The relaxin receptor RXFP1 signals through a mechanism of autoinhibition

Authors: Sarah C. Erlandson[1], Shaun Rawson[1], Kelly P. Brock[2], Xinyue Liu[1], Joao A. Paulo[1], Julian Mintseris[1], Steven P. Gygi[1], Debora S. Marks[2], Xiaojing Cong[3], Andrew C. Kruse[1,*]

Affiliations:

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

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

[3]Institut 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 relaxins[1]. The LGRs are involved in a variety of physiological processes across reproductive and developmental biology. RXFP1, the receptor for the relaxin-2 hormone[2], 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 parturition[3–5]. 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 diseases[8–10].

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-2[13]. 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 conformation[14]. 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.

63 We optimized receptor expression using fusions to an engineered minimal G_α protein[15], then
64 formed a larger complex with the addition of G protein β_1 and γ_2 subunits, and determined the
65 structure using cryo-electron microscopy. Unexpectedly, the structure revealed that RXFP1's
66 extracellular loop 2 (ECL2) occupies the GPCR orthosteric ligand-binding pocket in the active
67 state. Results from structural and functional studies define a mechanism in which ECL2
68 conformation is regulated by the receptor's LRRs and hinge region, a short segment between the
69 LRRs and 7TMs. Collectively, these studies identify several conformational switches in both the
70 receptor ectodomain and 7TMs, showing that the concerted action of multiple receptor domains
71 controls the transduction of RXFP1 signaling by its agonist, relaxin-2.
72
73 **Cryo-EM structure of the RXFP1–G protein complex**
74 Wild-type (WT) full-length RXFP1 receptor could be expressed only at very low levels
75 in mammalian cells. To enable structural studies, we cloned a fusion of RXFP1 to the engineered
76 G_α protein minimal G_s (mini-G_s)[15]. Truncations of RXFP1's flexible C-terminus further
77 increased expression levels, with the optimal expression construct having a C-terminal truncation
78 of 20 amino acids (**Fig. S1**). The binding of the catalytically inactive mini-G_s protein blocks
79 RXFP1 signaling through endogenous G_s proteins and likely stabilizes the receptor, leading to
80 higher purification yields. The fusion protein of RXFP1 and mini-G_s was purified in complex
81 with human G protein β_1 and γ_2 subunits, the camelid antibody VHH fragment nanobody 35
82 (Nb35)[16], and an engineered version of the agonist relaxin-2 (SE001; see Methods) to form an
83 agonist–GPCR–G protein complex (hereafter referred to as RXFP1–G_s).
84 Cryo-electron microscopy (cryo-EM) was used for structural studies of RXFP1–G_s.
85 Initial two-dimensional classification analysis revealed averages that showed clear density for
86 RXFP1's 7TM domain and the heterotrimeric G protein. In contrast, the density for RXFP1's
87 ectodomain was weak and poorly defined, indicating flexibility of the ectodomain with respect to
88 the transmembrane domain. Due to the conformational heterogeneity of RXFP1–G_s, we analyzed
89 the cryo-EM data using two different approaches. The first utilized masking of RXFP1's 7TM
90 and G proteins to obtain a high-resolution cryo-EM map of these domains at 3.2 Å, allowing us
91 to build an atomic model (**Fig. S2**). The second approach used focused classifications of
92 RXFP1's ectodomain to obtain a cryo-EM map of the entire complex (**Fig. S3**). As a result of the

93 ectodomain's flexibility, the cryo-EM map of the full-length receptor is lower resolution, with an
94 overall resolution of 4.2 Å and local resolution of the ectodomain between 5-8 Å.
95 RXFP1's 7TM domain displays characteristic hallmarks of the active state for class A
96 GPCRs (**Fig. 1**). Most notable is the outward conformation of the intracellular end of
97 transmembrane helix 6 (TM6), which creates the binding site for the α5 helix of $G_{\alpha s}$[17,18].
98 Additional active-state features include the hydrogen bond between Tyr681$^{7.53}$ of the conserved
99 NPxxY motif in TM7 and Tyr599$^{5.58}$ in TM5, and an open conformation of the "ionic lock"
100 between Glu623$^{6.30}$ and Lys510$^{3.50;18}$ (superscript indicates Ballesteros-Weinstein numbering
101 system)[19]. In RXFP1–G_s, the active state of the 7TMs displays a canonical interaction with
102 heterotrimeric G_s, similar to previously reported GPCR–G protein structures[20].
103
104

