5-RSPO1 signaling remains unclear as it appeared independent from G-proteins coupled signaling.

More recently, stem-cells specific transmembrane E3 ligases ZNRF3 and homologue RNF43 were identified as negative regulators of Wnt signaling [19,20]. Both E3 ligases are Wnt target genes that ubiquitinate LRP6 and FZD receptors, serving as a negative feedback to restrict stem cells population. Intriguingly, R-spondin 1 was found to induce the formation of a ternary complex between LGR5 and ZNRF3/RNF43. The formation of such a ternary complex induced the membrane clearance of E3 ligases. Removal of E3 ligases resulted in an increase of Wnt receptor availability on cell surface, hence potentiating Wnt signaling [20]. The discovery of E3 ligases ZNRF3 (and RNF43) thus elucidates the mechanistic role of LGR5-RSPO1 in Wnt signaling.

The discoveries of LGR5, R-spondin and E3 ligases ZNRF3/RNF43 provide valuable insights into the role of Wnt signaling in adult stem cell proliferation and regulation (described in **chapter 1**). Crystallographic studies, by our group and others, provide molecular basis for RSPO1-LGR4/5 and RSPO1-ZNRF3/RNF43 interactions [21-26]. Various biochemical and functional studies offer insights into the dynamic of receptor-ligands interaction on the membrane. In addition, identification of novel interacting partners have shed light on other possible signaling mechanisms of LGR5 in adult stem-cells proliferation [27-29].

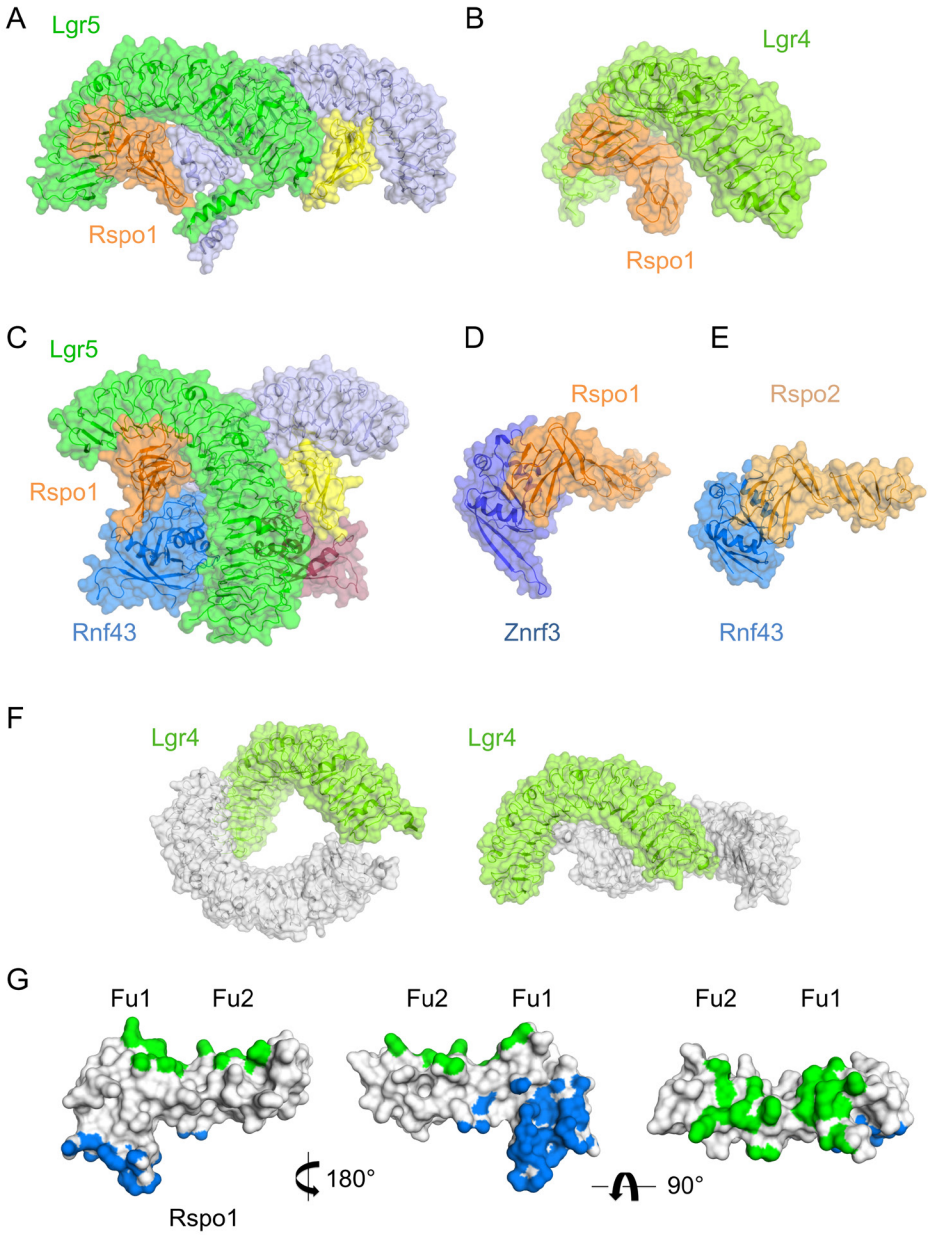


Fig 1. Crystal structures of LGR4/5, RSPO1/2 and ZNRF3/RNF43 complexes.
A, LGR5-RSPO1 (PDB 4BSR). B, LGR4-RSPO1 (PDB 4KT1). C, LGR5-RSPO1-RNF43 (PDB 4KNG). D, ZNRF3-RSPO1 (PDB 4CDK). E, RNF43-RSPO2 (PDB 4C9V). F, LGR4 (PDB 4LI1). G, Contact footprints of LGR5 on RSPO1 (green) and RNR43 on RSPO1 (blue). RNF43-RSPO1 interface overlapped with trans LGR5-RSPO1 dimer interface. Figure is reproduced from de Lau et al., Genes and Development 28:305-316.

