

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/263586045>

Mapping Key Regions of the RXFP2 Low-Density Lipoprotein Class-A Module That Are Involved in Signal Activation

ARTICLE in **BIOCHEMISTRY** · JULY 2014

Impact Factor: 3.02 · DOI: 10.1021/bi500797d · Source: PubMed

CITATIONS

3

READS

58

6 AUTHORS, INCLUDING:



Ross Bathgate

University of Melbourne

231 PUBLICATIONS 6,170 CITATIONS

[SEE PROFILE](#)



Shoni Bruell

The Florey Institute of Neuroscience and Mental Health

2 PUBLICATIONS 8 CITATIONS

[SEE PROFILE](#)



John Wade

The Florey Institute of Neuroscience and Mental Health

292 PUBLICATIONS 6,702 CITATIONS

[SEE PROFILE](#)



Emma J Petrie

Walter And Eliza Hall Institute For Medical Research

16 PUBLICATIONS 230 CITATIONS

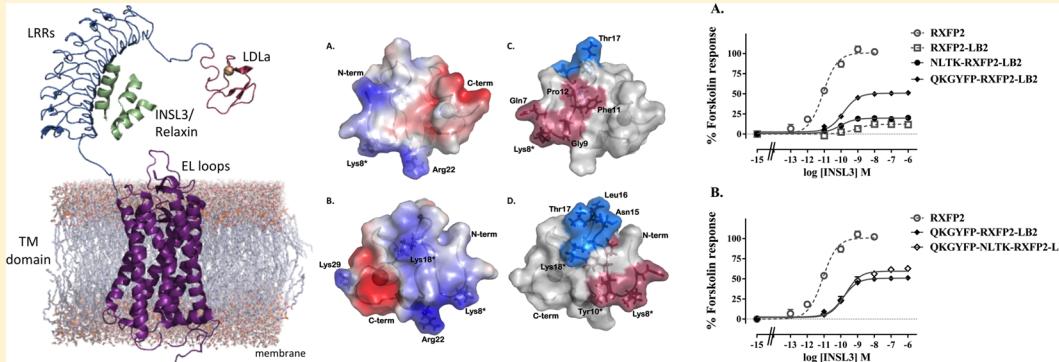
[SEE PROFILE](#)

Mapping Key Regions of the RXFP2 Low-Density Lipoprotein Class-A Module That Are Involved in Signal Activation

Roy C. K. Kong,^{†,‡} Ross A. D. Bathgate,^{*,†,‡} Shoni Bruell,^{†,‡} John D. Wade,^{‡,§} Paul R. Gooley,^{*,†} and Emma J. Petrie[†]

[†]Department of Biochemistry and Molecular Biology, The Bio21 Molecular Science and Biotechnology Institute, [‡]Florey Institute of Neuroscience and Mental Health, and [§]School of Chemistry, University of Melbourne, Parkville, Victoria 3010, Australia

Supporting Information



ABSTRACT: The peptide hormone INSL3 and its receptor, RXFP2, have co-evolved alongside relaxin and its receptor, RXFP1. Both RXFP1 and RXFP2 are G protein-coupled receptors (GPCRs) containing the hallmark seven transmembrane helices in addition to a distinct ectodomain of leucine-rich repeats (LRRs) and a single low-density lipoprotein class-A (LDLa) module at the N-terminus. RXFP1 and RXFP2 are the only mammalian GPCRs known to contain an LDLa, and its removal does not perturb primary ligand binding to the LRRs; however, signaling is abolished. This presents a general mechanism whereby ligand binding induces a conformational change in the receptor to position the LDLa to elicit a signal response. Although the LDLa interaction site has not been identified, the residues important to the action have been mapped within the RXFP1 LDLa module. In this study, we comprehensively study the RXFP2 LDLa module. We determine its structure using nuclear magnetic resonance (NMR) and concurrently investigate the signaling of an RXFP2 with the LDLa removed (RXFP2-short), confirming that the LDLa is essential to signaling. We then replaced the LDLa with the second ligand binding module from the LDL receptor, LB2, creating the RXFP2–LB2 chimera. Unlike that in the equivalent RXFP1–LB2 chimera, signaling is rescued albeit modestly. Guided by the NMR structure, we dissected regions of the RXFP2 LDLa to identify specific residues that are important to signal activation. We determine that although the module is important to the activation of RXFP2, unlike the RXFP1 receptor, specific residues in the N-terminus of the domain are not involved in signal activation.

The relaxin and insulin-like peptide-3 (INSL3) receptors, relaxin family peptide receptor-1 and -2 (RXFP1 and -2, respectively), belong to the leucine-rich repeat (LRR)-containing G protein-coupled receptor (LGR) family. This family consists of three subclasses, defined by the number of LRRs and the presence or absence of an N-terminal LDLa module. Type A LGRs contain five to seven LRRs and include the glycoprotein hormone receptors for follicle stimulating hormone (FSHr), thyroid stimulating hormone (TSHr), and luteinizing hormone (LHr). Type B LGRs contain 11–15 LRRs and have been deorphanized as the receptors for R-spondins.¹ Type C LGRs are distinguished by the addition of a low-density lipoprotein class-A (LDLa) module N-terminal to the LRRs and are further subtyped as C1 (with a single LDLa module) or C2 (with multiple modules).² There are only two C1 LGRs found in mammals, LGR7 and LGR8, with all mammalian

species having orthologs of these, now known as RXFP1 and RXFP2, respectively (Figure 1). C2 LGRs, on the other hand, are found in species such as the pond snail, human body lice, and sea urchin.² The evolution of type C LGRs remains unclear as RXFP1 and RXFP2 are the only type C LGRs that have been identified in humans,³ and no other human GPCRs containing a LDLa module have been described.

LDLa modules are well-characterized for their roles in lipoprotein metabolism for they were initially described as the seven repeating domains in the low-density lipoprotein (LDL) receptor⁴ and other related receptors.⁵ These modules are typically 4 kDa in size, contain three conserved disulfide bonds,

Received: June 26, 2014

Revised: June 30, 2014

Published: July 1, 2014

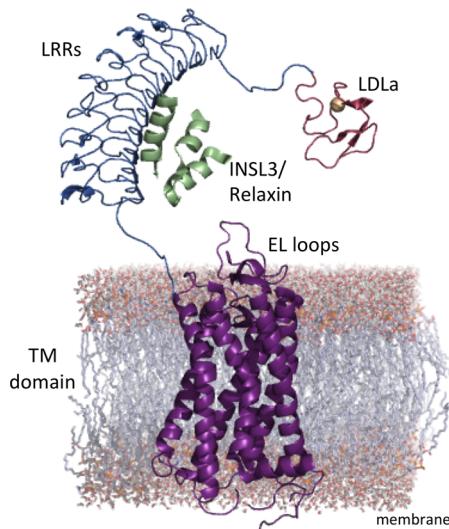


Figure 1. Schematic of the domain structure of RXFP1 and -2. The rhodopsin-like seven TM helices are embedded in the membrane. An uncharacterized linker presents the ectodomain of LRR followed by an LDLa module tethered to the LRRs via another linker. The ligand binds to the LRRs, but little is known about the molecular mechanism that leads to signal activation.

and possess a structurally essential calcium ion within the C-terminal region, ligated via the consensus motif, DxxxDxxDx-xDE, where x is any residue. Since their initial characterization, other LDLa modules have been described in a number of functionally distinct proteins involved in processes as diverse as viral entry,⁶ breast cancer invasion and metastasis,⁷ normal development of the nervous system,⁸ normal cell differentiation,⁹ and interaction with collagen.¹⁰

The hormones relaxin and INSL3 are historically associated with roles in reproductive function. The actions of relaxin on the female reproductive organs are well-characterized (reviewed in ref 11) as are its pleiotropic roles in vasodilation¹² and collagen turnover.¹³ INSL3 is predominantly expressed in the Leydig cells of the testis in males and in the ovarian thecal cells in females (reviewed in ref 11). In males, INSL3 is essential for testicular descent *in utero* (reviewed in ref 14) and in adults has been implicated in Leydig cell function (reviewed in ref 15), germ cell maturation,¹⁶ and bone metabolism (reviewed in ref 17). Interestingly, recent studies suggest that RXFP2 is associated with horn growth in sheep¹⁸ and cows,¹⁹ suggesting the action of INSL3 on bone is a common feature in male mammals. In females, thecal-derived INSL3 may have important roles in follicle development.^{20,21}

Relaxin and INSL3 have an insulin-like fold and are produced as a single chain with B-C-A domains. It is thought that upon maturation, the connecting peptide (C-peptide) is cleaved to reveal the A- and B-chains tethered by two disulfide bonds. Numerous structure-function studies have led to the current model of relaxin and INSL3 binding to their respective receptors.^{22–27} Both hormones bind to the LRRs with high affinity through key residues in the B-chain, while the A-chain is thought to interact with the transmembrane (TM) exoloops.^{28,29} For GPCRs without extensive ectodomains, binding of peptide ligands within the exoloops is sufficient to induce a conformational change within the TM helices, in particular, TMS and TM3, TM7 and TM2, and TM6, that leads to G protein coupling (reviewed in ref 30). However, removal of

either the RXFP1 or RXFP2 LDLa module results in receptors with full hormone binding affinity but a nonsignaling phenotype.³¹ Additionally, a chimeric receptor in which the LDLa module of RXFP1 was replaced with the second ligand binding domain of the LDL receptor (LB2) bound relaxin with an affinity comparable to that of RXFP1 but was unable to signal.³² These data suggest that the LDLa module of RXFP1 and RXFP2 is involved in a novel mechanism of GPCR activation and that this action is mediated through specific side chain interactions at an interface involving the LDLa and a second unknown site.³³

The unique paradigm of signal activation via the LDLa module presents novel therapeutic potential for agonists and/or antagonists targeting RXFP1 and RXFP2. We have previously demonstrated that the recombinant LDLa module can antagonize RXFP1 signaling.³¹ Additionally, expression of the recombinant RXFP1 LDLa module was shown recently to downregulate markers associated with the tumorigenesis of prostate cancer cells.³⁴ The RXFP2-INSL3 system is implicated in fertility regulation³⁵ and in germ cell survival in both sexes.¹⁶ Importantly, treatment of rats with an RXFP2 antagonist was demonstrated to disrupt testis function.³⁶ This raises the potential of exploiting the LDLa module to therapeutically modulate receptor activity. However, the interface at which the LDLa module binds and the molecular details behind how it triggers signaling are unknown, making it difficult to design therapeutic molecules to block this interaction. For this reason, it is essential to understand the surface of the LDLa module and the key residues of the molecule that are important to whole receptor activity.

