The relaxin receptor RXFP1 signals through a mechanism of autoinhibition

Authors: Sarah C. Erlandson¹, Shaun Rawson¹, Kelly P. Brock², Xinyue Liu¹, Joao A. Paulo¹, Julian Mintseris¹, Steven P. Gygi¹, Debora S. Marks², Xiaojing Cong³, Andrew C. Kruse^{1,*}

Affiliations:

¹Department of Biological Chemistry and Molecular Pharmacology, Blavatnik Institute, Harvard Medical School, Boston, Massachusetts 02115, USA.

²Department of Systems Biology, Blavatnik Institute, Harvard Medical School, Boston, Massachusetts 02115, USA.

³Institut de Génomique Fonctionnelle, CNRS UMR5203-INSERM U1191-Université de Montpellier, 34090 Montpellier Cedex 5, France

*Correspondence and requests for materials should be addressed to
andrew_kruse@hms.harvard.edu.

Abstract

The relaxin family peptide receptor 1 (RXFP1) is the receptor for relaxin-2, an important regulator of reproductive and cardiovascular physiology and an established therapeutic target. RXFP1 is a multi-domain G protein-coupled receptor (GPCR) with an ectodomain consisting of an LDLa module and leucine-rich repeats. The molecular mechanism of RXFP1 signal transduction is clearly distinct from that of other GPCRs, but remains very poorly understood, hindering the development of therapeutics targeting the receptor. Here we present the cryo-electron microscopy structure of active-state human RXFP1, bound to the endogenous agonist relaxin-2 and heterotrimeric G_s proteins. Evolutionary coupling analysis and structure-guided functional experiments reveal that RXFP1 signals through a mechanism of autoinhibition, wherein the receptor's extracellular loop 2 occupies the orthosteric site in the active state but is inhibited by the ectodomain in the absence of relaxin-2. Together, these data provide a model for relaxin receptor signaling, explaining how an unusual but important GPCR sub-family functions and providing a path to rational drug development targeting the relaxin receptors.

Introduction

RXFP1 is a member of the leucine-rich repeat-containing GPCRs (LGR), a small subset of family A GPCRs that have remained a notable exception to our understanding of GPCR signaling, despite substantial progress in studies of other GPCR families. In LGRs, the leucine-rich repeats (LRRs) function as the extracellular ligand-binding domain for three types of protein agonists: glycoprotein hormones, R-spondins, and relaxins¹. The LGRs are involved in a variety of physiological processes across reproductive and developmental biology. RXFP1, the receptor for the relaxin-2 hormone², plays an important role during pregnancy. In this setting, it is responsible for physiological changes including increasing cardiac output and remodeling reproductive tissues to facilitate parturition³⁻⁵. RXFP1 signaling also regulates the physiology of numerous organs in both males and females, particularly the heart, lungs, liver, and kidneys. Activation of RXFP1 by relaxin-2 in these organs leads to pleiotropic cellular effects, including vasodilation, angiogenesis, anti-inflammatory responses, and extracellular matrix remodeling through collagen degradation^{6,7}. Accordingly, the RXFP1 receptor has emerged as a promising therapeutic target for the treatment of numerous cardiovascular and fibrotic diseases⁸⁻¹⁰.

The relaxin receptors RXFP1 and RXFP2 are unique members of the LGR family, and are classified as type C LGRs due to an additional domain called an LDLa module at the receptors' distal N-termini, before the LRRs in sequence^{11,12}. These receptors are the only two mammalian GPCRs to contain an LDLa module, and the role of this domain in relaxin receptor signaling is poorly understood. The LDLa module is dispensable for relaxin-2 binding to the LRRs but essential for activation of RXFP1 signaling in response to relaxin-2¹³. The mechanisms that couple ligand binding in the LRRs to conformational changes within the 7TM domain required for G protein signaling remain undefined, largely due to an absence of structural data. Recent structures of the luteinizing hormone-choriogonadotropin receptor (LHCGR), one of the glycoprotein hormone receptors or type A LGRs, revealed conformational changes of the LRRs between inactive and active states. These studies proposed that large glycoprotein hormones signal through a steric “push-pull” mechanism that activates the receptor by driving changes in LRR conformation¹⁴. However, the small size (6 kDa) of the relaxin-2 peptide precludes such a mechanism, requiring an alternative explanation.

In order to elucidate the basis for RXFP1 signal transduction, we set out to determine the active-state structure of human RXFP1 bound to relaxin-2 and the heterotrimeric G protein G_s.

We optimized receptor expression using fusions to an engineered minimal G_{α} protein¹⁵, then formed a larger complex with the addition of G protein β_1 and γ_2 subunits, and determined the structure using cryo-electron microscopy. Unexpectedly, the structure revealed that RXFP1's extracellular loop 2 (ECL2) occupies the GPCR orthosteric ligand-binding pocket in the active state. Results from structural and functional studies define a mechanism in which ECL2 conformation is regulated by the receptor's LRRs and hinge region, a short segment between the LRRs and 7TMs. Collectively, these studies identify several conformational switches in both the receptor ectodomain and 7TMs, showing that the concerted action of multiple receptor domains controls the transduction of RXFP1 signaling by its agonist, relaxin-2.

Cryo-EM structure of the RXFP1–G protein complex

Wild type (WT) full-length RXFP1 receptor could be expressed only at very low levels in mammalian cells. To enable structural studies, we cloned a fusion of RXFP1 to the engineered G_{α} protein minimal G_s (mini- G_s)¹⁵. Truncations of RXFP1's flexible C-terminus further increased expression levels, with the optimal expression construct having a C-terminal truncation of 20 amino acids (**Fig. S1**). The binding of the catalytically inactive mini- G_s protein blocks RXFP1 signaling through endogenous G_s proteins and likely stabilizes the receptor, leading to higher purification yields. The fusion protein of RXFP1 and mini- G_s was purified in complex with human G protein β_1 and γ_2 subunits, the camelid antibody VHH fragment nanobody 35 (Nb35)¹⁶, and an engineered version of the agonist relaxin-2 (SE001; see Methods) to form an agonist–GPCR–G protein complex (hereafter referred to as RXFP1– G_s).

Cryo-electron microscopy (cryo-EM) was used for structural studies of RXFP1– G_s . Initial two-dimensional classification analysis revealed averages that showed clear density for RXFP1's 7TM domain and the heterotrimeric G protein. In contrast, the density for RXFP1's ectodomain was weak and poorly defined, indicating flexibility of the ectodomain with respect to the transmembrane domain. Due to the conformational heterogeneity of RXFP1– G_s , we analyzed the cryo-EM data using two different approaches. The first utilized masking of RXFP1's 7TM and G proteins to obtain a high-resolution cryo-EM map of these domains at 3.2 Å, allowing us to build an atomic model (**Fig. S2**). The second approach used focused classifications of RXFP1's ectodomain to obtain a cryo-EM map of the entire complex (**Fig. S3**). As a result of the

ectodomain's flexibility, the cryo-EM map of the full-length receptor is lower resolution, with an overall resolution of 4.2 Å and local resolution of the ectodomain between 5-8 Å.

RXFP1's 7TM domain displays characteristic hallmarks of the active state for class A GPCRs (**Fig. 1**). Most notable is the outward conformation of the intracellular end of transmembrane helix 6 (TM6), which creates the binding site for the $\alpha 5$ helix of $G_{\alpha s}$ ^{17,18}. Additional active-state features include the hydrogen bond between Tyr681^{7.53} of the conserved NPxxY motif in TM7 and Tyr599^{5.58} in TM5, and an open conformation of the “ionic lock” between Glu623^{6.30} and Lys510^{3.50;18} (superscript indicates Ballesteros-Weinstein numbering system)¹⁹. In RXFP1- G_s , the active state of the 7TMs displays a canonical interaction with heterotrimeric G_s , similar to previously reported GPCR-G protein structures²⁰.

Fig. 1 | Cryo-EM map and model of the RXFP1-G_s complex. a, Diagram of the primary structure of RXFP1 domains. **b-c**, Cryo-EM map (**b**) and model (**c**) of the RXFP1-G_s complex 7TM domain with heterotrimeric G_s proteins and Nb35.

Extracellular loop 2 is essential for signaling

The active-state structure of RXFP1 surprisingly revealed that ECL2 occupies the GPCR orthosteric ligand-binding pocket (**Fig. 2a**). The conformation of ECL2 was heterogeneous, and three-dimensional focused classifications of RXFP1's extracellular loops were required to obtain a cryo-EM map suitable for model building. The structure of ECL2 can be described in three segments. The first segment of ECL2, from Lys550^{ECL2} to Gly558^{ECL2}, interacts with TMs 3, 4, and 5, with the residues Tyr556^{ECL2} and Tyr557^{ECL2} making the most extensive contacts within the 7TMs. The second segment of ECL2, beginning at Thr559^{ECL2} until His567^{ECL2}, forms a loop structure which binds into the canonical GPCR orthosteric binding site. In particular, the side chains of the residues Phe564^{ECL2} and Leu566^{ECL2} fit into a hydrophobic cavity created by TMs 2, 3, 5, 6, and 7 (**Fig. 2b**). Within this segment, Cys563^{ECL2} forms a disulfide bond with Cys485^{3,25} in TM3, a highly conserved feature among family A GPCRs which stabilizes the conformation of the loop²¹. As a result of the disulfide bond and interactions with the hydrophobic cavity, the second segment is the most well-resolved region of ECL2. In contrast, four residues in the third segment of ECL2, from Ser568^{ECL2} to Ser573^{ECL2}, are not visible in the cryo-EM map, likely indicating a region of higher intrinsic flexibility.