Structure of Wnt agonist R-spondin

R-spondin has a signal peptide, two furin (Fu) domains (~100 residues), a thrombospondin type I (TSP-I) domain (~60 residues) and a polybasic tail (~60 residues). The sequence homology of the four R-spondins is about 60% [30]. Crystal structures of RSPO1 and RSPO2, comprising of the two Fu domains, were determined at a 2 Å resolution [21,25]. Both Fu1 and Fu2 domains contain three similarly constructed cysteine-knotted β-hairpins each with disulfide bond patterns that appear common to furin domains seen in insulin growth factor receptor (IGFR) and epidermal growth factor receptor (EGFR) (Garrett et al. 1998; Ogiso et al. 2002) (described in **chapter 2**). Overall, the two furin repeats that form a signaling-competent fragment (Kazanskaya et al. 2004) yield an extended structure, although a flexible hinge in between furin-1 and furin-2 results in various relative domain orientations [26].

Structural basis of R-spondin recognition by its receptor LGR4/5 offers insight into promiscuous receptor-ligand interaction

The ectodomain of LGR4/5 forms an anticipated horseshoe-like structure with 17 LRRs, flanked by LGR1-type N-terminal and C-terminal cysteine-rich caps [21,23,24,31] (Fig 1A-C). However, the horseshoe shape is twisted and kinked between LRR10 and LRR11. The binding interface of RSPO1 to the concave surface of LGR4 and LGR5 comprises LRR domains 3–9. The first β-hairpin of the Furin-2 repeat of RSPO1 contains a ^{106}FSHNF^{110} loop, of which aromatic residues Phe106 and Phe110 form a hydrophobic clamp (that we termed Phe-clamp) targeting Ala190 and participate in hydrophobic interaction with surrounding residues of LGR5. Single F106E or F110E mutations in the Phe-clamp completely abrogate receptor-ligand binding as well as the functional activity of R-spondin, as tested in Wnt reporter assays and small intestinal stem cell cultures [21]. Similarly, F106A and F110A mutations disrupt interaction with LGR4 [23]. A second LGR5 interaction site in R-spondin has a hydrophilic character and is mediated by Lys59 and Arg87 of Fu1 domain. Mutation of these residues reduced signaling, but the effects are less severe than above-mentioned mutations in Fu2. The R-spondin1/LGR interface is highly conserved among all four R-spondins and LGR4–6 in various species. Taken together, these studies provide a structural basis for the promiscuous interaction of LGR4-6 with RSPO1-4 [32] (described in **chapter 2**).

Structural basis of E3 ligase ZNRF3/RNF43 inhibition by R-spondin interaction offers insights into signaling mechanism

We and others presented the structure of RNF43/ZNRF3 alone, RSPO1 in complex with its co-receptor ZNRF3/RNF43, or ternary complex of LGR5-RSPO1-ZNRF3 [22,25,26] (Fig 1C-E). RNF43 and ZNRF3 ectodomains adopt a typical PA domain fold, similar to that of Grail (RNF 128), despite showing only 15% sequence identity. RSPO1 engages RNF43/ZNRF3 through its Fu1 domain, and only minimal contact is observed with Fu2. RNF43/ZNRF3 forms a shallow binding pocket for interaction with RSPO1. The key elements of the electrostatic interactions in RNF43/ZNRF3 are Gln84, His86, Lys108, and Glu110 (positions 100, 102, 125, and 127 in

human ZNRF3, respectively). Conserved residues in this interface on RSPO 1 are Ser48, Asn51, Arg66, and Gln71. Hence, RSPO1 interacts with LGR5 predominantly through Fu2 domain and with E3 ligase through Fu1 domain. In the ternary complex, LGR5 and RNF43 appear to be bridged by R-spondin, without any physical contact between the receptors [22], implying that the two receptors do not interact in the absence of R-spondin.

The structures of RSPO1-ZNRF3 provide a framework for understanding how disease mutations affect Wnt signaling and development. In anonychia patients, mutations in RSPO4 are found at the interface with ZNRF3 [22,25,26]. In Wnt reporter assays, these mutants showed abrogated signaling despite intact LGR5 binding [21]. In another study [33] and based on our unpublished observations, R66A and Q71A mutations in the Fu1 domain completely abolished binding to ZNRF3, and hence would explain the inactivity in signaling. Taken together, these structures provide the molecular basis of E3 ligase RNF43/ZNRF3 removal by R-spondin (described in **chapter 3**).

Quaternary arrangements of LGR4/5-RSPO1, RSPO1/2-ZNRF3 binary complexes and LGR5-RSPO1-ZNRF3 ternary complex

A puzzling aspect of the crystallographic analysis of several complexes is a variation in observed quaternary arrangements. In solution, LGR4/5-R-spondin complexes are monomeric [21,23,24], whereas LGR4 ECD by itself has been reported to behave as a dimer [24]. Xu et al. describe two 'N-C' LGR4 ectodomain arrangements in the asymmetric unit (ASU) of the crystal (Fig 1F, left panel) [24]. Inspection of this crystal lattice again shows a C-terminally "embraced" arrangement (Fig 1F, right panel). LGR4-RSPO1 complex, however, is a monomeric complex [24].

In multiple crystal forms of LGR5/R-spondin-1 complexes, we consistently observed dimeric, "embracing" complexes [21] (Fig 1A). R-spondin Fu1 domain in the complex appeared to bind in trans to the LGR5 C-terminal cysteine caps to mediate the formation of 2:2 complex. This arrangement has not been observed in LGR4 crystal structures reported by others. The RNF43/ZNRF3 interaction site on R-spondin overlaps with the dimer interface (Fig 1G), and hence binding of ZNRF3 to RSPO1 would disrupt the observed LGR5-RSPO1 dimers. How the 2:2 dimeric LGR5-RSPO1 complex arrangement affects binding of ZNRF3 is currently unclear.