Mutational analysis studies that aimed to dissect the role of the LDLa module in receptor activation have focused mainly on RXFP1.^{31–33,37} Consequently, we undertook a functional dissection on the LDLa module of RXFP2, the only other human GPCR with an LDLa module, to improve our understanding of the molecular mechanics and specificities surrounding its unique activation. We first tested if RXFP2 could signal through alternate pathways in an LDLa-independent manner. Subsequently, we determined the NMR solution structure of the RXFP2 LDLa module to understand the distribution of surface charge across the domain. We next generated the RXFP2-LB2 chimera by replacing the RXFP2 LDLa module with the structurally similar but functionally distinct module, the second ligand binding domain of the LDL receptor, LB2, to establish the need for specific LDLa residues in RXFP2 signaling. Finally, gain- and loss-of-function experiments were performed to identify important LDLa signaling regions and to define the side chain properties involved in protein–protein interactions that form the active receptor.

EXPERIMENTAL PROCEDURES

Generation of HEK293T Cell Lines Stably Expressing RXFP2 without the LDLa Module. A HEK293T cell line (ATCC #CRL-1573, American Type Tissue Culture Collection) stably expressing RXFP2 without the LDLa module (RXFP2-short) was generated via retroviral transduction.³⁸ The DNA expressing RXFP2-short was cloned in frame via a 5' BamHI and a 3' XhoI restriction site into the MSCV-IRES-GFP retroviral expression vector (a kind gift from R. Hannan, Peter MacCallum Cancer Centre, Melbourne, Australia) in the 5'-3' direction between the 5' long terminal repeats (LTRs) and the internal ribosomal entry site (IRES) sequence. Green fluorescent protein (GFP) follows the IRES. Following cloning

Table 1. Comparison of Receptor Cell Surface Expression (% RXFP2 expression), INSL3 Binding (pIC_{50}), Maximal cAMP Responses (E_{max} ; % forskolin-stimulated activity), and INSL3 Potency (pEC_{50}) of RXFP2–LB2 Chimeric Receptors and Several Single-Site Mutants to Those of RXFP2^a or RXFP2–LB2^{b,a}

receptor construct	cell surface expression	pIC_{50}	E_{max}	<i>p</i> value	pEC_{50}	<i>p</i> value
RXFP2	100.00 (16)	8.28 ± 0.06 (4)	103.10 ± 5.33	—	11.09 ± 0.09 (10)	—
RXFP2–LB2	107.7 ± 6.94 (5)	8.52 ± 0.10 (3)	11.98 ± 0.79	<0.001 ^a	9.38 ± 0.24 (4)	<0.001 ^a
Single-Region Gain-of-Function RXFP2–LB2 Chimeras						
QKGYFP–RXFP2–LB2	95.29 ± 5.14 (9)	8.52 ± 0.09 (3)	51.37 ± 2.55	<0.001 ^b	9.87 ± 0.08 (8)	<0.001 ^a
NLTK–RXFP2–LB2	94.34 ± 4.9 (6)	8.45 ± 0.11 (3)	20.91 ± 2.29	<0.05 ^b	9.94 ± 0.12 (9)	<0.001 ^a
Double-Region Gain-of-Function RXFP2–LB2 Chimeras						
QKGYFP NLTK–RXFP2–LB2	92.51 ± 4.54 (6)	8.46 ± 0.11 (3)	60.19 ± 1.85	<0.001 ^b	9.69 ± 0.26 (3)	<0.001 ^a
Single-Site RXFP2 Mutants						
K8A-RXFP2	107.30 ± 6.52 (5)	8.07 ± 0.13 (3)	108.00 ± 5.18	ns ^a	10.43 ± 0.08 (6)	<0.001 ^a
Y10M-RXFP2	98.76 ± 6.00 (7)	8.37 ± 0.15 (3)	110.80 ± 3.93	ns ^a	10.41 ± 0.07 (6)	<0.001 ^a

^aThe numbers of independent assays performed are given in parentheses.

and sequencing, the MSCV-RXFP2-short-IRES-GFP construct was utilized for retrovirus production by cotransfected HEK293T cells on a 10 cm cell culture dish (Nunc) with the construct together with the Amphi packaging plasmid (a kind gift from R. Hannan) for 48 h. Retrovirus secreted into the medium was collected and passed through a 30 mm diameter 0.45 μ m Durapore PVDF syringe filter (Millipore) to remove any cell debris prior to storage at –80 °C.

HEK293T cells seeded on a 10 cm cell culture dish were transduced with the harvested retrovirus by replacing the cell culture medium with 10 mL of the retrovirus-containing medium, mixed with Polybrene (Millipore) to increase the transduction efficiency. Two rounds of 24 h transduction were performed before the cells were allowed to recover for 48 h in Dulbecco's Modified Eagle's Medium (DMEM) (Multicel) supplemented with 10% heat-inactivated fetal bovine serum (FBS), 1% penicillin/streptomycin, and 1% L-glutamine (termed complete DMEM). The cells were then transferred and grown to confluence in a 175 cm² flask. Transduced cells were sorted into three distinct populations on the basis of the amount of GFP fluorescence using fluorescence-activated cell sorting (FACS) on a MoFlo XDP cell sorter (Beckman Coulter): high, medium, and low GFP (hence receptor) expression levels. These cells were grown to confluence, and the receptor expression on the cell surface was assessed as described below.

Secreted Alkaline Phosphatase (SEAP) Reporter Gene Assay. The SEAP reporter gene assay was performed as described in ref 39 using reporter gene constructs (Clontech) that contained the secreted alkaline phosphatase (SEAP) reporter gene under the control of specific transcription factor consensus sequences (CRE, cAMP response element; GRE, glucocorticoid response element; NF- κ B, nuclear factor of κ B cells; AP1, activator protein 1; SRE, serum response element; NFAT, nuclear factor of activated T cells; Myc, E-box DNA binding element; HSE, heat shock element). Briefly, HEK293T cells stably expressing RXFP2³⁹ or RXFP2-short seeded on 12-well plates (Nunc) were first transfected with the desired plasmids for 30 h at a 1:1 ratio for a reporter gene of interest and the mCherry plasmid. Cells were then partially serum starved for 18 h in DMEM supplemented with 0.5% FBS, 1% penicillin/streptomycin, and 1% L-glutamine prior to stimulation using 100 nM INSL3. Stimulation occurred for 18 h (a time point chosen as the peak response time),³⁹ after which medium samples were collected, frozen at –20 °C, and had their SEAP protein content quantified using the BD Great

EscAPE SEAP fluorescence detection kit (Clontech). The transfection efficiency of the reporter gene constructs could affect SEAP protein expression and hence bias the results. Therefore, the mCherry plasmid that constitutively expressed a red fluorescent monomer served as a control to measure transfection efficiency between and within experiments and also served as a transfection control. Upon medium sample collection, cells in each well were lysed in 500 μ L of lysis buffer [50 mM Tris base, 150 mM NaCl, 1 mM EDTA, and 1% Triton X-100 (pH 7.4)] and transferred to a 96-well OptiPlate with white opaque walls. mCherry fluorescence was then measured using a Victor Multilabel Plate Reader with excitation and emission wavelengths of 587 and 610 nm, respectively.

Data analysis was performed as described in ref 39 using GraphPad Prism 5. Briefly, discrepancies in cell transfection efficiency were first corrected by standardizing the SEAP protein measurement for the fluorescence measurement of the mCherry protein in each individual well. The transformed SEAP readings were then expressed as the fold change in response to peptide stimulation compared to the effect of vehicle alone. Data from three independent assays was pooled and displayed as means ± SEM (standard error of the mean). Unpaired two-tailed *t* tests were performed on the transfection-standardized and background-corrected data, comparing the peptide- and vehicle-treated cells for each reporter gene construct.

Cloning of the RXFP2–LB2 Chimera. Human RXFP2 with its LDLA module substituted with the LB2 module (RXFP2–LB2) was created by subcloning into the RXFP1–LB2 pcDNA3.1/Zeo construct.³² RXFP2 was amplified from Gly-45 to Ser-718 with flanking EcoRI and BamHI restriction sites for ligation into the RXFP1–LB2 chimera, in place of the RXFP1 sequence. As there are two artificial EcoRI restriction sites flanking the LB2 module, site-directed mutagenesis was used to delete the two additional amino acids (Glu-Phe) at each end of the module encoded by the restriction sites. The resulting RXFP2–LB2 chimera was sequenced to verify the correct insertion and lack of any additional modifications.

Site-Directed Mutagenesis. Site-directed mutagenesis was performed via the QuikChange PCR mutagenesis kit (Stratagene). Polymerase chain reaction was performed with RXFP2 or the RXFP2–LB2 chimera as the template using the primers listed in Table S1 of the Supporting Information. The complete sequence of each generated plasmid was verified via fluorescence-based cycle DNA sequencing.

Table 2. Comparison of Receptor Cell Surface Expression, Maximal cAMP Responses (E_{max} ; % RXFP2 forskolin-stimulated activity), and INSL3 Potency (pEC₅₀) of Single-Site Mutants to Those of RXFP2^a

receptor construct	pEC ₅₀	E_{max}	p value	expression	p value
Mutant Assay 1					
RXFP2	10.51 ± 0.06 (4)	100 ± 4.76	—	100 ± 3.9 (4)	—
K8M-RXFP2	10.26 ± 0.06 (3)	80.92 ± 5.37	ns	121.09 ± 6.8 (3)	<0.05
Y10F-RXFP2	10.47 ± 0.11 (3)	67.00 ± 5.04	<0.01	102.5 ± 4.49 (3)	ns
K8M/Y10F/K18M-RXFP2	10.37 ± 0.20 (3)	89.81 ± 14.84	ns	79.44 ± 4.44 (3)	<0.05
Mutant Assay 2					
RXFP2	10.29 ± 0.05 (4)	100 ± 11.59	—	100 ± 3.9 (4)	—
Q7A-RXFP2	10.49 ± 0.04 (3)	72.60 ± 11.37	ns	106.3 ± 3.49 (3)	ns
N15A-RXFP2	10.37 ± 0.21 (3)	71.29 ± 16.63	ns	40.99 ± 4.1 (3)	<0.001
K18A-RXFP2	10.31 ± 0.07 (3)	104.1 ± 24.16	ns	109.37 ± 6.45 (3)	ns
P12A-RXFP2	10.30 ± 0.01 (3)	57.92 ± 1.55	<0.05	99.64 ± 7.48 (3)	ns

^aThe numbers of independent assays performed are given in parentheses.

cAMP Activity Assay. cAMP activity assays were performed as previously described^{31,40} with minor modifications. For the receptor constructs listed in Table 1, HEK293T cells seeded on 96-well plates (Nunc) were transfected with plasmids of the receptor construct of interest and the pCRE (cAMP response element) β-galactosidase reporter in a 1:2.5 ratio with the empty pcDNA3.1 vector for 18 h. For the constructs listed in Table 2, cells were transfected with plasmids of the receptor construct of interest and the pCRE β-galactosidase reporter in a 1:100:100 ratio with the empty pcDNA3.1 vector for 18 h. It is important to note experiments with mutants listed in Table 2 were performed under slightly different transfection conditions, and at different times; however, they were conducted in parallel to an RXFP2 control.