The unusual conformation of ECL2 and deeply buried positions of Phe564^{ECL2} and Leu566^{ECL2} suggested that they may mimic the role played by exogenous agonists in other receptors. Indeed, the location of the residues Phe564^{ECL2} and Leu566^{ECL2} corresponds with the binding sites of small molecule and peptide orthosteric agonists of other family A GPCRs, such as adrenaline binding to the β_2 adrenergic receptor²² and angiotensin II analogs to the angiotensin II type I receptor²³ (**Fig. S4b,c**). Based on these observations, residues from the second segment of ECL2 were tested by mutagenesis for their contribution to signaling. ECL2 substitutions maintained above 50% of wild type RXFP1 expression, excepting the mutation of Leu566^{ECL2} to Asp, which expressed at 33% of wild-type levels (**Fig. S5a**). Mutation of Phe564^{ECL2} to Ala or Leu566^{ECL2} to Asp almost completely ablated RXFP1 signaling in response to relaxin-2, confirming the importance of these residues to receptor activation (**Fig. 2c**).

The discovery of the importance of ECL2 to RXFP1 signaling is reminiscent of the orphan receptor GPR52, which is directly activated by its own ECL2 as a tethered agonist²⁴. While ECL2 shows no conservation of sequence or detailed structure between RXFP1 and GPR52, the binding sites for ECL2 are in analogous positions within the 7TM domain for both

GPCRs (**Fig. S4a**). These structural parallels, along with RXPF1 mutagenesis in cell signaling assays, are consistent with ECL2 serving a critical role in activating RXFP1.

Fig. 2 | Regulation of receptor signaling by ECL2 and the ectodomain. **a**, The conformation of ECL2 and the hinge region in active-state RXFP1. **b**, Details of ECL2 in the 7TM orthosteric site and interactions between ECL2 and the hinge region. **c-d**, The effect of ECL2 (**c**) and Leu402 and Leu403 hinge region (**d**) mutations in an assay for G_s signaling by RXFP1. Data are mean $\pm$ s.e.m. from technical triplicates. **e-f**, Basal signaling and signaling in response to 50 nM relaxin-2 for RXFP1 ectodomain truncation constructs (**e**) and Ile396 and Ser397 hinge region mutations (**f**) in assay for G_s signaling. Data are mean $\pm$ s.e.m. from nine technical replicates.

The role of the hinge region in receptor activation

Our structure of the 7TM domain also includes 6 residues from the hinge region of RXFP1's ectodomain. The hinge region is composed of residues between the end of the LRRs and the beginning of TM1, and the six residues adjacent to TM1 are included in our refined model. We observed that the hinge region curves into the top of the 7TM domain to interact with ECL2 bound in the orthosteric pocket. In particular, the residues Leu402 and Leu403 interact with Val562^{ECL2}, which also packs against TM7 (**Fig. 2b**).

Functional analyses of type A LGRs, the glycoprotein hormone receptors (GPHRs), have previously established that the hinge region of those LGRs is critical for receptor signaling^{25,26}. The unstructured loop in the hinge region of GPHRs is approximately 60-125 residues long and plays a role in binding the receptors' glycoprotein hormone agonists²⁷. For GPHRs, a 10-residue section of their hinge region near the 7TM domain, P10, is also critical for receptor signaling^{25,26} and has been recently shown to adopt different conformations in inactive and active receptor states of LHCGR¹⁴. In contrast, the hinge loop of type C LGRs, such as RXFP1, is predicted to be about 15 residues long, has no established function, and shows no sequence conservation in comparison to type A LGRs¹. Despite the clear differences between the hinge regions of GPHRs and RXFP1, the interactions between Leu402, Leu403, and ECL2 in RXFP1's active-state structure unexpectedly suggested that these residues may also play a role in signaling.

To investigate the role of Leu402 and Leu403, we cloned single mutations to Ala and constructed a double mutant with Ala substitutions for both residues. These constructs expressed at low levels, so wild type receptor expression was reduced in order to compare receptor signaling (**Fig. S5b**). While the single Ala mutants each decreased the efficacy of relaxin-2, the double Ala mutant completely ablated RXFP1 signaling, despite maintaining expression levels at roughly 50% of wild type (**Fig. 2d**). These results indicated that Leu402 and Leu403 of the hinge region are essential for RXFP1 activation, likely functioning to stabilize ECL2 into its active-state conformation in the orthosteric site.

Autoinhibition of RXFP1 signaling

While structures of RXFP1 and GPR52 revealed that ECL2 is critical for the activation of both receptors, other aspects of their signaling suggest differing mechanisms. GPR52 is a self-activated orphan GPCR with very high basal activity, signaling at 90% of its $E_{\max}$ without any

agonist bound^{24,28}. The mechanism governing the intrinsic activity of GPR52 is clear, as the ECL2 tethered agonist is a component of the 7TM structure itself. In contrast, RXFP1 does not have high basal activity, but signals in response to the binding of relaxin-2 to its LRRs. These differences suggest that while both GPCR structures show ECL2 binding in the orthosteric site, RXFP1 likely uses additional mechanisms to prevent continuous self-activation of the 7TM domain. The major structural difference between these two receptors is that GPR52 is a conventional family A GPCR with an unstructured N-terminus, while RXFP1 is a type C LGR, containing a structured ectodomain of LRRs and an LDLa module. For this reason, we set out to investigate the role of RXFP1's ectodomain in modulating the activity of ECL2.

To address this question, we first cloned constructs of RXFP1 with deletions of the receptor's ectodomain. Basal signaling of RXFP1 was not increased by deletion of the LDLa module, deletion of the LDLa module and the 32-residue linker that connects it to the LRRs, or replacement of the linker with a 32 residue Gly-Gly-Ser linker. These results indicated that the LDLa module and linker of RXFP1's ectodomain do not play an inhibitory role in signaling. In contrast, these constructs are unable to signal in response to relaxin-2 binding, consistent with previous studies showing that the LDLa module and linker region are essential for receptor activation^{13,29}. We next focused on testing the role of RXFP1's LRRs in the receptor signaling mechanism. To remove the ectodomain, including the LRRs, constructs were designed to express RXFP1's 7TM domain alone. However, these constructs showed very low cell-surface expression, at 12% of wild type. Although the 7TM domain expressed very poorly, it showed similar basal signaling to wild type RXFP1 in a G_s signaling assay, at 11% of the wild type relaxin-2 E_{max} . To increase expression of the 7TMs, we cloned a fusion of RXFP1's 7TM domain to the unstructured N-terminus of the high-expressing β_2 adrenergic receptor. This fusion rescued 7TM domain expression to essentially wild type levels (**Fig. S5c**). When tested in a G_s signaling assay, the fusion showed a high level of basal activity. In the absence of ligand, the 7TM domain signaled at 70% of the maximum level of agonist-induced signaling for wild-type RXFP1 (**Fig. 2e**). As expected, RXFP1's 7TM domain alone does not show any change in activity in response to relaxin-2, since the relaxin-2 binding site in the LRRs is deleted in this construct³⁰.

To establish whether the high basal activity of RXFP1's 7TM alone is due to ECL2, we introduced the Phe564^{ECL2} to Ala mutation that greatly reduced full-length RXFP1 signaling in

response to relaxin-2. The Phe564^{ECL2} to Ala mutation was able to completely ablate the 7TM domain's high basal activity, confirming that ECL2 constitutively activates the 7TMs in the absence of RXFP1's LRR domain. Likewise, Ala mutations of the hinge region residues Leu402 and Leu403 (which are included in the 7TM domain fusion construct) were also able to reduce the high basal signaling. These data suggested a model of signaling in which the LRRs play an autoinhibitory role in regulating the active state of ECL2 and the hinge region. As a result, deletions of the LRRs allow constitutive activation the receptor, leading to high basal signaling.