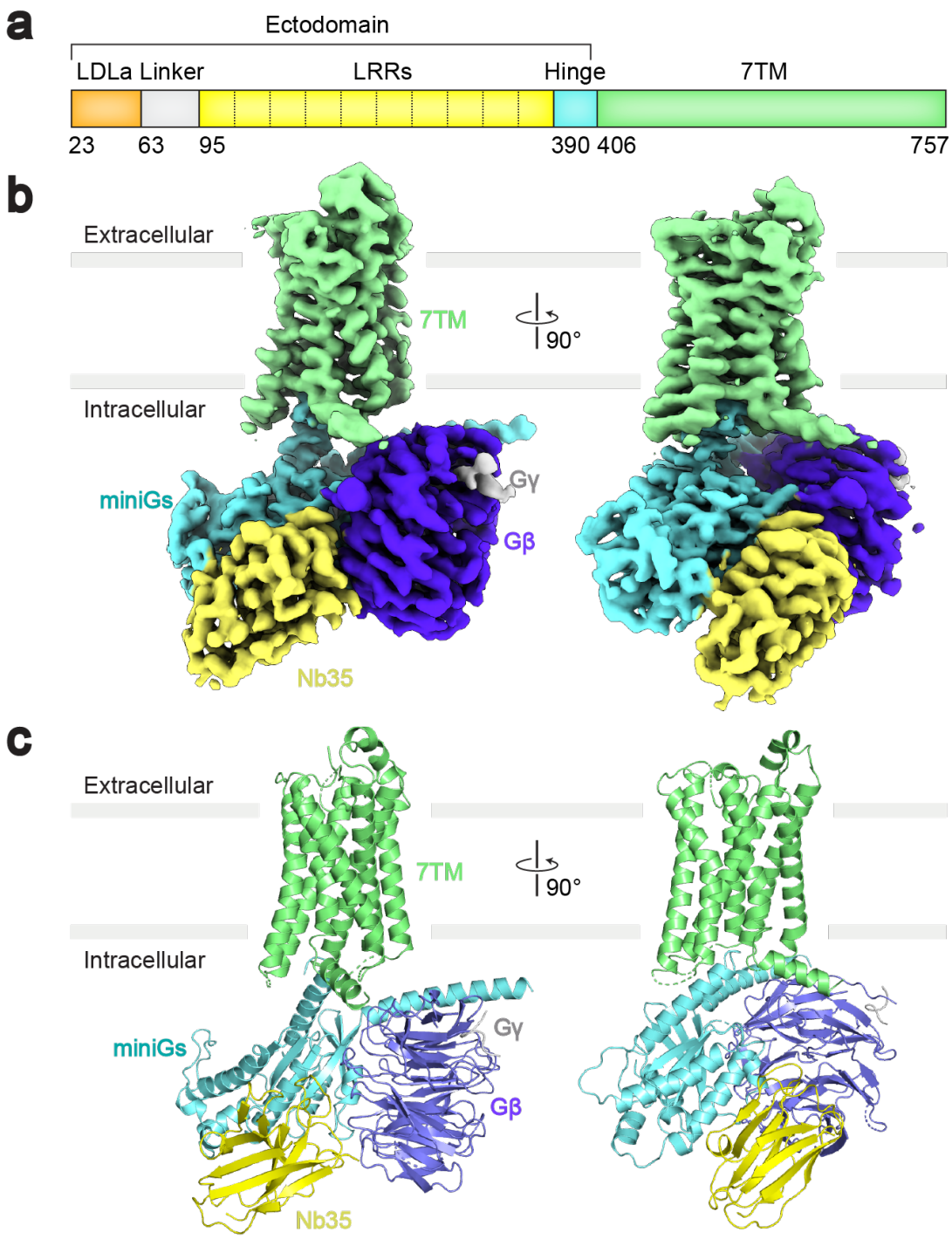

Fig. 1 | Cryo-EM map and model of the RXFP1–G_s complex. a, Diagram of the primary structure of RXFP1 domains. **b-c**, Cryo-EM map (**b**) and model (**c**) of the RXFP1–G_s complex 7TM domain with heterotrimeric G_S proteins and Nb35.