After transfection, cells were incubated with 100 μL of synthetic human INSL3 (prepared in house)⁴¹ in increasing concentrations, made up in complete DMEM. Positive and negative control stimulations using 100 μL of 5 μM forskolin (Sigma-Aldrich, St. Louis, MO) and 100 μL of complete DMEM, respectively, were also performed. After 6 h, the medium was aspirated and cells were frozen overnight at −80 °C. Measurement of cAMP-driven β-galactosidase expression based on its interaction with its substrate, chlorophenol red-β-D-galactopyranoside (CPRG) (Roche Diagnostics, Basal, Switzerland), then followed. Data were analyzed using GraphPad Prism 5 and are presented as means ± SEM of the percent normalized response compared to that of 5 μM forskolin for constructs listed in Table 1 and 5 μM forskolin RXFP2 responses listed in Table 2 in each case from at least three experiments with triplicate determinations within each assay. Significance was determined using one-way analysis of variance (ANOVA) followed by a Tukey–Kramer multiple-comparison post-test.

Receptor Expression Assay. All constructs were fused with an N-terminal FLAG epitope tag to allow monitoring of receptor expression using an anti-FLAG antibody assay as previously described.⁴⁰ HEK293T cells on 24-well plates (Nunc) were transfected with the plasmid of the receptor construct of interest and empty pcDNA3.1/Zeo in a 1:1 ratio. After 18 h, receptor expression was assessed in triplicate using an anti-FLAG mouse monoclonal primary antibody (Sigma-Aldrich) and a goat anti-mouse ALEXA Fluo 488 secondary antibody. Data were from at least three independent assays and analyzed using GraphPad Prism 5 and are displayed as means ± SEM of the wild-type receptor percentage expression. Unpaired two-tailed t tests were performed to gauge the significance of

the expression of each receptor of interest above the background as well as compared to the expression of RXFP2.

Whole Cell Competition Binding Assay Using Europium-Labeled INSL3 (Eu-INSL3). HEK293T cells on 48-well plates (Nunc) were transfected with the plasmid of the receptor construct of interest for 18 h. Ligand competition binding was then assessed in triplicate as described in ref 42 using increasing concentrations of synthetic human INSL3 to compete with Eu-labeled INSL3 (1 nM)⁴³ in receptor binding buffer [20 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), 1.5 mM CaCl₂, 50 mM NaCl, and 0.01% NaN₃] containing 1% BSA for 1 h at 22 °C. At the end of the incubation period, the buffer was aspirated and cells were washed with PBS; 100 μL of a Delfia enhancement solution (PerkinElmer) was then added to each well, and the cells were transferred to a 96-well OptiPlate with white opaque walls for fluorescence measurement on a Victor Multilabel Plate Reader with excitation and emission wavelengths of 395 and 619 nm, respectively. Data from at least three independent experiments were analyzed using GraphPad Prism 5 and are presented as means ± SEM of the percentage of specific binding. Significance was determined using one-way ANOVA followed by a Tukey–Kramer multiple-comparison post-test.

Cloning, Expression, and Purification of the RXFP2 LDLa Module. The RXFP2 module was cloned and prepared as previously reported for the RXFP1 LDLa module.³² Briefly, the module was cloned from the pcDNA3.1/Zeo plasmid containing the human RXFP2 sequence. Residues Met-1–Gly-44 were amplified via PCR with BamHI and XhoI restriction sites to allow ligation into the pGEV2 vector. The resulting protein is a fusion to the GB1 protein followed by a thrombin cleavage site and the LDLa module. The construct was designed so that the Gly and Ser residues that remain after thrombin cleavage are N-terminal to Met-1 of the native sequence of RXFP2. The resulting plasmid was transformed into T7 bacterial expression strain BL21(DE3)trxB (Novagen) for the expression of the GB1–RXFP2LDLa fusion protein. Freshly transformed cells were used for all protein expression. Protein generated for structure determination was uniformly isotopically labeled with ¹⁵N or with ¹³C and ¹⁵N by growing cultures in a 2 L Biostat A-plus fermenter (Sartorius) containing 1 L of minimal medium with ¹⁵NH₄Cl and d-[¹³C]glucose as the sole sources of nitrogen and carbon, respectively. Fermentation was conducted as described by Cai et al.⁴⁴ Protein expression was induced by the addition of 1 mM isopropyl 1-thio-β-D-galactopyranoside (IPTG) for 2.5 h, at

which point cells were harvested, pelleted, and stored at -20°C until they were further purified.

Fusion protein was purified using IgG-Sepharose (GE Healthcare) via the manufacturer's instructions. The eluted protein was buffer-exchanged into 50 mM Tris-HCl and 150 mM NaCl (pH 7.5) and concentrated to 100 $\mu\text{g}/\text{mL}$ using a 3 kDa cutoff Vivapin centrifugal concentrator (Sartorius). The protein concentration was adjusted to 100–300 $\mu\text{g}/\text{mL}$ in refolding buffer [3 mM GSH, 0.3 mM GSSG, 50 mM Tris-HCl, 150 mM NaCl, and 2.5 mM CaCl₂ (pH 7.5)]. The sample was incubated overnight at 4°C while being stirred.

The GB1 fusion protein was cleaved from the LDLa module overnight by incubating with 1 unit of thrombin protease (GE Healthcare) per milligram of fusion protein. The cleaved GB1 was separated from the LDLa module via reverse phase high-performance liquid chromatography (RP-HPLC) using a Jupiter Proteo 4 μm , 90 Å column, using a 20% (buffer A) to 70% (buffer B) gradient, where buffer A consisted of 0.1% trifluoroacetic acid and buffer B consisted of 80% acetonitrile and 0.1% trifluoroacetic acid. The eluted protein was lyophilized and stored at -20°C .

Collection of NMR Data. For ¹H, ¹⁵N, and ¹³C assignments, NMR data were collected on 2.0 mM ¹⁵N-labeled and 2.5 mM ¹³C- and ¹⁵N-labeled RXFP2 LDLa in 50 mM imidazole (pH 5.5) and 10 mM CaCl₂ at 25°C on a Bruker Avance II 800 MHz spectrometer equipped with a triple-resonance cyroprobe. Assignments were made from three-dimensional (3D) CBCA(CO)NH, C(CO)NH, ¹³C NOESY-HSQC (150 ms mixing time), HBHA(CO)NH, HC(CO)NH, HCCCH-TOCSY, HNCACB, HNCO, HNHB, and ¹⁵N NOESY-HSQC (150 ms mixing time) spectra. Additionally, aromatic residues were confirmed through the collection of two-dimensional TOCSY and NOESY spectra on an unlabeled 0.5 mM RXFP2 LDLa in 50 mM *d*₄-imidazole (pH 5.5) and 10 mM CaCl₂. All spectra were processed using NMRPipe⁴⁵ and typically Fourier-transformed after applying Lorentz-to-Gauss window functions in the direct dimension and cosine bells in the indirect dimensions. Data were analyzed using CCPNMR Analysis 2.15.⁴⁶

Structure Calculations. Distance constraints were obtained from 3D ¹⁵N-edited NOESY-HSQC and ¹³C-edited HSQC-NOESY spectra. NOE data were assigned using CYANA 2.0^{47,48} with rounds of both manual and automatic assignment. In the final round of structure calculation, a total of 389 upper-distance constraints and 28 dihedral angles generated from TALOS+⁴⁹ (Table 2 of the Supporting Information) were used to calculate the structure. In addition, a calcium ion was included via the generation of "pseudolinks" from the C-terminus and inclusion of pseudo-NOEs based on the ligation of the calcium ion from the crystal structure for LB5 [Protein Data Bank (PDB) entry 1AJJ]⁵⁰ to assigned residues D27(OD1), D31(OD2), D37(OD2), and E38(OE2). Calculations generated 100 structures, with the 20 lowest-energy structures being selected for analysis. The quality of structures was validated using PROCHECK-NMR⁵¹ and MacPyMol (PyMOL Molecular Graphics System, version 1.5.0.4, Schrödinger, LLC) and MOLMOL⁵² were used to inspect structures, calculate ring current shifts and surface areas, and produce figures.

RESULTS

NMR Structure of the RXFP2 LDLa Module. The sequences of the LDLa modules of RXFP1 and RXFP2 are

72% identical, between the first and sixth cysteine residues (Figure 2), so unsurprisingly, the structures are very similar and



Figure 2. Alignment of human RXFP2 and RXFP1 LDLa module sequences compared to LB2. The numbering is based on the RXFP2 sequence in which methionine is known as residue 1. Identical residues are highlighted in black, and similar residues are highlighted in gray; conserved cysteine residues are highlighted in red.

have a fold typical of other LDLa modules found in the Protein Data Bank. The final ensemble reveals that between residues Cys-6 and Cys-41 the structure is well-defined (Figure 3A), with a backbone root-mean-square deviation (rmsd) of 0.54 Å and a heavy atom rmsd of 0.90 Å. There are no violations of experimental distance and angle constraints exceeding 0.2 Å and 3°, respectively (Table 2 of the Supporting Information). The N-terminal region consists of an antiparallel β -sheet running from residue Phe-11 to Cys-13 and from residue Lys-18 to Pro-21, immediately followed by a short 3₁₀ helix involving Arg-22–Phe-24. The overall fold is dominated by three conserved disulfide bonds and ligation of a calcium ion by the acidic motif in the C-terminal region (Figure 3B). Because of the size of the molecule, those residues with side chains not directly involved in the ligation of the calcium ion in the C-terminus or the cysteines that form disulfide bonds are largely surface-exposed. The side chains of Tyr-10, from the loop that precedes the first β -strand, and Pro-21, at the end of the 3₁₀ helix, pack against each other and are 27 and 20% surface-exposed, respectively, values that are relatively low compared to those of other residues; however, the hydroxyl group of Tyr-10 is oriented toward the solvent. The aromatic side chain of Phe-11 is oriented toward the core of the structure, showing NOEs to Asp-37, which also ligates calcium, and Leu-20 extending from the second β -strand, suggesting that Phe-11 makes significant structural contributions, and this is consistent with good conservation of phenylalanine residues at the equivalent position in most LDLa modules. The mean structure of the RXFP2 LDLa module overlays well with the second ligand binding repeat (LB2) from the LDLR (backbone rmsd of 1.8 Å) and the module of RXFP1 (backbone rmsd of 1.3 Å), allowing us to make meaningful comparisons between these structures (Figure 3C).