Mechanism of RXFP1 7TM autoactivation by ECL2

We performed enhanced-sampling molecular dynamics (MD) simulations (see Methods for details) to study the role of ECL2 in the basal activity of RXFP1's 7TM domain. An inactive-state model was first obtained by MD simulations of RXFP1 7TM deactivation, starting from a model based on the cryo-EM structure. In brief, a sodium ion was placed in the sodium-binding site to favor the sampling of inactive states during the simulations. We obtained an inactive state that was strikingly similar to the inactive-state AlphaFold2 model³¹, in which the intracellular end of TM6 bent toward TM3 to form the classic Lys510^{3,50}–Glu623^{6,30} ionic lock (**Fig. S6a-c**). The second segment of ECL2 maintained the same conformation throughout the simulations, whereas Ser568^{ECL2} to Ser573^{ECL2} in the third segment were highly mobile, consistent with absence of clear density in the cryo-EM map. Starting from the inactive state, we simulated autoactivation of the RXFP1 7TMs by removing the sodium ion and protonating the sodium anchor, Asp451^{2,50}. As negative controls, we performed the same simulations with two mutants with low basal activity, F564^{ECL2}A and L566^{ECL2}D. The wild type RXFP1 7TM exhibited autoactivation with outward movements of TM6 on the intracellular side, destabilization of the ionic lock, and frequent side-chain flips of the toggle switch W641^{6,48} (**Fig. S6d-f**). In contrast, the two mutants remained in inactive conformations. The wild type receptor exhibited distinct shapes of the orthosteric pocket compared to the two mutants, owing to ECL2–TM7 interactions. Namely, Phe564^{ECL2} in the wild type 7TM domain “pushed” TM7 toward TM2, which likely altered W641^{6,48} conformations and triggered autoactivation (**Fig. S6d-f**). The F564^{ECL2}A mutation directly eliminates this steric effect, whereas substituting Leu566^{ECL2} with Asp reorients the charged side chain away from the pocket, leaving space for F564^{ECL2} and diminishing its impact on TM7 conformation (**Fig. S6d,e**).

Inhibitory interactions revealed by evolutionary coupling analysis

An additional insight into the regulation of ECL2 by the ectodomain arose from evolutionary coupling (EC) analysis of RXFP1^{32,33}. The strongest ECs, or evolutionary coupled residues, are derived from applying a global probability model to multiple sequence alignments and typically indicate residue pairs that are in contact in 3D³⁴, including residues involved in conformational changes^{35,36}. Residues of ECL2 had multiple strong ECs pairing them with residues within the 7TM helices, supporting our active-state structure. However, ECL2 also showed ECs with residues from RXFP1's hinge region not present in our model (**Fig. S7**). In total, two residues from the hinge region, Ile396 and Ser397, had ECs with three residues of ECL2, including the critical Phe564^{ECL2}. ECL2 and the hinge region are likely involved in other interactions than those observed in our structure of the active state. If ECL2 contacts Ile396 and Ser397 in an inactive-state conformation, we predicted that those interactions may contribute to the inhibition of ECL2 in the absence of relaxin-2 binding to the LRRs.

Ile396 and Ser397 were each mutated to Ala to test the effects on the activation state of ECL2. Despite having low expression levels, the Ile396 and Ser397 mutations each showed a significant increase in basal signaling, at 21% and 55%, respectively, of wild type RXFP1's $E_{\max}$ (**Fig. 2f, S5d**). The Ser397 to Ala mutant similarly increased the basal signaling of RXFP1 with a deletion of the LDLa module and linker, confirming that those domains are not involved in the mechanism of signaling inhibition. Interestingly, the hinge mutants showed a reduced signaling response to relaxin-2, suggesting that these residues involved in a potential inactive-state ECL2 interface are also important for allosteric communication between the ectodomain and 7TM domain. Addition of the Phe564^{ECL2} or Leu402 and Leu403 substitutions to the Ile396 and Ser397 mutants was able to abolish the increase in basal signaling, confirming that RXFP1 activation is dependent on these residues (**Fig. 2f**).

Relaxin-2 binding to the leucine-rich repeats

The cryo-EM map of the full-length RXFP1-G_s complex is limited by lower resolution due to the dynamic nature of the receptor's ectodomain. Continuous heterogeneity present in the final particle stack for RXFP1-G_s was visualized through three-dimensional variability analysis in cryoSPARC³⁷. The resulting movies showed that the LRRs are very flexible in the relaxin-2-bound state, moving at the hinge region between the LRRs and 7TM domain. Despite these

limitations, the cryo-EM map offered several new insights into the overall ectodomain architecture and relaxin-2 binding.

In the refined cryo-EM map of active-state RXFP1, the LRRs are extended away from the transmembrane domain, at an angle of 40 degrees from the membrane plane (**Fig. 3a**). The orientation of the ectodomain also rotates the concave ligand-binding side of the LRRs away from the extracellular side of the 7TM domain (**Fig. 3b**). Each of these features physically separates the relaxin-2 binding site on the LRRs from the 7TMs. While a secondary binding site between relaxin-2 and the extracellular loops (ECLs) has been proposed^{38,39}, the active state that we captured through cryo-EM does not show any direct interaction between relaxin-2 and the ECLs.

The low-resolution map shows relaxin-2 bound to the concave side of the LRRs. To aid with modeling the relaxin-2–LRR interaction, we used crosslinking mass spectrometry (CLMS) with the RXFP1–G_s complex. Using the Extended-EDC approach with an EDDA crosslinker⁴⁰, Glu14^{B-chain} of relaxin-2 crosslinked with three residues on the LRRs, Glu206, Glu299, and Glu351 (**Fig. 3c,d**). Several residues involved in relaxin-2 binding have been previously characterized through mutational analysis in radioligand binding assays. In those experiments, a series of conserved residues on the B-chain of relaxin-2, Arg13, Arg17, and Ile/Val20, a motif known as the relaxin binding cassette, were proposed to be part of the interface^{41,42}. Multiple residues on the concave side of the LRRs were also found to be important for relaxin-2 binding³⁰. Interestingly, the Glu14^{B-chain} residue of relaxin-2 identified by CLMS is adjacent to the Arg13^{B-chain} of the relaxin binding cassette. Additionally, the crosslinked Glu299 residue on the LRRs had been previously identified by mutational studies, highlighting the close agreement between our CLMS data and prior functional analyses.

The residues from our CLMS experiment were used in combination with the mutational data as restraints for docking the relaxin-2–LRR interaction in HADDOCK⁴³ (**Fig. 3c,d**). The highest scoring HADDOCK models agreed well with our low resolution cryo-EM map and showed the B-chain of relaxin-2 bound to the concave side of the LRRs, while the A-chain made limited contacts (**Fig. 3e**). Additional density in the cryo-EM map is present next to the A-chain, which may belong to the ectodomain's linker region (**Fig. S8f**), which the A-chain has been proposed to bind²⁹. However, the low resolution and absence of crosslinks for these domains prevented further characterization of A-chain interactions.

The docked model of relaxin-2 bound to the LRRs predicted that one of the relaxin binding cassette residues, Arg17^{B-chain}, interacts with Glu206, a residue on the LRRs identified by CLMS that was not previously known to be involved in the binding site. To verify this interaction, we mutated Glu206 to Ala and tested the effect of the mutation on binding of an Fc-tagged relaxin-2 protein (SE301; see Methods). The Glu206 to Ala mutant expressed at equivalent levels to wild type receptor, but the single mutant reduced relaxin-2 binding to 42% of wild type levels (**Fig. 3f,g**), validating the proposed interaction.

Fig. 3 | Cryo-EM and crosslinking mass spectrometry reveal interactions between relaxin-2 and the leucine-rich repeats. **a**, Local resolution cryo-EM map of the full-length RXFP1-G_s complex. **b**, The relaxin-2 binding site is above and rotated away from the 7TM orthosteric site. **c**, Model of the relaxin-2-LRR interaction from HADDOCK. **d**, Details of the relaxin-2-LRR interface with residues identified in published binding studies in magenta, residues from CLMS in green, and Glu299 from both CLMS and published binding studies in yellow. Crosslink distances: Glu14^{B-chain}-Glu206 = 14.6 Å, Glu14^{B-chain}-Glu299 = 10 Å, Glu14^{B-chain}-Glu351 = 11.4 Å **e**, The relaxin-2-LRR model fit into the low resolution cryo-EM map. **f-g**, Receptor expression (**f**) and Fc-tagged relaxin-2 binding data (**g**) for the Glu206 to Ala mutation. Data are mean ± s.e.m. from technical triplicates.

Discussion

Our active-state cryo-EM structure of RXFP1–G_s revealed two new structural motifs involved in activation of RXFP1’s 7TM domain. RXFP1’s ECL2 was revealed to occupy the GPCR orthosteric ligand binding site, with the key residues Phe564^{ECL2} and Leu566^{ECL2} playing an essential role in receptor activation. Residues Leu402 and Leu403 of the hinge region, which interact with ECL2, were also required for RXFP1 signaling. These two signaling motifs were shown to be autoinhibited in the absence of relaxin-2 binding by both the LRRs and two residues of the receptor hinge region, Ile396 and Ser397, with the minimal inhibitory ectodomain construct containing both of these features. We also characterized the binding site of relaxin-2 on the LRRs, showing that it is physically separated from the 7TM domain.