In addition, we and Zebisch et al. observed dimers of ZNRF3 alone and 2:2 ZRNF3-RSPO1 complexes [25,26] in many crystal forms, but not in RNF43. A model for LGR5-RSPO1-ZNRF3 dimer based on ZNRF3-RSPO1 dimer was proposed [25]. It is currently unclear what the physiological relevance is of higher-order organization of the receptors studied. It has been reported that in many E3 ligase, RING domain dimerization is required for ubiquitination activity [34]. ZNRF3 RING domain shared 30-33% sequence identity with the RING domains of RNF4 and RNF8. Homodimers of RING domain of RNF4 and RNF8 have been reported by crystallographic studies [35,36]. It may be possible that dimerization of E3 ligase ZNRF3 PA-domain has a functional implication and is not merely a coincidence of crystals packing.

Frizzled receptor, a substrate for E3 ligase

Frizzled receptor is a substrate for RNF43/ZNRF3 ubiquitination and FZD:ZNRF3/RNF43 were co-immunoprecipitated [19,20]. Antibody binding to ZNRF3 ectodomain prevents E3 ligase activity on FZD, indicating that the ectodomain interaction is required for substrate recognition [20]. In the absence of crystal structures, the molecular detail of FZD-RNF43/ZNRF3 interaction is unknown. The FZD:ZNRF3 interaction could also be indirect, requiring one or more additional factors. Further, the mechanism and regulation of FZD ubiquitination by E3 ligases is not well characterized. Study on C. elegans E3 ligase PLR-1 shed some light on this aspect [37]. FZD-CRD domain is required for interaction with PLR-1, while the Lys residue in the second intracellular of FZD is required for ubiquitination by PLR-1 [37]. In addition, PLR-1 is observed predominantly in the endosome with FZD, indicating that co-internalization and degradation process is constitutively active. Notably, PLR-1 is not detected on the membrane when FZD is present, but accumulates on cell surface when FZD expression is suppressed.

Frizzled receptor family, which consists of 10 members (FZD 1-10), is involved in both canonical and non-canonical Wnt signaling [38]. The substrate specificity for stem-cell specific E3 ligases ZNRF3/RNF43 is currently not known. Hao et al. detected ZNRF3 interaction with FZD4, FZD6 and FZD8 through co-immunoprecipitation [20]. Koo et al. detected FZD5 ubiquitination by RNF43 and observed FZD 1-3 downregulation after RNF43 overexpression [19]. Based on these observations it is possible that both E3 ligases recognize multiple FZD members. In addition, LRP5 abundance on the membrane decreases when RNF43 is present and LRP6 lacking its ectodomain is not inhibited by RNF43 [19], suggesting that LRP6 is targeted by E3 ligase for degradation along with FZD. At present, it is not clear if E3 ligase targets only free LRP6/FZD receptors, reducing the number of receptors available for signaling, or also targets activated Wnt-LRP5/6-FZD ternary complex, directly 'switching off' signaling.

The current mechanism of LGR5-RSPO1 mediated E3 ligase removal suggests a molecular competition between LGR5-RSPO1 complex and FZD receptor for E3 ligase. Since R-spondin prevents FZD ubiquitination, the affinity of LGR5-RSPO1:ZNRF3 interaction must be higher than or comparable to the FZD:ZNRF3 interaction. ZNRF3 would be expected to associate weakly with FZD, in a reversible manner that can be blocked upon R-spondin binding to ZNRF3. This is consistent with structural analysis that ZNRF3 has only one conserved binding platform, and possibly ZNRF3 binds FZD or R-spondin using the same or an overlapping interface.

RSPO1-LGR5 versus RSPO1-ZNRF3 binding affinity and implications for signaling

LGR5 and homologues bind R-spondin with high affinity (K_D of 2-5 nM) as determined by several studies [10,13,39]. On the contrary, RSPO1 interacts weakly with RNF43 (K_D of 7-10 μM) [22] and ZNRF3 (K_D of 0.8-2.0 μM) [26,33]. The ~1000-fold difference in RSPO1-LGR5 and RSPO1-ZNRF3/RNF43 binding affinities can partly be explained by different characteristics of the Fu1 and Fu2 binding interfaces, i.e. hydrophilic in RSPO1-ZNRF3/RNF43 and hydrophobic in RSPO1-LGR5. On the other hand, presence of the LGR5 ectodomain did not change the

binding affinity of RSPO1 to E3 ligase ZNRF3/RNF43 significantly, which can be attributed to the lack of physical contact between LGR5 and E3 ligases. This observation prompted Chen et al. to suggest that LGR5 served solely as an high-affinity engagement receptor for R-spondins [22]. It is likely that the nanomolar binding affinity allows R-spondin to associate efficiently with LGR4-6 (almost irreversibly) at low concentration, resulting in a greater local concentration on the membrane and enhanced R-spondin interaction with E3 ligase. The ternary complex is subsequently removed from the cell surface by endocytosis [20].

The weak binding of R-spondin to E3 ligase ZNRF3/RNF43 has some broad implications on stem-cell proliferation. E3 ligases function as negative regulators of adult stem cell proliferation [18,19]. Down-regulation of ZNRF3/RNF43 or inactivating mutations are often found in adenoma and cancer [40-42]. E3 ligases are targeted by LGR5 for removal, upon R-spondin binding. The lack of physical interaction between the two receptors, LGR5 and E3 ligases, ensures that E3 ligase is not removed by LGR5 in the absence of R-spondin. Hence, regulation of E3 ligase activity by R-spondin activity is critical for stem-cells proliferation. A strong association of R-spondin to E3 ligase is undesirable, as R-spondin would bind to (and inhibit) E3 ligase activity when (and where) LGR5 is not present. On the contrary, weak association of R-spondin to ZNRF3 allows its activity to be modulated by LGR5, e.g. up-regulation of LGR5 expression to kickstart proliferation, or down-regulation of LGR5 to slow down proliferation. Interestingly, LGR5 and ZNRF3 are both Wnt target genes with opposing functions. LGR5 functions to stimulate stem-cell proliferation (yielding a positive feedback loop), while ZNRF3 serves as a 'brake' to restrict stem cell population size (negative feedback loop). How the expression of these two receptors are regulated to achieve optimal stem-cell proliferation remains a matter of interest.