Considering the charge distribution on the surface of the molecule, the C-terminal region is predominantly acidic via the contributions of the residues that ligate the calcium ion. The remainder of the surface is largely hydrophobic apart from a few notable protrusions of basic residues. On the first face of the molecule, Lys-8 and Arg-22 extend from the structure to create a largely basic "base" to the molecule. On the converse side of the molecule, Lys-29 contributes a single basic patch within the largely acidic C-terminal region while Lys-18 extends from the loop that connects the β -strands (Figure 4A,B).

Investigating the Role of the RXFP2 LDLa Module in Signal Activation. The signaling profile of RXFP2 was previously established via reporter gene assays.³⁹ These assays measure eight different response elements driving the expression of SEAP: cAMP response element (CRE), activator protein 1 (AP1), serum response element (SRE), nuclear factor of activated T-cells (NFAT), E-box DNA binding element (myc), heat shock response element (HSE), glucocorticoid

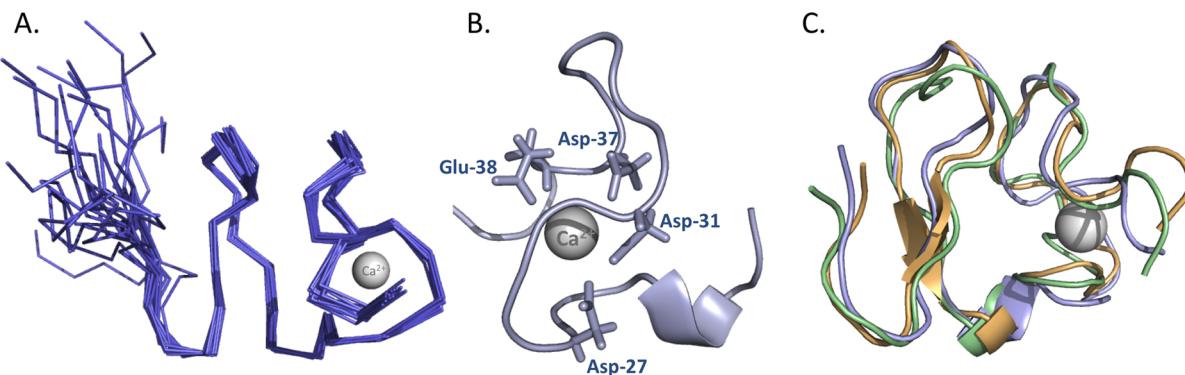


Figure 3. Structure of the RXFP2 LDLa module. (A) Overlay of the 20 lowest-energy structures. The calcium ion is shown as a gray sphere. (B) The calcium ion is ligated by the side chain carbonyls of four residues. The structure is rotated around the y axis by 180° compared to panel A. (C) Overlay of the RXFP2 LDLa module (purple) with the RXFP1 LDLa (gold) and LB2 (green) modules. The unstructured residues before the first Cys are not shown. The orientation of the structures is the same as in panel A. Figures generated using PyMol.

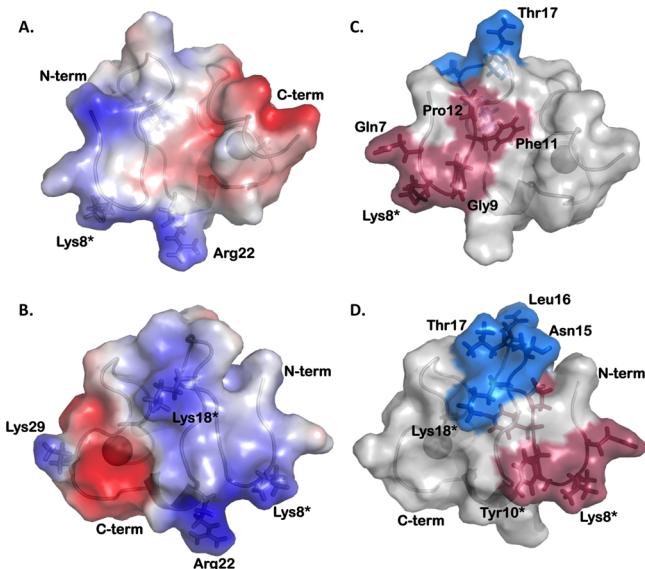


Figure 4. (A and B) Surface charge distribution of the RXFP2 LDLa module. The surface charge of the module is heterogeneous because of the acidic residues that ligate calcium located in the C-terminal region and the basic residues of Lys-8, Lys-18, and Arg-22 located in the N-terminal region. (A) “Front” face of the molecule with the N-terminus positioned on the left and the C-terminus on the right. (B) “Back” of the molecule, where the N-terminus is positioned on the right and the C-terminus on the left. (C and D) Surface of the module investigated in the “add-back” RXFP2–LB2 chimeras. Region 1 is colored maroon and consists of Glu-7, Lys-8, Gly-9, Tyr-10, Phe-11, and Pro-12. Region 2 is colored teal and consists of residues Asn-15, Leu-16, Thr-17, and Lys-18. Panel C shows the front of the molecule and panel D the back as in panels A and B, respectively. The residues that are equivalent to the key resides in RXFP1 (Leu-7, Tyr-9, and Lys-17) are denoted with asterisks. Figures and charge distribution represented by PyMol.

response element (GRE), and nuclear factor κ B (NF- κ B). RXFP2 was previously demonstrated to activate only the CRE reporter, in contrast to RXFP1, which activates both CRE and GRE while inhibiting NF- κ B.^{33,39} We have previously demonstrated that the RXFP2 receptor lacking the LDLa module (RXFP2-short) is still able to bind INSL3 but unable to signal via the cAMP pathway,³¹ which is confirmed in the study presented here (Table 1). As the incorporation of the LDLa module into the signaling mechanism of these GPCRs is

unique, we were interested in testing the possibility that binding of INSL3 to RXFP2-short may result in the activation of alternative pathways in an LDLa-independent manner.

To ensure a stable level of receptor expression at the cell surface and limit assay variability, cell lines stably expressing each receptor of interest were used. An RXFP2 stable cell line was already available,³⁶ and an RXFP2-short cell line was generated using retrovirus to transduce HEK293T cells. Transduced cells were sorted on the basis of the amount of GFP fluorescence into three populations: high, medium, and low GFP (hence receptor) expression levels. Receptor expression and ligand binding assays demonstrated that the strongly expressing cell line did indeed have the highest level of expression of RXFP2-short and importantly that these cells bound INSL3 with the same affinity as the RXFP2-expressing cells (data not shown). Thus, this cell line was used for the SEAP reporter assays.

RXFP2 cells were stimulated with 100 nM INSL3 to screen for GPCR-relevant pathways associated with RXFP2 activation. Consistent with previous results,³⁹ only the CRE reporter gene was affected, resulting in significantly elevated levels of secretion of SEAP into the cell culture medium [3.26 ± 0.34 -fold of the vehicle treatment ($n = 3$; $p < 0.01$)] (Figure 5A). In contrast, stimulation of RXFP2-short stable cells with 100 nM INSL3 did not induce any change in the transcriptional activity of any of the reporters; hence, SEAP levels in the cell culture medium were comparable between the peptide-stimulated and vehicle-treated cells (Figure 5B). Importantly, the positive control construct used in every experiment produced a transfection-standardized SEAP response that was significantly higher than that of the negative control construct (data not shown). These results confirm that RXFP2 activates only the CRE reporter in HEK293T cells and that the LDLa module is essential for this activity.

Construction and Characterization of the RXFP2–LB2 Chimera. The sequences of the LDLa modules of RXFP1 and RXFP2 are weakly identical (~38%) to that of the second ligand binding domain (LB2) of the LDL receptor; however, the LDLa modules of RXFP1 and RXFP2 have very similar backbone folds (Figures 2 and 3C). We previously exploited the similarities in fold to demonstrate that the key to signal activation by the RXFP1 LDLa module lies within the presentation of specific residues within the N-terminal region of the domain.^{32,33} By generating an RXFP1–LB2 chimera, we demonstrated that relaxin binding was unaltered by the

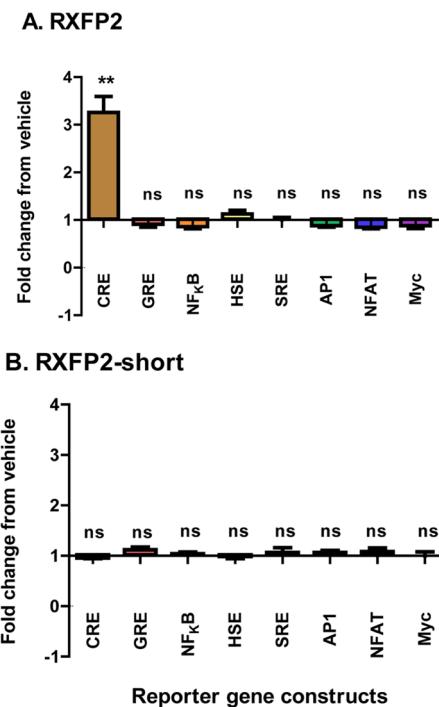


Figure 5. Reporter gene responses (fold SEAP change from vehicle) upon stimulation of (A) RXFP2 or (B) RXFP2-short stably expressed in HEK293T cells using 100 nM INSL3. Data are the fold changes of response from vehicle and are means of three independent experiments.

replacement of the module but the chimera was unable to signal.³² In this study, we have produced an RXFP2–LB2 construct and tested its ability to bind ligand and signal.