Based on these results, it appears that relaxin-2’s mechanism of RXFP1 signaling involves indirectly releasing the autoinhibition of ECL2 (**Fig. 4**). As the LDLa module of the ectodomain is essential for RXFP1 signaling by relaxin-2¹³, it must play a role in signal transduction between relaxin-2 binding to the LRRs and transition of the ECL2/hinge interface to the active state. NMR studies with soluble constructs of RXFP1’s ECLs showed that the LDLa module can interact with ECL2^{29,39,44}. However, the LDLa module was not resolved in our map, indicating that it is likely mobile in the state that was captured in our cryo-EM studies.

An inactive-state structure will be necessary to define structural changes between receptor states, such as conformational changes of the LRR domain. Nonetheless, our results show that RXFP1 is not controlled by the steric “push-pull” mechanism proposed for GPHRs¹⁴, since the small size of relaxin-2 compared to glycoprotein hormones is incompatible with direct steric activation of the receptor. Our better understanding of RXFP1’s activation mechanism suggests small molecule or antibody agonists could be created to either mimic ECL2-induced activation or relieve inhibitory interactions between the ectodomain and ECL2. Such molecules could be useful therapeutics for the treatment of numerous cardiovascular and fibrotic diseases.

Fig. 4 | Model of RXFP1 activation by relaxin-2. The inactive state of RXFP1 is characterized by inhibitory interactions between ECL2, the hinge region, and the LRRs that prevent ECL2 from binding in the orthosteric pocket. Relaxin-2 binds to the concave side of the LRRs, away from the 7TM domain, also interacting with the linker region. This binding event is likely communicated to the ECL2/hinge/LRR interface by the LDLa module. In the active state of RXFP1, residues of the hinge region stabilize the binding of ECL2 in the GPCR orthosteric pocket, activating the 7TM domain.

Acknowledgements

We thank Marie Bao for critical reading of the manuscript, and the SBGrid Consortium for computational support of structural biology software. Cryo-EM data was collected at the Harvard Center for Cryo-Electron Microscopy at Harvard Medical School, and we thank them for their support and advice during data collection. This work was funded by an NIH Ruth L. Kirschstein Predoctoral fellowship (F31 GM128233) to S.C.E and a Blavatnik Biomedical Accelerator grant from Harvard Medical School to A.C.K. We thank GENCI-TGCC (France) for providing access to the supercomputer IRENE (grant no. Spe00015).

Competing interests statement

A.C.K. and S.C.E are inventors on a patent application for engineered single-chain relaxin proteins. A.C.K. is a co-founder and consultant for Tectonic Therapeutic and Seismic Therapeutic and for the Institute for Protein Innovation, a non-profit research institute.

Data availability

The cryo-EM model and maps for RXFP1-G_s-TM are deposited under accession codes 7TMW (PDB) and EMDB-26003, respectively. The cryo-EM map for RXFP1-G_s-FL is deposited under the accession code EMDB-26004. The rigid-body docking model of the RXFP1 LRRs bound to relaxin-2 is available on the website for the Kruse lab at Harvard Medical School.

Author contributions

The molecular cloning, protein expression and purification, cryo-EM grid preparation, flow cytometry, and cell signaling assays were performed by S.C.E. with supervision from A.C.K. The cryo-EM data were processed by S.C.E. and S.R. Model building and refinement were performed by S.C.E. with supervision from A.C.K. The evolutionary coupling analysis was performed by K.P.B. with supervision from D.S.M. The crosslinking mass spectrometry was performed by X.L., J.A.P., and J.M. with supervision from S.P.G. The molecular dynamics simulations were performed by X.C. The manuscript was written by S.C.E. with editing by A.C.K. and input from all authors.

Methods

Cloning of RXFP1 constructs

Residues 23-757 of human RXFP1 were cloned with an N-terminal hemagglutinin signal sequence, FLAG tag, and 3C protease site into the pcDNA-Zeo-tetO vector⁴⁵. Fusions to the miniG_s-399 protein with C-terminal truncations to RXFP1 were constructed using PCR followed by NEBuilder HiFi DNA Assembly (New England Biolabs). Truncations removed 10, 15, 20, 25, 30, or 35 residues from the receptor C-terminus. For signaling assays, human RXFP1 residues 23-757 with an N-terminal hemagglutinin signal sequence and FLAG tag were cloned into pcDNA-Zeo-tetO. Mutations to ECL2, hinge region, or LRR residues were introduced using Quikchange Lightning PCR (Agilent). Ectodomain truncations were constructed using PCR and NEBuilder HiFi DNA Assembly. Residue numbering is based on the canonical RXFP1 sequence beginning at the initiating Met residue (UniProt ID Q9HBX9).

Cell surface expression tests

RXFP1 signaling assay constructs were tested for cell surface expression using flow cytometry. HEK293T cells (ATCC) were maintained in Dulbecco's Modified Eagle Medium (DMEM) (Corning) with 10% (v/v) fetal bovine serum (FBS) (Sigma-Aldrich). Cells were plated at 100,000 cells/well into 12-well plates (Thermo Fisher Scientific). The following day, cells were transfected with 220 ng/well (unless otherwise stated) of human RXFP1 or empty vector DNA using FuGENE, according to the manufacturer's instructions (Promega). For **Fig. S5b**, 4.4 ng of wild type RXFP1 DNA was used per well, plus 215.6 ng of empty vector DNA, to lower wild type receptor expression levels. For **Fig. S5c**, 110 ng of 7TM + β_2 -Nterm DNA was used per well, plus 110 ng of empty vector DNA, in order to lower the expression of this construct to be equivalent to wild type levels. After twenty-four hours, the media was aspirated, and cells were detached by pipetting in phosphate-buffered saline (PBS) with 1% (v/v) FBS and 2 mM calcium chloride (Buffer A). Cells were distributed at 100,000 cells/well in 200 μ L Buffer A into a V-bottom 96-well plate (Corning). Cells were washed once and blocked in Buffer A by incubation for 30 minutes at 4°C. M1 anti-FLAG antibody (In house) labeled with Alexa Fluor 647 (Thermo Fisher Scientific) was incubated with the cells at 2.5 μ g/mL for 1 hour at 4°C. Cells were washed twice in Buffer A and resuspended in 100 μ L Buffer A. Fluorescence intensity was quantified using a CytoFLEX flow cytometer (Beckman Coulter) and around 2,000 events per

sample were collected. The gating strategy used is illustrated in **Fig. S9b**. Data was normalized using the wild type and empty vector mean fluorescence intensities as 100% and 0%, respectively, and plotted using GraphPad Prism.

cAMP signaling assay

The GloSensor assay from Promega, a live-cell signaling assay that detects cellular cAMP levels, was used to measure activation of G_s signaling through RXFP1. White, clear-bottom 96-well plates (Thermo Fisher Scientific) were coated with 30 μ L of 10 μ g/mL poly-D-lysine (Sigma-Aldrich), and washed once with PBS. HEK293T cells were then plated at 2×10^4 cells/well in DMEM with 10% (v/v) FBS. The following day, cells were transfected with 20 ng of GloSensor reporter plasmid and 20 ng (unless otherwise stated) of human RXFP1 or empty vector DNA per well using FuGENE, according to the manufacturer's instructions. For **Fig. 2d**, 0.4 ng of wild type RXFP1 DNA was used per well, plus 19.6 ng of empty vector DNA, according to ratios determined during cell surface expression tests. For **Fig. 2e**, 10 ng of 7TM + β_2 -Nterm DNA was used per well, plus 10 ng of empty vector DNA, according to ratios determined during cell surface expression tests. Twenty-four hours later, the media was replaced with 40 μ L of CO₂-independent media (Thermo Fisher Scientific) with 10% (v/v) FBS and 2 mg/mL D-luciferin (Goldbio) and incubated for 2 hours at room temperature (RT) in the dark. Measurements of luminescence with 1 second integration times were taken before ligand addition using a SpectraMax M5 microplate reader. For signaling curves, dilution series of native relaxin-2 (R&D Systems), were added to the cells and luminescence measurements were taken 30 minutes after relaxin-2 addition. For measurements of RXFP1 basal signaling versus relaxin-2 signaling, vehicle control (PBS + 0.1% bovine serum albumin) or relaxin-2 at 50 nM final concentration were added to the cells and the luminescence measurement was taken after 30 minutes. The maximum signaling response of relaxin-2 at RXFP1 for each experiment was normalized to 100%, and the percentages were plotted using GraphPad Prism.

Expression of RXFP1-miniG_s

RXFP1-miniG_s constructs for protein purification were expressed in inducible Expi293F tetR cells (Thermo Fischer Scientific). Stable cells lines were generated for RXFP1-miniG_s399-20res by plating Expi293F tetR cells adherently in DMEM with 10% (v/v) FBS. Cells were

transfected with linearized RXFP1-miniG_s399-20res pcDNA-Zeo-tetO DNA using Lipofectamine, according to the manufacturer's protocols (Thermo Fischer Scientific). Stable integrations were selected using 200 µg/mL Zeocin. After selection, polyclonal adherent RXFP1-miniG_s399-20res cells were readapted to suspension culture in antibiotic-free Expi293 media (Thermo Fischer Scientific), then maintained in Expi293 media with 10 µg/mL Zeocin. Stable RXFP1-miniG_s399-20res Expi293F tetR cells were expanded for protein expression in antibiotic-free Expi293 media and induced with 4 µg/mL doxycycline, 0.4% glucose, and 5 mM sodium butyrate. After 48 hours of induction, cells were harvested by spinning at 4,000 xg for 30 minutes at 4°C. The pellet was flash frozen in liquid nitrogen and stored at -80°C.