In addition, R-spondin-induced membrane clearance of ZNRF3 is dependent on E3 ligase RING domain and presence of LGR4/5, suggesting that E3 ligase activity is required for self-internalization. Clathrin-mediated internalization has been reported for LGR5 [13,39], which may serve as a mechanism by which ZNRF3 is removed from the cell surface. It is currently unknown if LGR4/5 itself is ubiquitinated by E3 ligase.

LGR4 engages IQGAP to phosphorylate LRP6 for activating Wnt signaling

In a very recent development, IQ-motif containing GTPase-activating protein 1 (IQGAP1) was found to bridge LGR4-RSPO1 to the Wnt receptors complex. LGR4-RSPO1 engage IQGAP to enhance LRP6 phosphorylation, mediated by MEK1/2 [29]. Physical interaction between LGR4 and Wnt signaling complex was detected, which may explain the earlier observation made by mass-spectrometry analysis [10]. This study provides an additional signaling role for LGR4-RSPO1, in addition to the removal of E3 ligase.

G-protein coupled signaling by LGR5

LGR4-6 belong to the family of GPCR proteins with large ectodomains, which includes

glycoprotein hormone receptors such as follicle-stimulating hormone receptor (FSHR), luteinizing hormone receptor (LHR) and thyroid-stimulating hormone receptor (TSHR) [43,44] (also known as LGR1-3, respectively). G-protein coupled signaling through $G_{\alpha s}$ is well established for the glycohormone protein receptors. However, several studies reported that LGR5-R-spondin signaling is not coupled to G-proteins (Gi, Gq and Gs) and is not associated with cAMP elevation or β-arrestin internalization.

In a recent study, Snyder et al. reported that LGR5, but not LGR4 or LGR6, possesses "NPXXY" motif, which is important for G-protein signaling and engaging β-arrestin [45]. They observed β-arrestin2 recruitment by LGR5, mediated by G-protein coupled receptor kinase (GRK 4-6). In addition, an "SSS" motif present on the C-terminal tail of LGR5 is required for β-arrestin2 recruitment. In another study, Kwon et al. observed that LGR5 activates $G_{12/13}$-Rho GTPase signaling, but not LGR4/6 [46]. Surprisingly, both studies observed that LGR5 exhibits a classical GPCR behavior, which is not present in LGR4 or LGR6 and is independent of R-spondin. The physiological relevance of GPCR signaling would require further verification and may depend on yet unidentified ligands.

Beyond E3 ligase: LGR5, a multi-task regulator

Norrin, a small secreted cysteine-rich ligand interacts with FZD4 receptor and LRP5 co-receptor in the presence of transmembrane protein TSPAN12 to activate canonical Wnt signaling [47-49]. Norrin is expressed in the retina, ear, fetal and adult brain, and is functionally similar to Wnt [47], i.e. it induces β-catenin stabilization.

In drosophila, dLGR2, the invertebrate orthologue of LGR4-6, has been found to associate with heterodimeric ligand Bursicon [50]. In a recent study, Hsueh and colleagues noted that Norrin is an orthologue for invertebrate Bursicon [27] and associates strongly with LGR4-6. However, Norrin only activates Wnt signaling through LGR4. The mechanism of LGR4-Norrin remains unclear, and is likely to be different from LGR5-RSPO1. Norrin may function to bridge LGR4 to LRP5/FZD4 complex, leading to the formation of a signalosome complex. Notably, this study provides concrete evidence that LGR4-6 can interact with ligands other than R-spondin to potentiate Wnt signaling.

In another study, TROY, a tumor necrosis tumor family member, was identified as a negative regulator of Wnt signaling in the intestine [28]. The authors also noted that TROY is inhibited by LGR5. Proliferation of TROY-knockout organoid is less dependent on exogenous R-spondin than wild-type organoid. Although the mechanism is not known, it could be similar to E3 ligase removal by LGR5-R-spondin.

CONCLUDING REMARKS

Adult stem cell marker LGR5 and transmembrane E3 ligase ZNRF3/RNF43 are Wnt target genes exclusively expressed in Wnt-dependent adult stem cells. In the absence of R-spondin, E3 ligase

targets FZD receptor for degradation, negatively regulating Wnt signaling. LGR5 associates with R-spondin, a potent Wnt agonist to induce membrane clearance of E3 ligases. This leads to membrane accumulation of Wnt receptor and consequently boosts Wnt signal strength to kickstart stem cell proliferation. Crystallographic, biochemical and functional studies depict how LGR5 receptors serve to efficiently recruit R-spondin ligands to the membrane and bring them into position for interaction with RNF43/ZNRF3. The canonical Wnt signaling cascade is encoded by all animal genomes and a RNF43/ZNRF3 homologue capable of down-regulating Frizzled surface expression has been described in at least one invertebrate; i.e. PLR-1 in the roundworm C. elegans. However, the LGR5-RSPO1 module appears to be a relatively recent evolutionary "add-on" seen only in vertebrates. In addition, LGR4-6 are versatile receptors, interacting with many other ligands, such as Norrin and TROY, suggesting multiple roles and signaling mechanisms of LGR5 in activating Wnt signaling. Interestingly, more studies have shown that LGR5 exhibits classical GPCR behavior, despite early observations that LGR5-RSPO1 signaling is not mediated by downstream G-proteins. It would be fascinating to understand how these different signaling mechanisms come together to regulate Wnt signaling.