As expected, the RXFP2–LB2 chimeric receptor expressed at the cell surface at levels similar to that of RXFP2 [$107.69 \pm 6.94\%$ of the RXFP2 expression level (Table 1)] and bound INSL3 with a similar affinity (Figure 6A and Table 1). The LB2 module therefore does not disrupt cell surface trafficking or ligand binding as has previously been demonstrated for the RXFP1–LB2 chimera.³² However, unlike the RXFP1–LB2 chimera, which is incapable of generating a signaling response, the RXFP2–LB2 chimeric receptor was able to elicit a cAMP response, albeit with significantly reduced INSL3 potency [pEC₅₀ of 9.38 ± 0.24 compared to a value of 11.09 ± 0.09 for RXFP2 ($p < 0.001$)] (Figure 6B and Table 1) and markedly reduced maximal cAMP levels ($11.98 \pm 0.79\%$ forskolin) compared to that of the wild type [$103.10 \pm 5.33\%$ forskolin ($p < 0.001$)] (Figure 6B and Table 1). These results suggest that the LB2 module can compensate for the activity of the native RXFP2 module in a way that it cannot in RXFP1.

Identification of the RXFP2 LDL_a Activation Surface via a “Gain-of-Function” Approach. Although the RXFP2–LB2 chimeric receptor retained the ability to activate signaling, we considered the reduction significant enough to use it as a scaffold to reintroduce regions of the native RXFP2 LDL_a sequence and look for an increase in the level of intracellular cAMP production in an effort to globally map the surface of the module important for activity.

Using our NMR structure and sequence alignments, we selected two regions to add-back to the RXFP2–LB2 chimera. As the residues before the first cysteine are unstructured and not well conserved across species, we focused on the regions after these residues. Therefore, we targeted the region

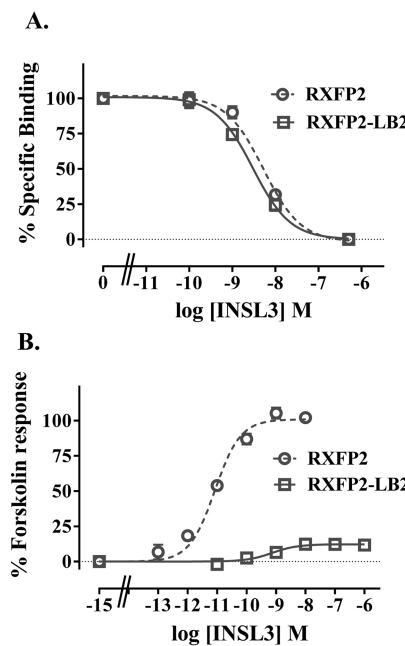


Figure 6. Activity of the RXFP2–LB2 chimera compared to that of RXFP2. (A) Competition binding using Eu-labeled INSL3. (B) INSL3-induced cAMP responses. cAMP activity is expressed as the percentage of the 5 μ M forskolin-stimulated response for each receptor, which has also been normalized for cell surface expression. Symbols represent means, and vertical bars represent the SEM of triplicate determinations from at least three independent experiments.

encompassing residues Gln-7–Pro-12, representing the first loop and the first β -strand of the structure (Figures 3 and 4). We also selected the loop between the two β -strands, which spans residues Asn-15–Lys-18 and constitutes a putative N-linked glycosylation site. The residues suggestively protruding from the 3_{10} helix, Ala-25, Phe-26, and His-27, were not included in this study as we have found mutation of these residues adversely affected folding of the recombinant LDL_a module (data not shown). The NOE network of Phe-26 is consistent with a strong role in maintaining structure, including NOEs to calcium-ligating residues Asp-29 and Asp-33, and clearly packs against Phe-13 within the core of the structure. Mutations that disrupt the structure of the module are difficult to interpret in the context of chimeric receptors; we therefore initially generated two receptor chimeras, QKGYFP–RXFP2–LB2 and NLTK–RXFP2–LB2.

All receptor chimeras were expressed at the cell surface at levels similar to that of RXFP2 (Table 1). Additionally, the INSL3 binding ability of all the receptor chimeras and mutants was unaffected by any of the mutations introduced (Table 1). Both QKGYFP–RXFP2–LB2 and NLTK–RXFP2–LB2 chimeras responded to INSL3 stimulation, generating cAMP responses stronger than that of the RXFP2–LB2 chimera, although the pEC₅₀ values did not improve. The QKGYFP–RXFP2–LB2 chimera produced a maximal response of $51.37 \pm 2.55\%$ forskolin-induced cAMP activity ($n = 8$; $p < 0.001$) (Figure 7A and Table 1), while the NLTK–RXFP2–LB2 chimera yielded a modest increase in the maximal response of $20.91 \pm 2.29\%$ forskolin-induced cAMP activity ($n = 9$; $p < 0.05$) (Figure 7A and Table 1).

When the add-back of the regions in the QKGYFP–RXFP2–LB2 and NLTK–RXFP2–LB2 chimeras were combined, the QKGYFP NLTK–RXFP2–LB2 chimera produced

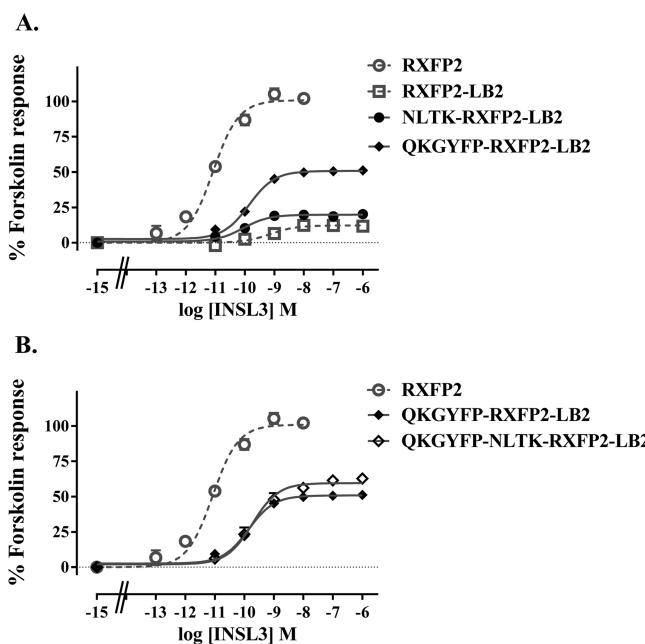


Figure 7. INSL3-induced cAMP response of (A) RXFP2–LB2, QKGYFP–RXFP2–LB2, and NLTK–RXFP2–LB2 chimeras and (B) QKGYFP–RXFP2–LB2 and QKGYFP NLTK–RXFP2–LB2 chimeras compared to that of wild-type RXFP2. cAMP activity is expressed as a percentage of the 5 μ M forskolin-stimulated response for each receptor. Symbols represent means, and vertical bars represent the SEM of triplicate determinations from at least three independent experiments.

$60.19 \pm 1.85\%$ forskolin-induced cAMP activity ($n = 3$) upon INSL3 stimulation (Figure 7B and Table 1), which was significantly higher than that of either mutation alone ($p < 0.05$ compared to QKGYFP–RXFP2–LB2 and $p < 0.001$ compared to NLTK–RXFP2–LB2) (Table 1).

The pEC₅₀ values did not improve from that of the RXFP2–LB2 chimera for any of the chimeras, suggesting the addition of these residues has not improved the efficacy for signaling; however, the increase in maximal cAMP activity suggests that the reintroduction of these regions has made a significant contribution to signal activation.

Single-Site Mutations within the N-Terminal Region of the RXFP2 LDLa Module. The increase in the level of cAMP production upon reintroduction of the two regions onto the RXFP2–LB2 chimera suggests that elements that are important to receptor activation are located within these regions. To this end, we dissected these regions with single-site mutations. Within the first loop between Cys-6 and Cys-13 are residues Gln-7, Lys-8, Gly-9, Tyr-10, Phe-11, and Pro-12. Because of the conservation of the equivalent Gly and Phe residues generally within the LDLa module, we did not mutate these residues on the basis of the assumption that they are structurally important. Our initial studies concentrated on Lys-8 and Tyr-10 as we have previously demonstrated these positions to be important for the function of the RXFP1 LDLa module. We therefore produced the mutant receptor K8A–RXFP2 to test the removal of the lysine side chain. We mutated Tyr-10 to methionine to remove the benzene ring only as the mutation to alanine in RXFP1 resulted in a global disruption to the structure of the LDLa module.³² Both mutant receptors demonstrated normal cell surface expression and INSL3 binding (Table 1). However, both receptors demonstrated

lower activity in response to INSL3 [pEC₅₀ values of 10.43 ± 0.08 for K8A, 10.41 ± 0.07 for Y10M, and 11.09 ± 0.09 for RXFP2 (for both $p < 0.001$)] while having unchanged maximal responses (Figure 8). These minor shifts in INSL3 activity suggested some role for Lys-8 and Tyr-10 and hence were further investigated.

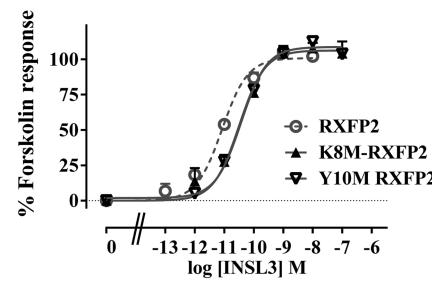


Figure 8. INSL3-induced cAMP response of K8A–RXFP2 and Y10M–RXFP2 compared to that of RXFP2. cAMP activity is expressed as a percentage of the 5 μ M forskolin-stimulated response for each receptor. Symbols represent means, and vertical bars represent the SEM of triplicate determinations from at least three independent experiments.