Purification of relaxin proteins:

To generate relaxin-2 proteins for RXFP1 complex purifications and flow cytometry binding assays, single-polypeptide relaxin-2 constructs were cloned with linkers connecting the B-chain and A-chain. Sequences for these proteins can be found in **Table S5**. Briefly, a His-tagged single-chain relaxin-2 (SE001) was expressed as a secreted protein from inducible Expi293F tetR cells. SE001 was purified from cell supernatants after 5 days of induction using Nickel Excel resin (GE Healthcare), followed by size exclusion chromatography with a Sephadex S200 column (GE Healthcare). Purified SE001, in 30 mM MES pH 6.5 and 300 mM sodium chloride, was aliquoted, flash frozen in liquid nitrogen, and stored at -80°C. A single-chain relaxin-2 fused to an antibody IgG1 Fc fragment (SE301) was expressed as described above in Expi293F tetR cells. SE301 was purified using Protein G resin (GE Healthcare) in 20 mM HEPES pH 7.5, 150 mM sodium chloride. Purified SE301 was aliquoted, flash frozen in liquid nitrogen, and stored at -80°C.

Purification of the RXFP1-G_s complex

Purification of the RXFP1-G_s complex used a cell pellet from a 2 L induction of the RXFP1-miniG_s399-20res Expi293F tetR stable cell line, and the purified proteins SE001 (single-chain relaxin-2), human Gβ₁γ₂, and Nb35-His-PrC. Gβ₁γ₂ and Nb35-His-PrC were expressed and purified according to previously published protocols¹⁶. The RXFP1-miniG_s399-20res cell pellet was lysed through osmotic shock by stirring in 250 mL cold Lysis buffer containing 20 mM HEPES pH 7.5, 2 mM magnesium chloride, 1 µL benzonase (Sigma Aldrich), 1 protease

inhibitor tablet (Thermo Fisher Scientific), and 100 nM SE001. Once stirring, iodoacetamide was added at a final concentration of 2 mg/mL. Lysed cells were centrifuged at 50,000 xg for 30 minutes at 4°C and the supernatant was decanted. Membranes were homogenized with a glass dounce (Thermo Fisher Scientific) in 270 mL Complexation buffer containing 20 mM HEPES pH 7.5, 350 mM sodium chloride, 20% (v/v) glycerol, 1 μ L benzonase, and 1 protease inhibitor tablet. After homogenization, 960 μ g G β ₁ γ ₂, 740 μ g Nb35-His-PrC (1:2:4 ratio of RXFP1:G β ₁ γ ₂:Nb35-His-PrC), 100 nM SE001, and 0.2 U apyrase (New England Biolabs) were added to the membranes and the solution was stirred for 1 hour at 4°C. Next, 30 mL of 10X Detergent buffer [10% (w/v) lauryl maltose neopentyl glycol (L-MNG; Anatrace) and 1% (w/v) cholesterol hemisuccinate (CHS; Anatrace)] were added to the solution, dounce homogenization was repeated, and the solution was stirred for 2 hours at 4°C to extract the RXFP1-G_s complex from the membrane. After 2 hours, the solution was centrifuged at 50,000 xg for 30 minutes at 4°C and the supernatant was filtered with a glass fiber prefilter (Millipore). Calcium was added to the solubilized membranes at a final concentration of 2 mM and the solution was loaded by gravity flow over 3 mL M1 anti-FLAG resin (In house). This step of the purification was based on the N-terminal FLAG tag of RXFP1-miniG_s399-20res. The M1 resin was equilibrated with Wash buffer 1 [0.1% (w/v) L-MNG, 0.01% (w/v) CHS, 350 mM sodium chloride, 20 mM HEPES pH 7.5, 2 mM calcium chloride, and 100 nM SE001]. After loading the solubilized membrane fraction, the column was subsequently washed with 50 mL each of Wash buffer 1, Wash buffer 2 [0.1% (w/v) L-MNG, 0.01% (w/v) CHS, 350 mM sodium chloride, 20 mM HEPES pH 7.5, 2 mM calcium chloride, 2 mM magnesium chloride, 5 mM adenosine 5'-triphosphate magnesium salt, and 100 nM SE001], and Wash buffer 3 [0.01% (w/v) L-MNG, 0.001% (w/v) CHS, 350 mM sodium chloride, 20 mM HEPES pH 7.5, 2 mM calcium chloride, and 100 nM SE001]. M1 resin was eluted with 0.01% (w/v) L-MNG, 0.001% (w/v) CHS, 350 mM sodium chloride, 20 mM HEPES pH 7.5, 0.5 mg/mL FLAG peptide (GenScript), and 100 nM SE001, then 2 mM calcium chloride was added to the eluted protein. Next, 0.5 mL anti-Protein C resin (In house) was equilibrated with Protein C wash buffer [0.005% (w/v) L-MNG, 0.0005% (w/v) CHS, 350 mM sodium chloride, 20 mM HEPES pH 7.5, 2 mM calcium chloride, and 100 nM SE001]. The M1 elution with 2 mM calcium chloride was loaded onto the anti-Protein C resin by gravity flow, washed with 10 mL Protein C wash buffer, and eluted with 0.005% (w/v) L-MNG, 0.0005% (w/v) CHS, 350 mM sodium chloride, 20 mM HEPES pH 7.5,

0.5 mg/mL Protein C peptide (GenScript), and 100 nM SE001. This step of the purification was based on the Protein C tag of Nb35-His-PrC. The elution was concentrated with a 3 kDa molecular weight cutoff centrifugal concentrator (Millipore) and loaded onto a Sephadex S200 column (GE Healthcare) in SEC buffer [0.005% (w/v) L-MNG, 0.0005% (w/v) CHS, 350 mM sodium chloride, 20 mM HEPES pH 7.5, and 100 nM SE001]. Peak fractions from size exclusion were concentrated with a 3 kDa molecular weight cutoff centrifugal concentrator and either stored overnight at 4°C prior to grid freezing or immediately used to freeze cryo-EM grids.

Cryo-EM grid preparation and data collection

Cryo-EM grids were prepared using QUANTIFOIL[®] holey carbon grids (400-mesh, copper, R1.2/1.3; Electron Microscopy Sciences). Grids were washed with ethyl acetate (Sigma Aldrich) and then glow-discharged at -20 mA for 60 seconds with a Pelco Easiglow. 3 µL of sample at 0.2 to 0.4 mg/mL was applied to the grids. Grids were plunge-frozen in liquid ethane using a Mark IV Vitrobot (Thermo Fisher Scientific) at 10°C and 100% humidity with a 10 second wait time, 3 to 4 second blot time, and blot force of 15.

Cryo-EM data for the RXFP1-G_s complex were collected in four separate sessions. Grids were imaged with a Titan Krios microscope (Thermo Fisher Scientific) at 300 kV with a Gatan BioQuantum GIF/K3 direct electron detection camera in counting mode. Movies were collected with 50 frames each at 81,000x, corresponding to 1.06 Å per pixel, and a total dose of around 52 electrons per Å². Defocus values ranged from -0.8 to -2.3 µm.

Cryo-EM data processing

Cryo-EM data was collected in four separate session and initially processed individually, then particle stacks were merged to generate the final maps. Motion correction was carried out with MOTIONCOR2⁴⁶ and CTF parameters were estimated with CTFFIND-4.1⁴⁷. Particles were picked in RELION 3.1⁴⁸ using Laplacian of Gaussian autopicking. A map from a previous small dataset collected on a Talos Arctica microscope and processed in RELION was used as an initial model. Particles were extracted with a box size large enough to include the entire complex (288 pixels). Multiple rounds of 2D classifications in RELION used a circular mask of 150 Å around only the micelle and G proteins as the first step of particle sorting for both 3D reconstructions. Masking of the micelle and G proteins was used to initially sort high quality particles from

contamination because 2D classifications attempted for the entire particle showed weak signal for the ectodomain, likely indicating a region of higher flexibility. It became evident at this stage of data processing that the datasets showed a preferred orientation for one of the side views of the RXFP1–G_s complex. Processing 2D classifications of the preferred view separately from other views of the complex resulted in a larger number of particles from different orientations in the final particle stacks.

For the map of the 7TMs, masked 3D classifications of the 7TM domain bound to G proteins were performed in RELION. Following these classifications, iterative 3D refinements (using a mask of the micelle and G proteins) and Bayesian particle polishing steps were carried out in RELION. At this stage, particle stacks from different datasets were joined together, followed by additional 2D classifications with a 150 Å circular mask and masked 3D classifications. Finally, 3D focused classifications without alignments were carried out using masks for either the ECLs, TM helices, or helix8. These particle stacks were imported into cryoSPARC v3.1.0³⁷ and used with Non-uniform Refinement (New). The cryoSPARC refined maps were post-processed in DeepEMhancer⁴⁹ using the half-maps as input. This data processing workflow is described in **Fig. S2**.