REFERENCES

1. Barker N, Clevers H (2010) Leucine-rich repeat-containing G-protein-coupled receptors as markers of adult stem cells. Gastroenterology 138: 1681–1696. doi:10.1053/j.gastro.2010.03.002.
2. van der Flier LG, Sabates-Bellver J, Oving I, Haegebarth A, De Palo M, et al. (2007) The Intestinal Wnt/TCF Signature. Gastroenterology 132: 628–632. doi:10.1053/j.gastro.2006.08.039.
3. Barker N, van Es JH, Kuipers J, Kujala P, van den Born M, et al. (2007) Identification of stem cells in small intestine and colon by marker gene Lgr5. Nature 449: 1003–1007. doi:10.1038/nature06196.
4. Barker N, Huch M, Kujala P, van de Wetering M, Snippert HJ, et al. (2010) Lgr5(+ve) stem cells drive self-renewal in the stomach and build long-lived gastric units in vitro. Cell Stem Cell 6: 25–36. doi:10.1016/j.stem.2009.11.013.
5. Huch M, Bonfanti P, Boj SF, Sato T, Loomans CJM, et al. (2013) Unlimited in vitro expansion of adult bi-potent pancreas progenitors through the Lgr5/R-spondin axis. EMBO J 32: 2708–2721. doi:10.1038/emboj.2013.204.
6. Snippert HJ, Haegebarth A, Kasper M, Jaks V, van Es JH, et al. (2010) Lgr6 marks stem cells in the hair follicle that generate all cell lineages of the skin. Science 327: 1385–1389. doi:10.1126/science.1184733.
7. Huch M, Dorrell C, Boj SF, van Es JH, Li VSW, et al. (2013) In vitro expansion of single Lgr5+ liver stem cells induced by Wnt-driven regeneration. Nature 494: 247–250. doi:10.1038/nature11826.
8. Plaks V, Brenot A, Lawson DA, Linnemann JR, Van Kappel EC, et al. (2013) Lgr5-expressing cells are sufficient and necessary for postnatal mammary gland organogenesis. Cell Rep 3: 70–78. doi:10.1016/j.celrep.2012.12.017.
9. Barker N, Rookmaaker MB, Kujala P, Ng A, Leushacke M, et al. (2012) Lgr5(+ve) Stem/

Progenitor Cells Contribute to Nephron Formation during Kidney Development. Cell Rep 2: 540–552. doi:10.1016/j.celrep.2012.08.018.
10. de Lau W, Barker N, Low TY, Koo B-K, Li VSW, et al. (2011) Lgr5 homologues associate with Wnt receptors and mediate R-spondin signalling. Nature 476: 293–297. doi:10.1038/nature10337.
11. de Lau W, Peng WC, Gros P, Clevers H (2014) The R-spondin/Lgr5/Rnf43 module: regulator of Wnt signal strength. Genes Dev 28: 305–316. doi:10.1101/gad.235473.113.
12. Carmon KS, Gong X, Lin Q, Thomas A, Liu Q (2011) R-spondins function as ligands of the orphan receptors LGR4 and LGR5 to regulate Wnt/beta-catenin signaling. Proc Natl Acad Sci USA 108: 11452–11457. doi:10.1073/pnas.1106083108.
13. Glinka A, Dolde C, Kirsch N, Huang Y-L, Kazanskaya O, et al. (2011) LGR4 and LGR5 are R-spondin receptors mediating Wnt/β-catenin and Wnt/PCP signalling. EMBO Rep 12: 1055–1061. doi:10.1038/embor.2011.175.
14. Ruffner H, Sprunger J, Charlat O, Leighton-Davies J, Grosshans B, et al. (2012) R-Spondin potentiates Wnt/β-catenin signaling through orphan receptors LGR4 and LGR5. PLoS ONE 7: e40976. doi:10.1371/journal.pone.0040976.
15. Kim K-A, Kakitani M, Zhao J, Oshima T, Tang T, et al. (2005) Mitogenic influence of human R-spondin1 on the intestinal epithelium. Science 309: 1256–1259. doi:10.1126/science.1112521.
16. Kazanskaya O, Glinka A, del Barco Barrantes I, Stannek P, Niehrs C, et al. (2004) R-Spondin2 is a secreted activator of Wnt/beta-catenin signaling and is required for Xenopus myogenesis. Dev Cell 7: 525–534. doi:10.1016/j.devcel.2004.07.019.
17. Sato T, Vries RG, Snippert HJ, van de Wetering M, Barker N, et al. (2009) Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche. Nature 459: 262–265. doi:10.1038/nature07935.
18. Sato T, Clevers H (2013) Growing self-organizing mini-guts from a single intestinal stem cell: mechanism and applications. Science 340: 1190–1194. doi:10.1126/science.1234852.
19. Koo B-K, Spit M, Jordens I, Low TY, Stange DE, et al. (2012) Tumour suppressor RNF43 is a stem-cell E3 ligase that induces endocytosis of Wnt receptors. Nature 488: 665–669. doi:10.1038/nature11308.
20. Hao H-X, Xie Y, Zhang Y, Charlat O, Oster E, et al. (2012) ZNRF3 promotes Wnt receptor turnover in an R-spondin-sensitive manner. Nature 485: 195–200. doi:10.1038/nature11019.
21. Peng WC, de Lau W, Forneris F, Granneman JCM, Huch M, et al. (2013) Structure of Stem Cell Growth Factor R-spondin 1 in Complex with the Ectodomain of Its Receptor LGR5. Cell Rep 3: 1885–1892. doi:10.1016/j.celrep.2013.06.009.
22. Chen P-H, Chen X, Lin Z, Fang D, He X (2013) The structural basis of R-spondin recognition by LGR5 and RNF43. Genes Dev 27: 1345–1350. doi:10.1101/gad.219915.113.
23. Wang D, Huang B, Zhang S, Yu X, Wu W, et al. (2013) Structural basis for R-spondin recognition by LGR4/5/6 receptors. Genes Dev 27: 1339–1344. doi:10.1101/gad.219360.113.
24. Xu K, Xu Y, Rajashankar KR, Robev D, Nikolov DB (2013) Crystal Structures of Lgr4 and Its Complex with R-Spondin1. Structure. doi:10.1016/j.str.2013.07.001.