110 **Extracellular loop 2 is essential for signaling**

111 The active-state structure of RXFP1 surprisingly revealed that ECL2 occupies the GPCR
112 orthosteric ligand-binding pocket (**Fig. 2a**). The conformation of ECL2 was heterogeneous, and
113 three-dimensional focused classifications of RXFP1's extracellular loops were required to obtain
114 a cryo-EM map suitable for model building. The structure of ECL2 can be described in three
115 segments. The first segment of ECL2, from Lys550^{ECL2} to Gly558^{ECL2}, interacts with TMs 3, 4,
116 and 5, with the residues Tyr556^{ECL2} and Tyr557^{ECL2} making the most extensive contacts within
117 the 7TMs. The second segment of ECL2, beginning at Thr559^{ECL2} until His567^{ECL2}, forms a loop
118 structure which binds into the canonical GPCR orthosteric binding site. In particular, the side
119 chains of the residues Phe564^{ECL2} and Leu566^{ECL2} fit into a hydrophobic cavity created by TMs
120 2, 3, 5, 6, and 7 (**Fig. 2b**). Within this segment, Cys563^{ECL2} forms a disulfide bond with
121 Cys485$^{3.25}$ in TM3, a highly conserved feature among family A GPCRs which stabilizes the
122 conformation of the loop[21]. As a result of the disulfide bond and interactions with the
123 hydrophobic cavity, the second segment is the most well-resolved region of ECL2. In contrast,
124 four residues in the third segment of ECL2, from Ser568^{ECL2} to Ser573^{ECL2}, are not visible in the
125 cryo-EM map, likely indicating a region of higher intrinsic flexibility.

126 The unusual conformation of ECL2 and deeply buried positions of Phe564^{ECL2} and
127 Leu566^{ECL2} suggested that they may mimic the role played by exogenous agonists in other
128 receptors. Indeed, the location of the residues Phe564^{ECL2} and Leu566^{ECL2} corresponds with the
129 binding sites of small molecule and peptide orthosteric agonists of other family A GPCRs, such
130 as adrenaline binding to the β_2 adrenergic receptor[22] and angiotensin II analogs to the angiotensin
131 II type I receptor[23] (**Fig. S4b,c**). Based on these observations, residues from the second segment
132 of ECL2 were tested by mutagenesis for their contribution to signaling. ECL2 substitutions
133 maintained above 50% of wild type RXFP1 expression, excepting the mutation of Leu566^{ECL2} to
134 Asp, which expressed at 33% of wild-type levels (**Fig. S5a**). Mutation of Phe564^{ECL2} to Ala or
135 Leu566^{ECL2} to Asp almost completely ablated RXFP1 signaling in response to relaxin-2,
136 confirming the importance of these residues to receptor activation (**Fig. 2c**).

137 The discovery of the importance of ECL2 to RXFP1 signaling is reminiscent of the
138 orphan receptor GPR52, which is directly activated by its own ECL2 as a tethered agonist[24].
139 While ECL2 shows no conservation of sequence or detailed structure between RXFP1 and
140 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.