For the map of the full-length RXFP1–G_s complex, rounds of unmasked 3D classifications were used in RELION after the 2D classifications described above. Next, the micelle and G proteins were subtracted from the particles and 3D focused classifications without alignments were carried out using a mask on the entire ectodomain of RXFP1. 3D refinements with a mask of the entire complex and iterative Bayesian particle polishing steps were then performed in RELION. Following particle polishing, particle stacks from different datasets were combined and 3D classifications with a mask around the entire complex were performed in RELION, followed by 3D refinement in RELION with a mask around the entire complex. Finally, local resolution estimation and filtering were performed in RELION on the final maps. This data processing workflow is described in **Fig. S3**.

Model building and refinement

A combined focused map for the RXFP1 7TM domain bound to G proteins was generated using the DeepEMhancer post-processed maps as inputs in Phenix⁵⁰. Model building for the RXFP1 7TM domain bound to heterotrimeric G_s and Nb35 was performed using

DeepEMhancer focused and combined focused cryo-EM maps in *Coot*⁵¹. An initial model for the RXFP1 7TM domain (Gly395 to Gly709) was generated using trRosetta⁵² and manually fit into the DeepEMhancer focused cryo-EM map (for ECLs) in Chimera⁵³. Intracellular and extracellular loops and the hinge region were removed from the trRosetta model and manually rebuilt in *Coot*. ECL2 residues Lys554–Tyr557 and hinge region residues Glu400–Leu403 were manually built with the coordinates of the human RXFP1 AlphaFold2 model³¹ overlayed in *Coot*. Initial models for miniG_s399, Gβγ, and Nb35 were generated in MODELLER⁵⁴ using coordinates from PDB ID 6GDG⁵⁵. All models were refined with Phenix real-space refinement⁵⁰ using secondary structure restraints against the DeepEMhancer combined focused map. Statistics for the final model were evaluated using MolProbity in a Phenix comprehensive validation (cryo-EM) job that used the map from cryoSPARC v3.1.0 Non-uniform Refinement (New) and the final model as inputs. Figures were prepared using PyMOL and ChimeraX⁵⁶. Structural biology programs used in this work, other than cryoSPARC, were compiled and configured by SBGrid⁵⁷. Refinement statistics are present in **Table S1** and representative images of cryo-EM map and model quality are shown in **Fig. S8**.

Evolutionary coupling analysis

For comparing the constructed model to the strongest evolutionarily coupled pairs, we first downloaded the Uniprot canonical sequence for Q9HBX9 for residues 405-689 and then used the Jackhmmer software suite⁵⁸ to build multiple sequence alignments across multiple bitscore thresholds based on the June 2019 download of the Uniref100 database⁵⁹. We then chose an alignment with 352,511 sequences, with 90% of columns consisting of fewer than 30% gaps, to move forward with in analysis. We used the EVcouplings v0.0.5 software, available at <github.com/debbiemarkslab>, to identify evolutionary couplings (ECs) for this alignment.

To incorporate the LRR region, we ran the Q9HBX9 sequence for residues 120-757 in Jackhmmer, against the 02/2021 Uniref100 database for normalized bitscores between 0.1 and 0.9. We chose an alignment with 8,145 sequences with 86% of columns consisting of fewer than 30% gaps, and input it to v0.1.1 EVcouplings software. The strongest couplings were ranked based on assigning probabilities from a logistic regression model new to v0.1.1. This version of the pipeline was also used for plotting all contact maps.

Crosslinking mass spectrometry

The RXFP1-G_s complex used for CLMS was prepared as described above and crosslinking reactions were carried out the following day, after storage overnight at 4°C. CLMS was performed as previously described⁴⁰. Briefly, crosslinking reactions were carried out for 1 h at room temperature in 100 mM MOPS Buffer, pH 6.5 with 50 mM EDC, ~24 mM EDDA linker, and 20 mM sulfo-NHS. Reactions were quenched with hydroxylamine to a final concentration of 100 mM. Samples were reduced for 1 h in 2% SDS and 5 mM TCEP followed by alkylation with 10 mM iodoacetamide in the dark for 30 min and quenching with 5 mM DTT for 15 min. Samples were then processed with the SP3⁶⁰ method and digested with trypsin (Promega) at 1:25 enzyme:substrate ratio overnight at 37°C. Digested peptides were acidified with 10% formic acid to pH ~2 and desalted using stage tips with Empore C18 SPE Extraction Disks (3M) and dried under vacuum.

Sample was reconstituted in 5% formic acid (FA)/5% acetonitrile and analyzed in the Orbitrap Eclipse Mass Spectrometer (Thermo Fischer Scientific) coupled to an EASY-nLC 1200 (Thermo Fisher Scientific) ultra-high pressure liquid chromatography (UHPLC) pump, as well as a high-Field Asymmetric waveform Ion Mobility Spectrometry (FAIMS) interface. Peptides were separated on an in-house packed 100 µm inner diameter column packed with 35 cm of Accucore C18 resin (2.6 µm, 150 Å, ThermoFisher), using a gradient consisting of 5–35% (ACN, 0.125% FA) over 135 min at ~500 nL/min. The instrument was operated in data-dependent mode. FTMS1 spectra were collected at a resolution of 120K, with an automated gain control (AGC) target of 5×10^5 , and a max injection time of 50 ms. The most intense ions were selected for MS/MS for 1s in top-speed mode, while switching among three FAIMS compensation voltages (CV): -40, -60, and -80 V in the same method. Precursors were filtered according to charge state (allowed $3 \leq z \leq 7$), and monoisotopic peak assignment was turned on. Previously interrogated precursors were excluded using a dynamic exclusion window (60 s ± 7 ppm). MS2 precursors were isolated with a quadrupole mass filter set to a width of 0.7 m/z and analyzed by FTMS2, with the Orbitrap operating at 30K resolution, an AGC target of 100K, and a maximum injection time of 150 ms. Precursors were then fragmented by high-energy collision dissociation (HCD) at a 30% normalized collision energy.

Mass spectra were processed and searched using the PIXL search engine⁴⁰. The sequence database contained proteins identified at 1% FDR in non-cross-linked Comet⁶¹ search. For PIXL

search, precursor tolerance was set to 15 ppm and fragment ion tolerance to 10 ppm. Methionine oxidation was set as a variable modification in addition to mono-linked mass of +130.110613 for EDDA. Crosslinked peptides were searched assuming zero-length (-18.010565) and EDDA crosslinker +112.100048. Crosslinked searches considered 60 protein sequences to ensure sufficient statistics for FDR estimation. Matches were filtered to 1% FDR on the unique peptide level using linear discriminant features as previously described⁴⁰.

Docking

HADDOCK⁴³ was used to dock the relaxin-2–LRR interaction. Docking used the human relaxin-2 X-ray crystal structure (PDB ID: 6RLX)⁶² and a model of the LRRs from residues 104–391 of the AlphaFold2³¹ prediction for human RXFP1. Residues from CLMS studies that were part of crosslinks between relaxin-2 and RXFP1 were used as active restraints in the docking run. These CLMS residues were Glu14^{B-chain} of relaxin-2, and Glu206, Glu299, and Glu351 of RXFP1. Residues identified to be important for relaxin-2 binding from published mutations in radioligand binding assays were also used as active restraints^{30,41,42}. These residues included Arg13^{B-chain}, Arg17^{B-chain}, and Ile20^{B-chain} of relaxin-2 and Trp202, Ile204, Leu226, Asp253, Glu255, Glu299, and Asp301 of RXFP1. The resulting docking models were analyzed according to the HADDOCK scoring function and fit into the low resolution cryo-EM map of the RXFP1 ectodomain in ChimeraX⁵⁶.

Flow cytometry binding assay

A flow cytometry assay was used to measure the binding of an Fc-tagged relaxin-2 protein, SE301, to Expi293F cells transfected with human RXFP1 or empty pcDNA-Zeo-tetO vector. Expi293F tetR cells were grown in Expi293 media and transfected using FectoPRO (Polyplus), according to the manufacturer's protocols. The cells were enhanced 24 hours post-transfection with 0.4% glucose and induced 48 hours post-transfection with 4 µg/mL doxycycline and 5 mM sodium butyrate. After 24 hours of induction, cells were harvested by spinning at 200 xg for 5 minutes at 4°C and washed once with HBS with 1% (v/v) FBS and 2 mM calcium chloride (Buffer B). Cells were plated into a V-bottom 96-well plate (Corning) at 100,000 cells/well and blocked by incubation in Buffer B for 30 minutes at 4°C. After blocking, cells were centrifuged at 200 xg for 5 minutes at 4°C, resuspended in 100 µL of Buffer B

containing 500 nM SE301, and incubated for 1 hour at 4°C. Cells were then centrifuged at 200 xg for 5 minutes at 4°C, washed twice with 200 µL Buffer B, and resuspended in 100 µL Buffer B containing 100 nM M1 anti-FLAG antibody labeled with Alexa Fluor 488 (In house) and Alexa Fluor 647 anti-human IgG Fc (BioLegend) diluted 1:100 (v/v). Cells were incubated in secondary antibodies for 30 minutes at 4°C, washed once with 200 µL Buffer B, and resuspended in 100 µL Buffer B for flow cytometry.