25. Zebisch M, Xu Y, Krastev C, Macdonald BT, Chen M, et al. (2013) Structural and molecular basis of ZNRF3/RNF43 transmembrane ubiquitin ligase inhibition by the Wnt agonist R-spondin. Nat Commun 4: 2787. doi:10.1038/ncomms3787.
26. peng W, de Lau W, Madoori PK, Forneris F, Granneman JCM, et al. (2013) Structures of Wnt-Antagonist ZNRF3 and Its Complex with R-Spondin 1 and Implications for Signaling. PLoS ONE 8: e83110. doi:10.1371/journal.pone.0083110.
27. Deng C, Reddy P, Cheng Y, Luo C-W, Hsiao C-L, et al. (2013) Multi-functional norrin is a ligand for the LGR4 receptor. J Cell Sci. doi:10.1242/jcs.123471.
28. Fafilek B, Krausova M, Vojtechova M, Pospichalova V, Tumova L, et al. (2012) Troy, a Tumor Necrosis Factor Receptor Family Member, Interacts With Lgr5 to Inhibit Wnt Signaling in Intestinal Stem Cells. Gastroenterology. doi:10.1053/j.gastro.2012.10.048.
29. Carmon KS, Gong X, Yi J, Thomas A, Liu Q (2014) RSPO-LGR4 functions via IQGAP1 to potentiate Wnt signaling. Proceedings of the National Academy of Sciences. doi:10.1073/pnas.1323106111.
30. de Lau WBM, Snel B, Clevers HC (2012) The R-spondin protein family. Genome Biol 13: 242. doi:10.1186/gb-2012-13-3-242.
31. Cheng Z, Biechele T, Wei Z, Morrone S, Moon RT, et al. (2011) Crystal structures of the extracellular domain of LRP6 and its complex with DKK1. Nat Struct Mol Biol 18: 1204–1210. doi:10.1038/nsmb.2139.
32. Moad HE, Pioszak AA (2013) Reconstitution of R-spondin:LGR4:ZNRF3 adult stem cell growth factor signaling complexes with recombinant proteins produced in Escherichia coli. Biochemistry 52: 7295–7304. doi:10.1021/bi401090h.
33. Xie Y, Zamponi R, Charlat O, Ramones M, Swalley S, et al. (2013) Interaction with both ZNRF3 and LGR4 is required for the signalling activity of R-spondin. EMBO Rep. doi:10.1038/embor.2013.167.
34. Deshaies RJ, Joazeiro CAP (2009) RING Domain E3 Ubiquitin Ligases. Annu Rev Biochem 78: 399–434. doi:10.1146/annurev.biochem.78.101807.093809.
35. Liew CW, Sun H, Hunter T, Day CL (2010) RING domain dimerization is essential for RNF4 function. Biochem J 431: 23–29. doi:10.1042/BJ20100957.
36. Mattiroli F, Vissers JHA, van Dijk WJ, Ikpa P, Citterio E, et al. (2012) RNF168 ubiquitinates K13-15 on H2A/H2AX to drive DNA damage signaling. Cell 150: 1182–1195. doi:10.1016/j.cell.2012.08.005.
37. Moffat LL, Robinson RE, Bakoulis A, Clark SG (2014) The conserved transmembrane RING finger protein PLR-1 downregulates Wnt signaling by reducing Frizzled, Ror and Ryk cell-surface levels in C. elegans. Development 141: 617–628. doi:10.1242/dev.101600.
38. Huang H-C, Klein PS (2004) The Frizzled family: receptors for multiple signal transduction pathways. Genome Biol 5: 234. doi:10.1186/gb-2004-5-7-234.
39. Carmon KS, Lin Q, Gong X, Thomas A, Liu Q (2012) LGR5 interacts and cointernalizes with Wnt receptors to modulate Wnt/β-catenin signaling. Mol Cell Biol 32: 2054–2064. doi:10.1128/MCB.00272-12.
40. Zhou Y, Lan J, Wang W, Shi Q, Lan Y, et al. (2013) ZNRF3 acts as a tumour suppressor by the Wnt signalling pathway in human gastric adenocarcinoma. J Mol Histol. doi:10.1007/s10735-013-9504-9.

41. Jiang X, Hao H-X, Growney JD, Woolfenden S, Bottiglio C, et al. (2013) Inactivating mutations of RNF43 confer Wnt dependency in pancreatic ductal adenocarcinoma. Proceedings of the National Academy of Sciences. doi:10.1073/pnas.1307218110.
42. Ryland GL, Hunter SM, Doyle MA, Rowley SM, Christie M, et al. (2013) RNF43 is a tumour suppressor gene mutated in mucinous tumours of the ovary. J Pathol 229: 469–476. doi:10.1002/path.4134.
43. Vassart G, Pardo L, Costagliola S (2004) A molecular dissection of the glycoprotein hormone receptors. Trends Biochem Sci 29: 119–126. doi:10.1016/j.tibs.2004.01.006.
44. Pierce JG, Parsons TF (1981) Glycoprotein hormones: structure and function. Annu Rev Biochem 50: 465–495. doi:10.1146/annurev.bi.50.070181.002341.
45. Snyder JC, Rochelle LK, Barak LS, Caron MG (2013) The stem cell-expressed receptor Lgr5 possesses canonical and functionally active molecular determinants critical to β-arrestin-2 recruitment. PLoS ONE 8: e84476. doi:10.1371/journal.pone.0084476.
46. Kwon MS, Park B-O, Kim HM, Kim S (2013) Leucine-rich repeat-containing G-protein coupled receptor 5/GPR49 activates $G_{12/13}$-Rho GTPase pathway. Mol Cells. doi:10.1007/s10059-013-0173-z.
47. Xu Q, Wang Y, Dabdoub A, Smallwood PM, Williams J, et al. (2004) Vascular development in the retina and inner ear: control by Norrin and Frizzled-4, a high-affinity ligand-receptor pair. Cell 116: 883–895.
48. Junge HJ, Yang S, Burton JB, Paes K, Shu X, et al. (2009) TSPAN12 regulates retinal vascular development by promoting Norrin- but not Wnt-induced FZD4/beta-catenin signaling. Cell 139: 299–311. doi:10.1016/j.cell.2009.07.048.
49. Ke J, Harikumar KG, Erice C, Chen C, Gu X, et al. (2013) Structure and function of Norrin in assembly and activation of a Frizzled 4-Lrp5/6 complex. Genes Dev 27: 2305–2319. doi:10.1101/gad.228544.113.
50. Luo CW, Dewey EM, Sudo S, Ewer J, Hsu SY, et al. (2005) Bursicon, the insect cuticle-hardening hormone, is a heterodimeric cystine knot protein that activates G protein-coupled receptor LGR2. Proceedings of the National Academy of Sciences 102: 2820–2825. doi:10.1073/pnas.0409916102.