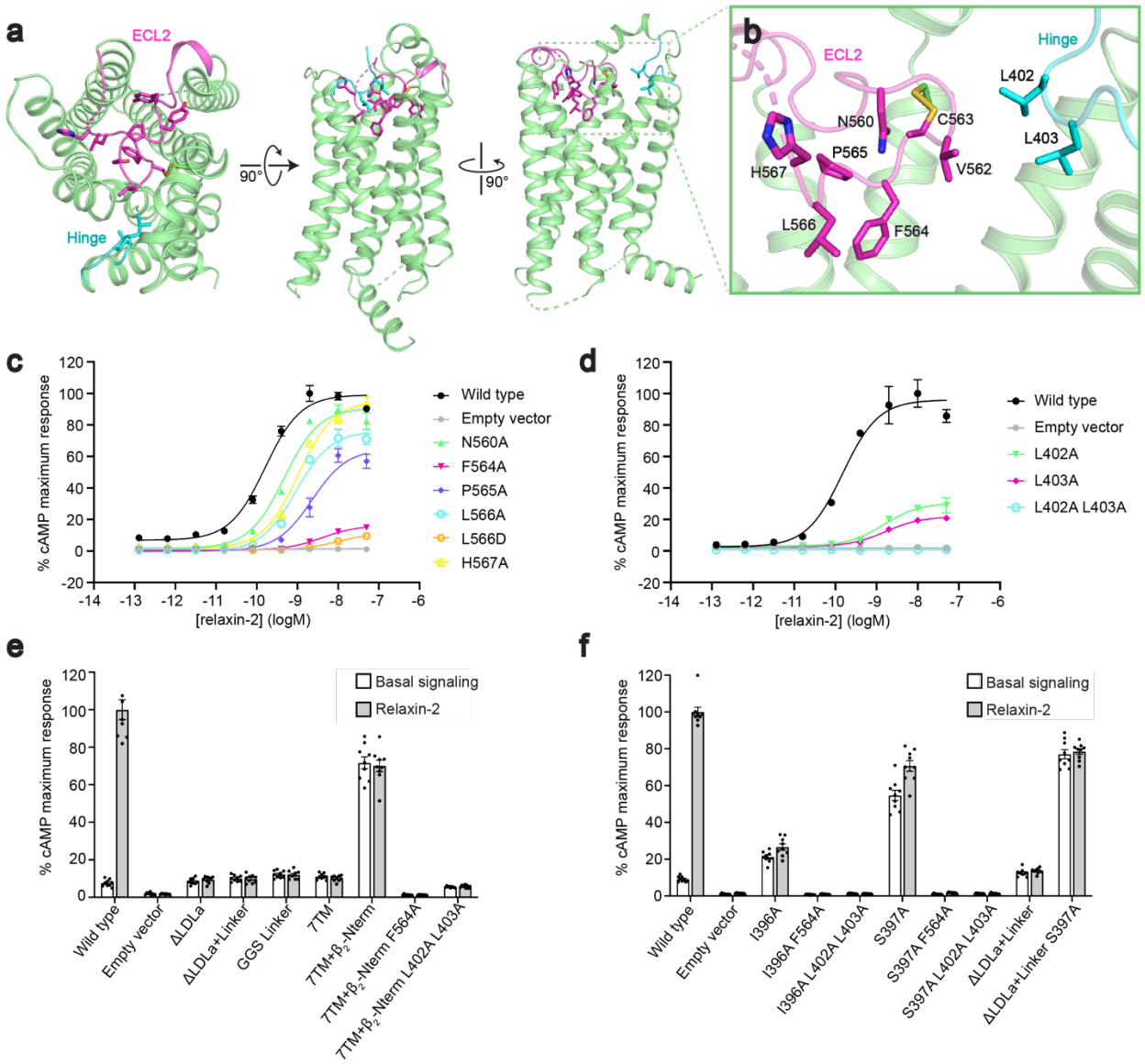

Fig. 2 | Regulation of receptor signaling by ECL2 and the ectodomain. a, The conformation of ECL2 and the hinge region in active-state RXFP1. **b**, Details of ECL2 in the 7TM orthosteric site and interactions between ECL2 and the hinge region. **c-d**, The effect of ECL2 (**c**) and Leu402 and Leu403 hinge region (**d**) mutations in an assay for G_s signaling by RXFP1. Data are mean ± 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 ± s.e.m. from nine technical replicates.

153 **The role of the hinge region in receptor activation**

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

160 Functional analyses of type A LGRs, the glycoprotein hormone receptors (GPHRs), have
161 previously established that the hinge region of those LGRs is critical for receptor signaling[25,26].
162 The unstructured loop in the hinge region of GPHRs is approximately 60-125 residues long and
163 plays a role in binding the receptors' glycoprotein hormone agonists[27]. For GPHRs, a 10-residue
164 section of their hinge region near the 7TM domain, P10, is also critical for receptor signaling[25,26]
165 and has been recently shown to adopt different conformations in inactive and active receptor
166 states of LHCGR[14]. In contrast, the hinge loop of type C LGRs, such as RXFP1, is predicted to
167 be about 15 residues long, has no established function, and shows no sequence conservation in
168 comparison to type A LGRs[1]. Despite the clear differences between the hinge regions of GPHRs
169 and RXFP1, the interactions between Leu402, Leu403, and ECL2 in RXFP1's active state
170 structure unexpectedly suggested that these residues may also play a role in signaling.

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

179

180 **Mechanism of RXFP1 7TM autoactivation by ECL2**

181 We performed enhanced-sampling molecular dynamics (MD) simulations (see Methods
182 for details) to study the role of ECL2 in the RXFP1 7TM basal activity. An inactive-state model
183 was first obtained by MD simulations of RXFP1 7TM deactivation, starting from a model based