Samples were analyzed on a CytoFLEX flow cytometer (Beckman Coulter) and gated according to plots of FSC-A/SCA-A, FSC-A/FSC-H, and receptor expression according to Alexa Fluor 488 M1 anti-FLAG antibody binding. Approximately 500 events/sample were collected from cells expressing receptor for human RXFP1-transfected cells or post-FSC-A/FSC-H gating for empty vector-transfected cells. Receptor expression levels were analyzed using the Alexa Fluor 488 M1 anti-FLAG mean fluorescence intensity (MFI) from post-FSC-A/FSC-H gated cells. SE301 binding was analyzed by plotting the MFI for Alexa Fluor 647 anti-human IgG Fc cells that had been gated for receptor expression. The gating strategy used is illustrated in **Fig. S9a**. The data were plotted and analyzed in GraphPad Prism.

Molecular dynamics

The initial model was built from the cryo-EM structure reported here. The missing segments in ECL2, ICL3 and TM6 were generated by MODELLER v9.15⁶³. PACKMOL-Memgen⁶⁴ was used to assign the side-chain protonation states and embed the models in a bilayer of POPC lipids. The system was solvated in a periodic $80 \times 80 \times 110 \text{ \AA}^{64}$ box of explicit water and neutralized with 0.15 M of Na⁺ and Cl⁻ ions. We used the Amber ff14SB⁶⁵ and lipid 14⁶⁶ force fields, the TIP3P⁶⁷ water model and the Joung-Cheatham⁶⁸ parameters the ions. For the simulations RXFP1 7TM deactivation, a Na⁺ ion was placed at the conserved Na⁺-binding site (between Asp451^{2,50} and Ser495^{3,35}). For the simulations of RXFP1 7TM autoactivation, the truncated AlphaFold2³¹ model was used to build the system and Asp451^{2,50} was protonated.

After energy minimization, all-atom MD simulations were carried out using Gromacs 5.1⁶⁹ patched with the PLUMED 2.3 plugin⁷⁰. The LINCS algorithm⁷¹ was applied to constrain bonds involving hydrogen atoms, allowing a time step of 2 fs. Each system was gradually heated to 310 K and pre-equilibrated during 10 ns of brute-force MD in the *NPT*-ensemble. The replica exchange with solute scaling (REST2)⁷² technique was used to enhance the conformational

sampling. A total of 64 replicas of simulations were performed in the *NVT* ensemble. REST2 is a type of Hamiltonian replica exchange simulation scheme. Besides the original simulation, many replicas of the same system were simulated simultaneously. The additional replicas have modified free energy surfaces, in which the energy barriers are easier to cross than in the original simulation system. By frequently swapping the replicas and the original system during the MD, the simulations “travel” on different free energy surfaces and easily visit various conformational zones. Finally, only the samples on the original free energy surface are collected. The additional replicas are artificial to ease barrier crossing, which are discarded after the simulations. REST2, in particular, modifies the free energy surfaces by scaling (reducing) the force constants of the “solute” molecules in the simulation system. In this case, the protein was considered as “solute”—the force constants of its van der Waals, electrostatic and dihedral terms were subject to scaling—in order to facilitate the conformational changes. The scaling factors were generated using the Patriksson-van der Spoel approach⁷³ and effective temperatures ranging from 310 K to 1000 K. Exchange between replicas was attempted every 1000 simulation steps. This setup resulted in an average exchange probability of ~40%. We performed 80 ns × 64 replicas of MD in the *NVT* ensemble for each system. The first 30 ns were discarded for equilibration. The simulation trajectory on the original unscaled free energy surface was reassembled and analyzed.

References

1. Petrie, E. J., Lagaida, S., Sethi, A., Bathgate, R. A. D. & Gooley, P. R. In a class of their own - RXFP1 and RXFP2 are unique members of the LGR family. *Front. Endocrinol. (Lausanne)*. **6**, 1–9 (2015).
2. Sheau, Y. H. *et al.* Activation of orphan receptors by the hormone relaxin. *Science* (80-.). **295**, 671–674 (2002).
3. Hisaw, F. L. Experimental relaxation of the pubic ligament of the guinea pig. *Proc. Soc. Exp. Biol. Med.* **23**, 661–663 (1926).
4. MacLennan, A. H., Nicolson, R. & Green, R. C. Serum relaxin in pregnancy. *Lancet* **328**, 241–243 (1986).
5. McGuane, J. T. *et al.* Role of relaxin in maternal systemic and renal vascular adaptations during gestation. *Ann. N. Y. Acad. Sci.* **1160**, 304–312 (2009).
6. Samuel, C. S. *et al.* Anti-fibrotic actions of relaxin. *Br. J. Pharmacol.* **174**, 962–976 (2017).
7. Sarwar, M., Du, X. J., Dschietzig, T. B. & Summers, R. J. The actions of relaxin on the human cardiovascular system. *Br. J. Pharmacol.* **174**, 933–949 (2017).
8. Devarakonda, T. & Salloum, F. N. Heart disease and relaxin: New actions for an old hormone. *Trends Endocrinol. Metab.* **29**, 338–348 (2018).
9. Leo, C. H., Jelinic, M., Ng, H. H., Tare, M. & Parry, L. J. Serelaxin: A Novel Therapeutic for Vascular Diseases. *Trends Pharmacol. Sci.* **37**, 498–507 (2016).
10. Bennett, R. G. Relaxin and its role in the development and treatment of fibrosis. *Transl. Res.* **154**, 1–6 (2009).
11. Hsu, S. Y. *et al.* The three subfamilies of leucine-rich repeat-containing G protein-coupled receptors (LGR): Identification of LGR6 and LGR7 and the signaling mechanism for LGR7. *Mol. Endocrinol.* **14**, 1257–1271 (2000).
12. Van Hiel, M. B., Vandersmissen, H. P., Van Loy, T. & Vanden Broeck, J. An evolutionary comparison of leucine-rich repeat containing G protein-coupled receptors reveals a novel LGR subtype. *Peptides* **34**, 193–200 (2012).
13. Scott, D. J. *et al.* 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 (2006).
14. Duan, J. *et al.* Structures of full-length glycoprotein hormone receptor signalling complexes. *Nature* **598**, 688–692 (2021).
15. Carpenter, B. & Tate, C. G. Engineering a minimal G protein to facilitate crystallisation of G protein-coupled receptors in their active conformation. *Protein Eng. Des. Sel.* **29**, 583–593 (2016).
16. Rasmussen, S. G. F. *et al.* Crystal structure of the β_2 adrenergic receptor-Gs protein complex. *Nature* **477**, 549–555 (2011).
17. Hilger, D., Masureel, M. & Kobilkat, B. K. Structure and dynamics of GPCR signaling complex. *Nat. Struct. Mol. Biol.* **25**, 4–12 (2018).
18. Erlandson, S. C., McMahon, C. & Kruse, A. C. Structural basis for G protein-coupled receptor signaling. *Annu. Rev. Biophys.* **47**, 1–18 (2018).
19. Ballesteros, J. A. & Weinstein, H. Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G protein-coupled receptors. *Methods Neurosci.* **25**, 366–428 (1995).