SAMENVATTING

De wand van de dunne darm is geplooid en bestaat uit vingervormige uitstulpingen, de villi en daartussen gelegen instulpingen, de crypten. De dunne darm is verantwoordelijk voor de opname van voedingsstoffen. De hierbij benodigde enzymen brengen ook schade toe aan de darmwandcellen. Dit veroorzaakt een sterke beperking in de levensduur van deze cellen. Deze celdood maakt het noodzakelijk dat de epitheelcellen aan de oppervlakte van het darmweefsel voortdurend vernieuwd worden. De crypt is een “celproductiefabriek”, waarin door deling vanuit stamcellen uiteindelijk gedifferentieerde cellen worden aangemaakt. De meesten hiervan migreren naar het villus compartiment waar ze hun gespecialiseerde functie uitvoeren. De drijvende kracht voor dit massieve proces van zelfvernieuwing wordt gevormd door een evolutionair oud moleculair mechanisme, Wnt signaling genaamd. Het is belangrijk dat de stamcelproliferatie nauwkeurig geregeld is zodat er een balans bestaat tussen het aantal dode cellen en het aantal nieuw geproduceerde cellen in gezond darmweefsel. Dit wordt ook wel homeostase genoemd. Als er teveel nieuwe cellen worden geproduceerd vergroot dit de kans op de vorming van een tumor. Wnt signaling wordt “aangezet” wanneer Wnt eiwitten, aanwezig buiten de cel, binden aan Frizzled receptoren en hun hulpreceptoren LRP5/6, aanwezig op het oppervlak van de stamcel. Wnt signaling wordt sterk gereguleerd door LGR5, R-spondin en E3-ligase ZNRF3. Dit proefschrift beschrijft de moleculaire interactie van LGR5 met R-spondin en van R-spondin met ZNRF3 door kristallografische studies te combineren met functionele analyses. Samenvattend geeft deze studie inzicht in de regulatie van Wnt signaling door LGR5, R-spondin en ZNRF3 op moleculair niveau.

CURRICULUM VITAE

Weng Chuan Peng was born in Teluk Intan, Malaysia on 10 November 1982. He attended SMJK Sam Tet high school in Ipoh, Malaysia and studied Life Sciences in Nanyang Technological University (NTU), Singapore. He completed his undergraduate study with a First Class Honours (BSc) under the Accelerated Bachelor Program. During his undergraduate study, he worked in the laboratory of Prof. Dr. Jaume Torres, Division of Structural Biology and Computational Biology (NTU), focusing on the structural study of SARS coronavirus envelope protein in lipid bilayers using infrared spectroscopy. Upon graduation, he continued his research on the same topic, as a project officer (2006 to 2008) in Dr Jaume Torres' lab. During this time, he also initiated the study on the effects of transmembrane domain FGFR3 mutations on receptor activation, found in patients with dwarfism (achondroplasia and Crouzon syndrome). In 2009, he started his PhD training at the Bijvoet Center for Biomolecular Research, University Utrecht, under the supervision of Prof. Dr. Piet Gros. His research focuses on the structural studies of Wnt signaling regulation by LGR5, R-spondin 1 and E3 ligase ZNRF3 in adult stem cell activation, in collaboration with Prof. Dr. Hans Clevers and Dr Wim de Lau, Hubrecht Institute, KNAW.

PUBLICATIONS

1. de Lau W, **Peng WC**, Gros P, Clevers H (2014) The R-spondin/Lgr5/Rnf43 module: regulator of Wnt signal strength. Genes & development 28(4): 305–16.

2. **Peng WC**, de Lau W, Forneris F, Granneman JCM, Huch M, Clevers H, Gros P (2013) Structure of Stem Cell Growth Factor R-spondin 1 in Complex with the Ectodomain of Its Receptor LGR5. Cell Reports 3(6): 1885–1892.

3. **Peng WC**, de Lau W, Madoori PK, Forneris F, Granneman JCM, Clevers H, Gros P (2013) Structures of Wnt-antagonist ZNRF3 and its complex with R-spondin 1 and implications for signaling. PLoS ONE 8(12): e83110.

4. **Peng WC**, Lin X, Torres J (2009) The strong dimerization of the transmembrane domain of the fibroblast growth factor receptor 3 (FGFR3) is modulated by C-terminal juxtamembrane residues. Protein Science 18: 450–459.

5. Priya R, Tadwal, VS, Roessle MW, Gayen S, Hunke C, **Peng WC**, Torres J, Grüber G (2008) Low resolution structure of subunit b (b (22-156)) of Escherichia coli F(1)F(O) ATP synthase in solution and the b-delta assembly. Journal of Bioenergetics and Biomembranes 40: 245–255.

ACKNOWLEDGEMENTS

My PhD journey in Utrecht has been a very pleasant and unforgettable experience, accompanied by many amazing people that I met along the path. I would like to take this occasion to thank all those who have made this wonderful journey possible.

My deepest gratitude goes to my PhD supervisor, Piet Gros: Piet, thanks for your supervision, guidance, ideas, trust, and for providing such a stimulating environment for research. You have always been a very amiable and supportive supervisor. I remember the first thing you told me during our work discussion is to say 'no' to your suggestions! Hence it has always been easy communicating with you because you are very humble and willing to accept other ideas. You have great visions and always set sight on very challenging projects. Your optimism keeps difficult projects alive. I must admit that it is fun looking at crystal structures with you, as you would instantly come up with various mechanisms (and followed by many crazy thoughts). I did learn something about crystallography from you (remember crystal packing, B-factor, etc). And thanks for giving me the freedom in working on the topics I like in my final years. Work aside, we have had fun moments talking about random things at the coffee table. Your absence from the Christmas dinner was felt, and I am glad that you surprisingly showed up during the last labuitje. I wish you great success in sailing the lab in new direction.