To further test the potential role of Lys-8, we mutated the residue to methionine to assess the loss of the basic side chain charge. Interestingly, the mutant was expressed on the cell surface at levels ($121.09 \pm 6.8\%$; $p < 0.05$) slightly higher than that of RXFP2; however, there was no change in the pEC₅₀ in response to INSL3 (Table 2). K8M–RXFP2 did have a lower maximal activity (E_{max}) ($80.92 \pm 5.37\%$), which is potentially even lower given the significantly higher level of cell surface expression (Table 2 and Figure 9A). Next, a more conservative

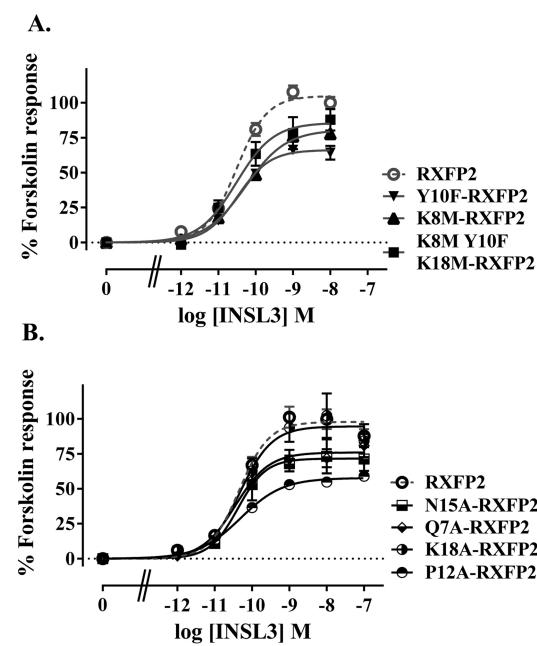


Figure 9. INSL3-induced cAMP responses of (A) K8M–RXFP2, Y10F–RXFP2, and K8M/Y10F/K18M–RXFP2 and (B) Q7A–RXFP2, P12A–RXFP2, N15A–RXFP2, and K18A–RXFP2 compared to that of RXFP2. cAMP activity is expressed as a percentage of the 5 μ M forskolin-stimulated response of RXFP2 for each receptor. Symbols represent means, and vertical bars represent the SEM of triplicate determinations from at least three independent experiments.

substitution of Tyr-10 with phenylalanine to assess the role of the polar OH group was tested. Cell surface expression and INSL3 activity of Y10F-RXFP2 were comparable to those of RXFP2; however, as shown for K8M, it demonstrated a significantly decreased maximal activity in response to INSL3 (67.00 ± 5.04 ; $p < 0.01$) (Table 2 and Figure 9A). The decreases in E_{max} observed with these mutations are more likely to reflect a structural perturbation of the LDLa fold as the pEC₅₀ values remain unchanged. Taken together, these results suggest that Lys-8 and Tyr-10 play a minor role in RXFP2 activation.

We next proceeded to test the other residues in the N-terminal region of the LDLa module potentially implicated in the gain-of-function studies. We mutated the first residue of the QKGYFP region, Gln-7, to alanine, generating Q7A-RXFP2, which was expressed at the cell surface normally and responded to INSL3 with an unaltered pEC₅₀ and a lower, but not significantly different, maximal response ($72.60 \pm 11.37\%$) (Table 2 and Figure 9B). As reintroduction of the section loop NLTK onto the RXFP2–LB2 scaffold produced a modest increase in the level of cAMP, we therefore chose to mutate Asn-15 to test the loss of glycosylation and Lys-18 because of its charge and protrusion from the surface of the molecule. Mutation of Asn-15 to alanine resulted in a large decrease in the level of cell surface expression ($40.99 \pm 4.1\%$; $p < 0.001$). However, there was no effect on the INSL3-stimulated potency but an expected decrease in the maximal activity due to the lower level of cell surface expression ($71.29 \pm 16.63\%$ of RXFP2 activity) (Table 2 and Figure 9B). Hence, this mutation likely disrupts glycosylation at the putative N-glycosylation site but does not effect INSL3-mediated activation. Mutation of Lys-18 to alanine (K18A-RXFP2) had no effect on cell surface expression or INSL3 activity (Table 2 and Figure 9B).

At this point, as none of the single-residue mutations tested in this study appeared to have a dramatic effect on INSL3 activity, we decided to test a combination of residues to potentially amplify any effect. To this end, we produced the triple mutant K8M/Y10F/K18M-RXFP2, which corresponds to the L7K/Y9M/K17A-RXFP1 mutant in which we observed a pronounced loss of activity.³³ Interestingly, although this mutant showed a lower level of cell surface expression ($79.44 \pm 4.44\%$; $p < 0.05$), there was no change in INSL3-mediated cAMP activity, although the maximal activity reflected the lower level of cell surface expression (Table 2 and Figure 9A).

As none of the additional mutants tested above exhibited any significant change in INSL3 potency, binding was not tested on any of the mutants. Taken together, none of the single-side chain mutations nor the triple mutant generated in this study had a major impact on the signal activation of RXFP2 in response to INSL3.

■ DISCUSSION

Activation of GPCRs generally involves direct binding of ligand to the TM domain, inducing a conformational change within the packing of the helices that ultimately leads to G protein coupling and activation of downstream signaling cascades. For the relaxin and INSL3 receptors, this appears to be a complex and multistep process. Both receptors, RXFP1 and RXFP2, are multidomain forms consisting of a GPCR class-A TM domain and a large extracellular region containing a domain of 10 LRRs and a LDLa module. The primary binding of relaxin and INSL3 is mapped to the LRRs²² through well-characterized interactions with residues in the B-chain of the peptides, and

a second interaction with the TM exoloops, potentially loop 2.^{28,53} A suite of chimeric receptors and mutations, primarily of RXFP1, have been characterized to highlight that the single LDLa module at the N-terminus of the receptor is essential for activation to occur through specific side chain interactions but only when the receptor is ligand-bound.^{32,33} For RXFP1, this activation is dependent on specific side chains, particularly within the N-terminal region, of the LDLa module.^{32,33} Previous studies^{22,39,54} have shown that the activation of INSL3-bound RXFP2 is also reliant on the LDLa module; however, the molecular details of this have not been investigated further. It is reasonable to assume, however, that the LDLa module makes contacts with the TM, as chimeric receptors reveal the best signaling is achieved when LDLa modules are paired with their native TM.⁵⁴ In this scenario, even though LDLa-driven activation is dependent on ligand binding, the true ligand that activates the receptor is the LDLa itself.

Initially, we wanted to confirm that removal of the LDLa module from RXFP2 results in a receptor that is unable to signal in response to ligand. We had previously tested an RXFP2 receptor construct lacking the LDLa module, RXFP2-short, and demonstrated that it binds INSL3 normally but did not signal as determined by direct cAMP measurements.³¹ As RXFP2-short is still able to bind ligand, it is possible that it can signal through an alternative GPCR signaling pathway. Therefore, we produced a stable cell line expressing RXFP2-short, confirmed that the receptor bound INSL3 normally, and then transfected a panel of reporter genes known to be associated with GPCR signaling to test if RXFP2-short was able to induce alternative signaling in comparison to that of RXFP2. We previously used this system to study signaling at RXFP2 and demonstrated that INSL3 stimulation of RXFP2 activated the CRE reporter gene but not other signaling pathways.³⁹ Importantly, in this study, we were able to confirm these results for RXFP2 while at the same time demonstrating that the RXFP2-short receptor is unable to signal through any signaling pathway. These results are consistent with those for the RXFP1 receptor in an equivalent experiment³³ and suggest in this context the LDLa module is the true “ligand” that leads to G protein coupling. These studies also gave us the confidence to utilize the CRE reporter gene in this study to monitor the ability of mutant receptors to signal.

From these results, we hypothesized that specific residues within the LDLa module are likely responsible for interactions that drive the active receptor conformation. To this end, we replaced the RXFP2 LDLa module with an unrelated LDLa module from the LDL receptor family, LB2, which has a sequence weakly identical to that of RXFP2 but a similar three-dimensional structure, creating the RXFP2–LB2 chimera. To our surprise, upon INSL3 stimulation the RXFP2–LB2 chimera produced a cAMP response, albeit with a much lower maximal activity (~12% of RXFP2) and with a compromised pEC₅₀. This observation is in contrast to our previous findings for an equivalent RXFP1–LB2 chimera, which was unable to signal in response to relaxin. This result immediately suggests that although the response of the RXFP2–LB2 chimera to INSL3 is much weaker than that of the wild-type receptor, in some way the LB2 module is able to compensate for the native RXFP2 module in a way that it cannot for RXFP1, and given the low level of sequence identity of the modules in the N-terminal region of the LDLa modules, it seems tempting to speculate that residues in the C-terminal

region that ligate the calcium ion are responsible for this activity.

Although the RXFP2–LB2 chimera was able to respond to INSL3 stimulation, we deemed this response weak enough to use LB2 as a scaffold to add-back two regions of the native RXFP2 sequence and look for a gain of function. Through this approach, we found that reintroduction of the residues of the first ordered loop into the N-terminus consisting of residues Gln-7, Lys-8, Gly-9, Tyr-10, Phe-11, and Pro-12 increased the maximal cAMP response (E_{max}) to ~50% while reintroduction of the second loop between the two antiparallel β -strands, Asn-15, Leu-16, Thr-17, and Lys-18, resulted in a modest increase in E_{max} from ~12 to ~21% and combining the two regions resulted in an additive increase in E_{max} of 60%. The incremental increase in cAMP production suggests that both regions independently contribute to receptor activation, and that certainly the majority of this effect is from the residues within the region of Gln-7–Pro-12 as increases attributable to the region of Asn-15–Lys-18 are modest. These observations again contrast with the results from our previous RXFP1 studies in which the second loop yielded a profound increase in E_{max} ,³³ which was attributed to Lys-17 in RXFP1. Importantly, none of the gain-of-function mutants showed any significant change in their pEC₅₀ values compared to that of the RXFP2–LB2 chimera. These results suggest we have improved the ability of the receptor to couple to G proteins, leading to an increased level of cAMP activation; however, the lack of an effect on INSL3 efficacy with no change in ligand affinity suggests that we are not reintroducing residues that increase the potency of the LDLa as the ligand.