- 767 20. Wang, J., Hua, T. & Liu, Z.-J. Structural features of activated GPCR signaling complexes.
768 *Curr. Opin. Struct. Biol.* **63**, 82–89 (2020).
- 769 21. Venkatakrishnan, A. J. *et al.* Molecular signatures of G-protein-coupled receptors. *Nature*
770 **494**, 185–194 (2013).
- 771 22. Ring, A. M. *et al.* Adrenaline-activated structure of β 2-adrenoceptor stabilized by an
772 engineered nanobody. *Nature* **502**, 575–579 (2013).
- 773 23. Wingler, L. M., McMahon, C., Staus, D. P., Lefkowitz, R. J. & Kruse, A. C. Distinctive
774 activation mechanism for angiotensin receptor revealed by a synthetic nanobody. *Cell*
775 **176**, 479–490.e12 (2019).
- 776 24. Lin, X. *et al.* Structural basis of ligand recognition and self-activation of orphan GPR52.
777 *Nature* **579**, 152–157 (2020).
- 778 25. Brüser, A. *et al.* The activation mechanism of glycoprotein hormone receptors with
779 implications in the cause and therapy of endocrine diseases. *J. Biol. Chem.* **291**, 508–520
780 (2016).
- 781 26. Schulze, A. *et al.* The intramolecular agonist is obligate for activation of glycoprotein
782 hormone receptors. *FASEB J.* **34**, 11243–11256 (2020).
- 783 27. Jiang, X., Dias, J. A. & He, X. Structural biology of glycoprotein hormones and their
784 receptors: Insights to signaling. *Mol. Cell. Endocrinol.* **382**, 424–451 (2014).
- 785 28. Martin, A. L., Steurer, M. A. & Aronstam, R. S. Constitutive activity among orphan class-
786 A G protein coupled receptors. *PLoS One* **10**, e0138463 (2015).
- 787 29. Sethi, A. *et al.* The complex binding mode of the peptide hormone H2 relaxin to its
788 receptor RXFP1. *Nat. Commun.* **7**, 11344 (2016).
- 789 30. Büllesbach, E. E. & Schwabe, C. The trap-like relaxin-binding site of the leucine-rich G-
790 protein-coupled receptor 7. *J. Biol. Chem.* **280**, 14051–14056 (2005).
- 791 31. Tunyasuvunakool, K. *et al.* Highly accurate protein structure prediction for the human
792 proteome. *Nature* **596**, 590–596 (2021).
- 793 32. Marks, D. S. *et al.* Protein 3D structure computed from evolutionary sequence variation.
794 *PLoS One* **6**, (2011).
- 795 33. Marks, D. S., Hopf, T. A. & Sander, C. Protein structure prediction from sequence
796 variation. *Nat. Biotechnol.* **30**, 1072–1080 (2012).
- 797 34. Sjødt, M. *et al.* Structure of the peptidoglycan polymerase RodA resolved by evolutionary
798 coupling analysis. *Nature* **556**, 118–121 (2018).
- 799 35. Zheng, S. *et al.* Structure and mutagenic analysis of the lipid II flippase MurJ from
800 *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 6709–6714 (2018).
- 801 36. Toth-Petroczy, A. *et al.* Structured states of disordered proteins from genomic sequences.
802 *Cell* **167**, 158–170.e12 (2016).
- 803 37. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: Algorithms for
804 rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).
- 805 38. Halls, M. L. *et al.* Multiple binding sites revealed by interaction of relaxin family peptides
806 with native and chimeric relaxin family peptide receptors 1 and 2 (LGR7 and LGR8). *J.*
807 *Pharmacol. Exp. Ther.* **313**, 677–687 (2005).
- 808 39. Diepenhorst, N. A. *et al.* Investigation of interactions at the extracellular loops of the
809 relaxin family peptide receptor 1 (RXFP1). *J. Biol. Chem.* **289**, 34938–34952 (2014).
- 810 40. Mintseris, J. & Gygi, S. P. High-density chemical cross-linking for modeling protein
811 interactions. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 93–102 (2020).
- 812 41. Büllesbach, E. E., Yang, S. & Schwabe, C. The receptor-binding site of human relaxin II:

- 813 A dual prong-binding mechanism. *J. Biol. Chem.* **267**, 22957–22960 (1992).
- 814 42. Büllesbach, E. E. & Schwabe, C. The relaxin receptor-binding site geometry suggests a
815 novel gripping mode of interaction. *J. Biol. Chem.* **275**, 35276–35280 (2000).
- 816 43. van Zundert, G. C. P. *et al.* The HADDOCK2.2 web server: User-friendly integrative
817 modeling of biomolecular complexes. *J. Mol. Biol.* **428**, 720–725 (2016).
- 818 44. Sethi, A. *et al.* Structural insights into the unique modes of relaxin-binding and tethered-
819 agonist mediated activation of RXFP1 and RXFP2: Defining the tethered agonist of
820 RXFP1 and RXFP2. *J. Mol. Biol.* **433**, 167217 (2021).
- 821 45. Staus, D. P. *et al.* Sortase ligation enables homogeneous GPCR phosphorylation to reveal
822 diversity in β -arrestin coupling. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 3834–3839 (2018).
- 823 46. Zheng, S. Q., Palovcak, E., Armache, J.-P., Cheng, Y. & Agard, D. A. Anisotropic
824 Correction of Beam-induced Motion for Improved Single-particle Electron Cryo-
825 microscopy. *bioRxiv* 1–30 (2016) doi:10.1101/061960.
- 826 47. Rohou, A. & Grigorieff, N. CTFFIND4: Fast and accurate defocus estimation from
827 electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015).
- 828 48. Scheres, S. H. W. RELION: Implementation of a Bayesian approach to cryo-EM structure
829 determination. *J. Struct. Biol.* **180**, 519–530 (2012).
- 830 49. Sanchez-Garcia, R. *et al.* DeepEMhancer: a deep learning solution for cryo-EM volume
831 post-processing. *Commun. Biol.* **4**, 1–8 (2021).
- 832 50. Liebschner, D. *et al.* Macromolecular structure determination using X-rays, neutrons and
833 electrons: Recent developments in Phenix. *Acta Crystallogr. Sect. D Struct. Biol.* **75**, 861–
834 877 (2019).
- 835 51. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot.
836 *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 486–501 (2010).
- 837 52. Du, Z. *et al.* The trRosetta server for fast and accurate protein structure prediction. *Nat.*
838 *Protoc.* **16**, 5634–5651 (2021).
- 839 53. Pettersen, E. F. *et al.* UCSF Chimera - A visualization system for exploratory research and
840 analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
- 841 54. Webb, B. & Sali, A. Comparative protein structure modeling using MODELLER. *Curr.*
842 *Protoc. Bioinforma.* **54**, 5.6.1–5.6.37 (2016).
- 843 55. García-Nafria, J., Lee, Y., Bai, X., Carpenter, B. & Tate, C. G. Cryo-EM structure of the
844 adenosine A2A receptor coupled to an engineered heterotrimeric G protein. *Elife* **7**, 1–19
845 (2018).
- 846 56. Pettersen, E. F. *et al.* UCSF ChimeraX: Structure visualization for researchers, educators,
847 and developers. *Protein Sci.* **30**, 70–82 (2021).
- 848 57. Morin, A. *et al.* Collaboration gets the most out of software. *Elife* **2**, e01456 (2013).
- 849 58. Potter, S. C. *et al.* HMMER web server: 2018 update. *Nucleic Acids Res.* **46**, W200–W204
850 (2018).
- 851 59. Suzek, B. E., Wang, Y., Huang, H., McGarvey, P. B. & Wu, C. H. UniRef clusters: A
852 comprehensive and scalable alternative for improving sequence similarity searches.
853 *Bioinformatics* **31**, 926–932 (2015).
- 854 60. Moggridge, S., Sorensen, P. H., Morin, G. B. & Hughes, C. S. Extending the compatibility
855 of the SP3 paramagnetic bead processing approach for proteomics. *J. Proteome Res.* **17**,
856 1730–1740 (2018).
- 857 61. Eng, J. K., Jahan, T. A. & Hoopmann, M. R. Comet: An open-source MS/MS sequence
858 database search tool. *Proteomics* **13**, 22–24 (2013).

62. Eigenbrot, C. *et al.* X-ray structure of human relaxin at 1.5 Å. Comparison to insulin and implications for receptor binding determinants. *J. Mol. Biol.* **221**, 15–21 (1991).
63. Mancebo, M. C., Eisen, J. L., Sibrava, N. J., Dyck, I. R. & Rasmussen, S. A. Comparative protein structure modeling using Modeller. in *Curr Protoc Bioinformatics* vol. Chapter 5 (2006).
64. Schott-Verdugo, S. & Gohlke, H. PACKMOL-Memgen: A simple-to-use, generalized workflow for membrane-protein-lipid-bilayer system building. *J. Chem. Inf. Model.* **59**, 2522–2528 (2019).
65. Lindorff-Larsen, K. *et al.* Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins Struct. Funct. Bioinforma.* **78**, 1950–1958 (2010).
66. Dickson, C. J. *et al.* Lipid14: The amber lipid force field. *J. Chem. Theory Comput.* **10**, 865–879 (2014).
67. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J Chem Phys* **79**, 926–935 (1983).
68. Joung, I. S. & Cheatham, T. E. Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations. *J. Phys. Chem. B* **112**, 9020–9041 (2008).
69. Van Der Spoel, D. *et al.* GROMACS: Fast, flexible, and free. *J. Comput. Chem.* **26**, 1701–1718 (2005).
70. Tribello, G. A., Bonomi, M., Branduardi, D., Camilloni, C. & Bussi, G. PLUMED 2: New feathers for an old bird. *Comput. Phys. Commun.* **185**, 604–613 (2014).
71. Hess, B., Bekker, H., Berendsen, H. J. C. & Fraaije, J. G. E. M. LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.* **18**, 1463–1472 (1997).
72. Wang, L., Friesner, R. A. & Berne, B. J. Replica exchange with solute scaling: a more efficient version of replica exchange with solute tempering (REST2). *J Phys Chem B* **115**, 9431–9438 (2011).
73. Patriksson, A. & van der Spoel, D. A temperature predictor for parallel tempering simulations. *Phys. Chem. Chem. Phys* **10**, 2073–2077 (2008).