I would like to extend my gratitude to my collaborators, Hans Clevers and Wim de Lau from the Hubrecht Institute. Wim, it has been a great pleasure working with you for the past 4 years. I learned a great deal from you about Wnt signaling, LGR5 and R-spondin. We visited each other lab on regular basis and brief updates often turned into hours of discussion. Your enthusiasm is inspiring and contagious. Despite your busy schedule, you have never turned down my request for experiments, suggestions, data, thesis writing, advice and etc. The support you have given is greatly appreciated.

My PhD in K&S is not just about crystallography, it is also defined by many great experiences: greasy Fri(ed)day, healthy Wednesday, Sichuan dinner, sushi, cake baking competition, CCP4 weekend in England, squash, table tennis, ice skating, canoeing, 6-day race in Ghent, lab-outing, Sinterklaas celebration, Christmas dinner and etc. Many of these activities have now become part of the lab tradition. To the PhD students, post-doc and staff who were present during my early years in K&S, thanks for making my stay here comfortable. Eric, you are instrumental in maintaining the lab in good order, thanks for all the advice on molecular biology. You often have very constructive ideas during group discussion. Bert, thanks for sharing your tips and tricks in solving structures. Your 'woohoo' scream at the corridor is unmistakeable. I wish you success with your 'neuro' team. Michael, my C9 mentor, Bhutjolokia is not the hottest pepper anymore, Naga Viper is your next challenge; Jin, thanks for the chinese dinners, chinese new year party and our trip to Cologne and Paris; Arjen, your heroic action in holding the train door in Antwerp is unforgettable, it was great seeing you again Hamburg; Dennis, the crystal farm manager, your sambal boontje is lekker, thanks for not farting in the cold room and have fun in Australia; Els, thanks for the Christmas dinner at Abel Tasmaniastraat, and advice for thesis

printing; Tom, thanks for the marmite; Eddie, the greasy master and Borrel Boy; Pascal, the French master chef, thanks for the baking lesson.

Federico and Pramod, we joined at the same time and we have gone through almost the entire K&S chapter together. Federico, thanks for everything that you taught me in the wet lab and in the terminal room. You are not only my lab bench mate, but also my terminal roommate, conference roommate and synchrotron partner. We endured many nights without sleep in the synchrotrons at ESRF and SLS. Your tiramisu, Italian 'expression', crazy stories about Italian politics/mafia, your passion for heavy metal, sushi, spicy food, and your enthusiasm, organizational skill, the number of constructs you made, never fail to impress me. Wish you lots of success in Pavia! Pramod, it has been fun working with you on Wnt signaling. Coming from another part of the world, we shared many things in common, e.g. food, culture, core values. It has been easy talking to you, and thanks for always offering me good food during lunch. Congratulation with your second baby! I would also like to offer my very special thanks to Joke; your help with cloning is critical for the success of my PhD. It started as a small favour soon turned into a long collaboration! I lost count of how many boxes of constructs we made and I apologize that we are always in hurry. I know I can text you anytime of the day and you would always respond immediately. Thanks for being my paranymph and helping with thesis proofreading, for all the Dutch translation, holiday photos and tips for Spain (Andalucía is awesome). You have been such a great colleague and friend! Chris, thanks for helping with the first C9 grid in Leiden. Christmas market in Aachen, after many years of planning, is an unforgettable experience. Thanks for being my paranymh!

XG, I wish we had more time to explore other cities by bikes. Bubble tea in London was worth queuing for although we almost missed our train. Remco, ice skating season starts again, you can slow down a bit, your joke is the best; Louris, when is the next BBQ in your garden? Matti, stop spinning the ball, good luck with the dimers. Hedwich, you always greet me with 'selamat pagi, apa khabar', your bahasa is impressive. Nadia, yoga is contagious, keep baking. Deniz, the lab movie is awesome, you deserve an Oscar. Camilla, your tiramisu is 'super super' good, I like your 'super super' expression for everything. Revina, the new Wnt mate, your constructs list is exploding, stop cloning. Tim, the bacon mints is awful! Tom, your complement storyboard is definitely impressive, another beer? Martin, Blackforest trip was awesome, I will not forget the maultasche. Viviana, I am sure you made the right choice to come here, the winter is cold but the people are warm. I would also like to thank to other people in K&S: Ton, Loes, Toine (thanks for integrating all complicated datasets), Arie (your FB post is always interesting), Cecile and Caroline (thanks for handling all the administrative issue. Caroline, you clay sculpture is the best), Harma (good to see you back in the Kruyt), Marian, Xin Xuan, the UPE guys (for the hundreds of transfections and grams of protein) : Wieger, Roland, Lucio, Smiriti. Smiriti, what is your next destination, after Malaysia?

I would like to extend my gratitude to my friends outside K&S who made my stay in Utrecht so colourful: Piero, Eleonora, Sara, Ana, Serena, Paul, Maria, RK, Shusil, Ramesh. The Italian dinner (panzerotti), BBQ, annual biking trip in Keukenhof, Queensday in Amsterdam,

Christmas dinner and games (Tombola), birthday celebrations, and New York will remain good memories. Naushad, it was great meeting you at the ING trips. Soon we started our routine Saturday lunch at Laplace with the rest of the gang, Sacha, Sean, Grace and Francisco. This acknowledgement would be incomplete with special thanks to Guru for being such a good company. I will not forget our dinners at TaiSoen and Saturday shopping. I hope everything is 'good only' with you.

To all my housemates in Brucknerlaan 74: RK, Fikadu, Guru, XG, Alessandro, Ramesh, Shusil, good luck with your scientific adventures. Elisa, thanks for accompanying me through the final chapter of my PhD, which is also the most challenging part.

I would also like to drop a quick note of thank you to the members of my former lab: Jaume Torres, Krupakar & Soni, Siok Wan, Ja, Lin Xin, Bindhu, Uma, Raja; and to my friends back home: Sau Siong & JY, Sze Min, Chuan Leng, Jien Sze, Voon Chieh, Phooi San, Kon Sang, Wei Xian, Sunny, Andre, Shubhendro, Kong Rong, Chin Ee, Wai Mun and Yong Jiann.

Lastly, my sincere thanks to my parents and family. I wish I could spend more time with you all. *Ibu ibu, engkaulah ratu hatiku.*