We therefore turned to loss-of-function mutations in the LDLa module to potentially determine residues of the N-terminal region that may be involved in the activity of the LDLa module. Somewhat surprisingly, and in contrast to mutations in the N-terminal region of the RXFP1 LDLa, there was little effect of any mutation on the activity of RXFP2. The mutation of Lys-8 and Tyr-10 alone led to minor changes in activity, but when combined into a triple mutant, K8M/Y10F/K18M-RXFP2, equivalent to the triple mutant that gave a marked reduction in the activity of RXFP1, there was no effect on receptor activity. The most pronounced effect was seen with mutant P12A, which reduced the E_{max} to 58% of the wild-type value. The lack of an effect on INSL3 potency suggests that this mutation rather than affecting an interaction has a structural effect on the module, weakening its ability to signal. Such structural perturbations in activity were seen with mutation in the LDLa of the RXFP1 receptor.³² The other mutation, which demonstrates a significant effect, was N15A, which resulted in a large decrease in the level of cell surface expression but had no effect on INSL3 activity. As this residue is a potential site for N-linked glycosylation, this would suggest that glycosylation is important for cell surface trafficking, but not receptor activation, and this is consistent with observations of the equivalent site in RXFP1.^{31,40,55}

Taken together, these results suggest that unlike RXFP1, the residues that are key to the activation surface are not located in the N-terminal region of the module and that the C-terminal region, which also ligates the calcium ion, is important. This is consistent with the observation that the RXFP2–LB2 chimera can generate a ligand-dependent response and that addition of RXFP2 LDLa N-terminal residues resulted in increases in E_{max} but not pEC₅₀. This leaves open the question of why an increase in E_{max} was observed upon the introduction of native

RXFP2 sequence into the RXFP2–LB2 construct. Given that it seems for RXFP2 the C-terminal region is involved in signal activation, it is reasonable to hypothesize that the increase in E_{max} is a reflection of an overall improved structure for the module and subsequently the receptor or an increase in the number of receptors at the cell surface that are sufficiently well structured to generate a response. When we previously correlated E_{max} responses from mutations in the whole receptor with the ability of the recombinant protein to refold, we found these mutations lowered the affinity or the ability of the protein to refold in the presence of calcium, as monitored by HPLC. This loss of structural integrity results in a reduction in E_{max} because fewer receptors are present at the cell surface with functional LDLa modules.³² To this end, it would appear that no change in pEC₅₀ through the mutations we have studied here suggests that we have not found key residues responsible for the potency of the “ligand”, which for this system is the LDLa module itself. The conclusion extends the idea that adding back of the RXFP2 regions to the RXFP2–LB2 chimera improves the ability of the module to remain structured at the cell surface and/or become more “RXFP2-like”. Indeed, when we determined the structure of the SLYFP-NITK–RXFP1–LB2 construct, we found that the structure within the C-terminal region of LB2 moved to become more “RXFP1-like”.³³

The hypothesis that the C-terminal region of the RXFP2 LDLa module is involved in receptor activation is consistent with the binding surface generally employed by LDLa modules for ligand interactions. There are several examples of LDLa modules using the same acidic residues that bind calcium in ligand binding: binding of ApoER2 to reelin,⁵⁶ binding of gentamicin to the 10th complement repeat of megalin,⁵⁷ and binding of the third domain of the receptor-associated protein (RAP) to LDLa modules 3 and 4 of the LDL receptor.⁵⁸ It is important to note that loss-of-function mutations in the N-terminal region of the RXFP1 LDLa module do not lead to a complete loss of activity. Hence, it is possible that some of the activity of the RXFP1 LDLa module resides in the C-terminal region. It is tempting to speculate that the N-terminal region has evolved to play a role in RXFP1 signaling because of the pleiotropic actions of relaxin that are associated with additional G protein-activated pathways (reviewed in ref 11). The distinct receptor conformations associated with this additional G protein coupling in RXFP1 may be driven by the residues in the N-terminal region of the RXFP1 LDLa module.

Unfortunately, it is difficult to probe the role of the C-terminal region of the LDLa module in either RXFP1 or RXFP2 as mutations in this region are not well-tolerated by the module because of the structural role of ligation of the calcium ion. Our challenge now is to identify the site at which the LDLa modules of both RXFP1 and RXFP2 bind as characterization of this site may aid our understanding of the mode of interaction and identify residues within the RXFP2 LDLa module that are important for signaling.

The outcomes of this work are essential to the development of small molecule agonists and/or antagonists based on the structure of the LDLa module. From these studies, it would seem that RXFP1 and RXFP2 have evolved to use their LDLa modules for the same function but through very different side chain interactions, presenting an opportunity to exploit these differences to develop receptor specific small molecules.

ASSOCIATED CONTENT

Supporting Information

Primers used for the generation of receptor constructs (Table 1) and statistics for the refinement of the NMR structure of the RXFP2 LDLa module (Table 2). This material is available free of charge via the Internet at <http://pubs.acs.org>.

Accession Codes

The structure of the RXFP2 LDLa module has been deposited as PDB entry 2M96 and BMRB entry 19282. It should be noted that because of cloning artifacts the numbering of the residues in the PDB is +2 with respect to the residue numbering discussed in this paper.

AUTHOR INFORMATION

Corresponding Authors

*Department of Biochemistry and Molecular Biology, The Bio21 Molecular Science and Biotechnology Institute, University of Melbourne, 30 Flemington Rd., Parkville, Victoria 3010, Australia. Telephone: 61-3-8344-2273. Fax: 61-3-9348-1421. E-mail: prg@unimelb.edu.au.

*Florey Institute of Neuroscience and Mental Health, University of Melbourne, Victoria 3010, Australia. E-mail: bathgate@floreys.edu.au.

Funding

This research was supported by National Health and Medical Research Council (NHMRC) of Australia Project Grants 628427 and 1043750 (R.A.D.B. and P.R.G.) and by the Victorian Government Operational Infrastructure Support Program. Instrument funding by the State of Victoria, ARCLIEF, and the Rowden White Foundation. R.A.D.B. and J.D.W. are recipients of an NHMRC Research Fellowship. R.C.K.K. is a recipient of a University of Melbourne International Research Scholarship and a University of Melbourne Fee Remission Scholarship. E.J.P. is a recipient of a Melbourne Research Fellowship (Career Interruptions).

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank Sharon Layfield and Tania Ferraro for technical assistance.

ABBREVIATIONS

RXFP, relaxin family peptide receptor; INSL3, insulin-like peptide-3; GPCR, G protein-coupled receptor; LDLa, low-density lipoprotein class-A; LRRs, leucine-rich repeats; LGR, leucine-rich repeat containing G protein-coupled receptors; TM, transmembrane region.

REFERENCES

- 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/β-catenin and Wnt/PCP signalling. *EMBO Rep.* **12**, 1055–1061.
- Van Hiel, M. B., Vandersmissen, H. P., Van Loy, T., and Vandenberghe, J. (2012) An evolutionary comparison of leucine-rich repeat containing G protein-coupled receptors reveals a novel LGR subtype. *Peptides* **34**, 193–200.
- Hsu, S. Y., Nakabayashi, K., Nishi, S., Kumagai, J., Kudo, M., Sherwood, O. D., and Hsueh, A. J. (2002) Activation of orphan receptors by the hormone relaxin. *Science* **295**, 671–674.
- Yamamoto, T., Davis, C. G., Brown, M. S., Schneider, W. J., Casey, M. L., Goldstein, J. L., and Russell, D. W. (1984) The human LDL receptor: A cysteine-rich protein with multiple Alu sequences in its mRNA. *Cell* **39**, 27–38.
- Sudhof, T. C., Goldstein, J. L., Brown, M. S., and Russell, D. W. (1985) The LDL receptor gene: A mosaic of exons shared with different proteins. *Science* **238**, 815–822.
- Bates, P., Young, J. A., and Varmus, H. E. (1993) A receptor for subgroup A Rous sarcoma virus is related to the low density lipoprotein receptor. *Cell* **74**, 1043–1051.
- Takeuchi, T., Misaki, A., Chen, B. K., and Ohtsuki, Y. (1999) H-cadherin expression in breast cancer. *Histopathology* **35**, 87–88.
- Morris, S. W., Naeve, C., Mathew, P., James, P. L., Kirstein, M. N., Cui, X., and Witte, D. P. (1997) ALK, the chromosome 2 gene locus altered by the t(2;5) in non-Hodgkin's lymphoma, encodes a novel neural receptor tyrosine kinase that is highly related to leukocyte tyrosine kinase (LTK). *Oncogene* **14**, 2175–2188.
- Demczuk, S., Aledo, R., Zucman, J., Delattre, O., Desmaze, C., Dauphinot, L., Jalbert, P., Rouleau, G. A., Thomas, G., and Aurias, A. (1995) Cloning of a balanced translocation breakpoint in the DiGeorge syndrome critical region and isolation of a novel potential adhesion receptor gene in its vicinity. *Hum. Mol. Genet.* **4**, 551–558.
- Kallunki, P., and Tryggvason, K. (1992) Human basement membrane heparan sulfate proteoglycan core protein: A 467-kD protein containing multiple domains resembling elements of the low density lipoprotein receptor, laminin, neural cell adhesion molecules, and epidermal growth factor. *J. Cell Biol.* **116**, 559–571.
- Bathgate, R. A., Halls, M. L., van der Westhuizen, E. T., Callander, G. E., Kocan, M., and Summers, R. J. (2013) Relaxin family peptides and their receptors. *Physiol. Rev.* **93**, 405–480.
- Conrad, K. P. (2010) Unveiling the vasodilatory actions and mechanisms of relaxin. *Hypertension* **56**, 2–9.
- Unemori, E. N., and Amento, E. P. (1990) Relaxin modulates synthesis and secretion of procollagenase and collagen by human dermal fibroblasts. *J. Biol. Chem.* **265**, 10681–10685.
- Bay, K., Main, K. M., Toppari, J., and Skakkebaek, N. E. (2011) Testicular descent: INSL3, testosterone, genes and the intrauterine milieu. *Nat. Rev. Urol.* **8**, 187–196.
- Ivell, R., Heng, K., and Anand-Ivell, R. (2014) Insulin-Like Factor 3 and the HPG Axis in the Male. *Front. Endocrinol.* **5**, 6.
- Kawamura, K., Kumagai, J., Sudo, S., Chun, S. Y., Pisarska, M., Morita, H., Toppari, J., Fu, P., Wade, J. D., Bathgate, R. A., and Hsueh, A. J. (2004) Paracrine regulation of mammalian oocyte maturation and male germ cell survival. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 7323–7328.
- Ferlin, A., Selice, R., Carraro, U., and Foresta, C. (2013) Testicular function and bone metabolism: Beyond testosterone. *Nat. Rev. Endocrinol.* **9**, 548–554.
- Johnston, S. E., Gratten, J., Berenos, C., Pilkington, J. G., Clutton-Brock, T. H., Pemberton, J. M., and Slate, J. (2013) Life history trade-offs at a single locus maintain sexually selected genetic variation. *Nature* **502**, 93–95.
- Wiedemar, N., Tetens, J., Jagannathan, V., Menoud, A., Neuenschwander, S., Bruggmann, R., Thaller, G., and Drogemuller, C. (2014) Independent polled mutations leading to complex gene expression differences in cattle. *PLoS One* **9**, e93435.
- Glister, C., Satchell, L., Bathgate, R. A., Wade, J. D., Dai, Y., Ivell, R., Anand-Ivell, R., Rodgers, R. J., and Knight, P. G. (2013) Functional link between bone morphogenetic proteins and insulin-like peptide 3 signaling in modulating ovarian androgen production. *Proc. Natl. Acad. Sci. U.S.A.* **110**, E1426–E1435.
- Xue, K., Kim, J. Y., Liu, J. Y., and Tsang, B. K. (2014) Insulin-like 3-induced rat preantral follicular growth is mediated by growth differentiation factor 9. *Endocrinology* **155**, 156–167.
- Scott, D. J., Rosengren, K. J., and Bathgate, R. A. (2012) The different ligand-binding modes of relaxin family peptide receptors RXFP1 and RXFP2. *Mol. Endocrinol.* **26**, 1896–1906.
- Scott, D. J., Wilkinson, T. N., Zhang, S., Ferraro, T., Wade, J. D., Tregear, G. W., and Bathgate, R. A. (2007) Defining the LGR8 residues involved in binding insulin-like peptide 3. *Mol. Endocrinol.* **21**, 1699–1712.

- (24) Bullesbach, E. E., and Schwabe, C. (2005) The trap-like relaxin-binding site of the leucine-rich G-protein-coupled receptor 7. *J. Biol. Chem.* 280, 14051–14056.
- (25) Bullesbach, E. E., and Schwabe, C. (2000) The relaxin receptor-binding site geometry suggests a novel gripping mode of interaction. *J. Biol. Chem.* 275, 35276–35280.
- (26) Hossain, M. A., Rosengren, K. J., Haugaard-Jonsson, L. M., Zhang, S., Layfield, S., Ferraro, T., Daly, N. L., Tregear, G. W., Wade, J. D., and Bathgate, R. A. (2008) The A-chain of human relaxin family peptides has distinct roles in the binding and activation of the different relaxin family peptide receptors. *J. Biol. Chem.* 283, 17287–17297.
- (27) Bathgate, R. A., Zhang, S., Hughes, R. A., Rosengren, K. J., and Wade, J. D. (2012) The structural determinants of insulin-like Peptide 3 activity. *Front. Endocrinol.* 3, 11.
- (28) Sudo, S., Kumagai, J., Nishi, S., Layfield, S., Ferraro, T., Bathgate, R. A., and Hsueh, A. J. (2003) H3 relaxin is a specific ligand for LGR7 and activates the receptor by interacting with both the ectodomain and the exoloop 2. *J. Biol. Chem.* 278, 7855–7862.
- (29) Kong, R. C., Shilling, P. J., Lobb, D. K., Gooley, P. R., and Bathgate, R. A. (2010) Membrane receptors: Structure and function of the relaxin family peptide receptors. *Mol. Cell. Endocrinol.* 320, 1–15.
- (30) Deupi, X., Standfuss, J., and Schertler, G. (2012) Conserved activation pathways in G-protein-coupled receptors. *Biochem. Soc. Trans.* 40, 383–388.
- (31) Scott, D. J., Layfield, S., Yan, Y., Sudo, S., Hsueh, A. J., Tregear, G. W., and Bathgate, R. A. (2006) Characterization of novel splice variants of LGR7 and LGR8 reveals that receptor signaling is mediated by their unique low density lipoprotein class A modules. *J. Biol. Chem.* 281, 34942–34954.
- (32) Hopkins, E. J., Layfield, S., Ferraro, T., Bathgate, R. A., and Gooley, P. R. (2007) The NMR solution structure of the relaxin (RXFP1) receptor lipoprotein receptor class A module and identification of key residues in the N-terminal region of the module that mediate receptor activation. *J. Biol. Chem.* 282, 4172–4184.
- (33) Kong, R. C., Petrie, E. J., Mohanty, B., Ling, J., Lee, J. C., Gooley, P. R., and Bathgate, R. A. (2013) The relaxin receptor (RXFP1) utilizes hydrophobic moieties on a signaling surface of its N-terminal low density lipoprotein class A module to mediate receptor activation. *J. Biol. Chem.* 288, 28138–28151.
- (34) Feng, S., and Agoulnik, A. I. (2011) Expression of LDL-A module of relaxin receptor in prostate cancer cells inhibits tumorigenesis. *Int. J. Oncol.* 39, 1559–1565.
- (35) Spanel-Borowski, K., Schafer, I., Zimmermann, S., Engel, W., and Adham, I. M. (2001) Increase in final stages of follicular atresia and premature decay of corpora lutea in *Insl3*-deficient mice. *Mol. Reprod. Dev.* 58, 281–286.
- (36) Del Borgo, M. P., Hughes, R. A., Bathgate, R. A., Lin, F., Kawamura, K., and Wade, J. D. (2006) Analogs of insulin-like peptide 3 (INSL3) B-chain are LGR8 antagonists in vitro and in vivo. *J. Biol. Chem.* 281, 13068–13074.
- (37) Scott, D. J., Tregear, G. W., and Bathgate, R. A. (2005) LGR7-truncate is a splice variant of the relaxin receptor LGR7 and is a relaxin antagonist in vitro. *Ann. N.Y. Acad. Sci.* 1041, 22–26.
- (38) Cepko, C., and Pear, W. (2001) Overview of the retrovirus transduction system. *Current Protocols in Molecular Biology*, Chapter 9, Unit 9, 9, Wiley, New York.
- (39) Halls, M. L., Bathgate, R. A., and Summers, R. J. (2007) Comparison of signaling pathways activated by the relaxin family peptide receptors, RXFP1 and RXFP2, using reporter genes. *J. Pharmacol. Exp. Ther.* 320, 281–290.
- (40) Yan, Y., Scott, D. J., Wilkinson, T. N., Ji, J., Tregear, G. W., and Bathgate, R. A. (2008) Identification of the N-linked glycosylation sites of the human relaxin receptor and effect of glycosylation on receptor function. *Biochemistry* 47, 6953–6968.
- (41) Rosengren, K. J., Zhang, S., Lin, F., Daly, N. L., Scott, D. J., Hughes, R. A., Bathgate, R. A., Craik, D. J., and Wade, J. D. (2006) Solution structure and characterization of the LGR8 receptor binding surface of insulin-like peptide 3. *J. Biol. Chem.* 281, 28287–28295.
- (42) Shabanpoor, F., Hughes, R. A., Zhang, S., Bathgate, R. A., Layfield, S., Hossain, M. A., Tregear, G. W., Separovic, F., and Wade, J. D. (2010) Effect of helix-promoting strategies on the biological activity of novel analogues of the B-chain of INSL3. *Amino Acids* 38, 121–131.
- (43) Shabanpoor, F., Hughes, R. A., Bathgate, R. A., Zhang, S., Scanlon, D. B., Lin, F., Hossain, M. A., Separovic, F., and Wade, J. D. (2008) Solid-phase synthesis of europium-labeled human INSL3 as a novel probe for the study of ligand-receptor interactions. *Bioconjugate Chem.* 19, 1456–1463.
- (44) Cai, M., Huang, Y., Sakaguchi, K., Clore, G. M., Gronenborn, A. M., and Craigie, R. (1998) An efficient and cost-effective isotope labeling protocol for proteins expressed in *Escherichia coli*. *J. Biomol. NMR* 11, 97–102.
- (45) Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* 6, 277–293.
- (46) Vranken, W. F., Boucher, W., Stevens, T. J., Fogh, R. H., Pajon, A., Llinas, M., Ulrich, E. L., Markley, J. L., Ionides, J., and Lue, E. D. (2005) The CCPN data model for NMR spectroscopy: Development of a software pipeline. *Proteins* 59, 687–696.
- (47) Ellgaard, L., Bettendorff, P., Braun, D., Herrmann, T., Fiorito, F., Jelesarov, I., Guntert, P., Helenius, A., and Wuthrich, K. (2002) NMR structures of 36 and 73-residue fragments of the calreticulin P-domain. *J. Mol. Biol.* 322, 773–784.
- (48) Guntert, P. (2004) Automated NMR structure calculation with CYANA. *Methods Mol. Biol.* 278, 353–378.
- (49) Shen, Y., Delaglio, F., Cornilescu, G., and Bax, A. (2009) TALOS+: A hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J. Biomol. NMR* 44, 213–223.
- (50) Fass, D., Blacklow, S., Kim, P. S., and Berger, J. M. (1997) Molecular basis of familial hypercholesterolemia from structure of LDL receptor module. *Nature* 388, 691–693.
- (51) Laskowski, R. A., Rullmann, J. A., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) AQUA and PROCHECK-NMR: Programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR* 8, 477–486.
- (52) Koradi, R., Billeter, M., and Wuthrich, K. (1996) MOLMOL: A program for display and analysis of macromolecular structures. *J. Mol. Graphics* 14, 29–32, 51–55.
- (53) Halls, M. L., Bond, C. P., Sudo, S., Kumagai, J., Ferraro, T., Layfield, S., Bathgate, R. A., and Summers, R. J. (2005) Multiple binding sites revealed by interaction of relaxin family peptides with native and chimeric relaxin family peptide receptors 1 and 2 (LGR7 and LGR8). *J. Pharmacol. Exp. Ther.* 313, 677–687.
- (54) Bruell, S., Kong, R. C., Petrie, E. J., Hoare, B., Wade, J. D., Scott, D. J., Gooley, P. R., and Bathgate, R. A. (2013) Chimeric RXFP1 and RXFP2 Receptors Highlight the Similar Mechanism of Activation Utilizing Their N-Terminal Low-Density Lipoprotein Class A Modules. *Front. Endocrinol.* 4, 171.
- (55) Kern, A., Agoulnik, A. I., and Bryant-Greenwood, G. D. (2007) The low-density lipoprotein class A module of the relaxin receptor (leucine-rich repeat containing G-protein coupled receptor 7): Its role in signaling and trafficking to the cell membrane. *Endocrinology* 148, 1181–1194.
- (56) Yasui, N., Nogi, T., and Takagi, J. (2010) Structural basis for specific recognition of reelin by its receptors. *Structure* 18, 320–331.
- (57) Dagil, R., O'Shea, C., Nykjaer, A., Bonvin, A. M., and Kragelund, B. B. (2013) Gentamicin binds to the megalin receptor as a competitive inhibitor using the common ligand binding motif of complement type repeats: Insight from the NMR structure of the 10th complement type repeat domain alone and in complex with gentamicin. *J. Biol. Chem.* 288, 4424–4435.
- (58) Fisher, C., Beglova, N., and Blacklow, S. C. (2006) Structure of an LDLR-RAP complex reveals a general mode for ligand recognition by lipoprotein receptors. *Mol. Cell* 22, 277–283.