184 on the cryo-EM structure. In brief, a sodium ion was placed in the sodium-binding site to favor
185 the sampling of inactive states during the simulations. We obtained an inactive state that was
186 strikingly similar to the inactive-state AlphaFold2 model[28], in which the intracellular end of TM6
187 bent toward TM3 to form the classic Lys510$^{3.50}$–Glu615$^{6.30}$ ionic lock (**Fig. S6a-c**). The second
188 segment of ECL2 maintained the same conformation throughout the simulations, whereas
189 Ser568^{ECL2} to Ser573^{ECL2} in the third segment were highly mobile, consistent with absence of
190 clear density in the cryo-EM map. Starting from the inactive state, we simulated autoactivation
191 of the RXFP1 7TMs by removing the sodium ion and protonating the sodium anchor, Asp451$^{2.50}$.
192 As negative controls, we performed the same simulations with two mutants with low basal
193 activity, F564^{ECL2}A and L566^{ECL2}D. The wild type RXFP1 7TM exhibited autoactivation with
194 outward movements of TM6 on the intracellular side, destabilization of the ionic lock, and
195 frequent side-chain flips of the toggle switch W633$^{6.48}$ (**Fig. S6d-f**). In contrast, the two mutants
196 remained in inactive conformations. The wild type receptor exhibited distinct shapes of the
197 orthosteric pocket compared to the two mutants, owing to ECL2–TM7 interactions. Namely,
198 Phe564^{ECL2} in the wild type 7TM domain "pushed" TM7 toward TM2, which likely altered
199 W633$^{6.48}$ conformations and triggered autoactivation (**Fig. S6d-f**). The F564^{ECL2}A mutation
200 directly eliminates this steric effect, whereas substituting Leu566^{ECL2} with Asp reorients the
201 charged side chain away from the pocket, leaving space for F564^{ECL2} and diminishing its impact
202 on TM7 conformation (**Fig. S6d,e**).
203
204 **Autoinhibition of RXFP1 signaling**
205 While structures of RXFP1 and GPR52 revealed that ECL2 is critical for the activation of
206 both receptors, other aspects of their signaling suggest differing mechanisms. GPR52 is a self-
207 activated orphan GPCR with very high basal activity, signaling at 90% of its E_{max} without any
208 agonist bound[24,29]. The mechanism governing the intrinsic activity of GPR52 is clear, as the
209 ECL2 tethered agonist is a component of the 7TM structure itself. In contrast, RXFP1 does not
210 have high basal activity, but signals in response to the binding of relaxin-2 to its LRRs. These
211 differences suggest that while both GPCR structures show ECL2 binding in the orthosteric site,
212 RXFP1 likely uses additional mechanisms to prevent continuous self-activation of the 7TM
213 domain. The major structural difference between these two receptors is that GPR52 is a
214 conventional family A GPCR with an unstructured N-terminus, while RXFP1 is a type C LGR,

215 containing a structured ectodomain of LRRs and an LDLa module. For this reason, we set out to
216 investigate the role of RXFP1's ectodomain in modulating the activity of ECL2.
217 To address this question, we first cloned constructs of RXFP1 with deletions of the
218 receptor's ectodomain. Basal signaling of RXFP1 was not increased by deletion of the LDLa
219 module, deletion of the LDLa module and the 32-residue linker that connects it to the LRRs, or
220 replacement of the linker with a 32 residue Gly-Gly-Ser linker. These results indicated that the
221 LDLa module and linker of RXFP1's ectodomain do not play an inhibitory role in signaling. In
222 contrast, these constructs are unable to signal in response to relaxin-2 binding, consistent with
223 previous studies showing that the LDLa module and linker region are essential for receptor
224 activation[13,30]. We next focused on testing the role of RXFP1's LRRs in the receptor signaling
225 mechanism. To remove the ectodomain, including the LRRs, constructs were designed to express
226 RXFP1's 7TM domain alone. However, these constructs showed very low cell-surface
227 expression, at 12% of wild type. Although the 7TM domain expressed very poorly, it showed
228 similar basal signaling to wild type RXFP1 in a G_s signaling assay, at 11% of the wild type
229 relaxin-2 E_{max}. To increase expression of the 7TMs, we cloned a fusion of RXFP1's 7TM
230 domain to the unstructured N-terminus of the high-expressing β_2 adrenergic receptor. This fusion
231 rescued 7TM domain expression to essentially wild type levels (**Fig. S5c**). When tested in a G_s
232 signaling assay, the fusion showed a high level of basal activity. In the absence of ligand, the
233 7TM domain signaled at 70% of the maximum level of agonist-induced signaling for wild-type
234 RXFP1 (**Fig. 2e**). As expected, RXFP1's 7TM domain alone does not show any change in
235 activity in response to relaxin-2, since the relaxin-2 binding site in the LRRs is deleted in this
236 construct[31].
237 To establish whether the high basal activity of RXFP1's 7TM alone is due to ECL2, we
238 introduced the Phe564^{ECL2} to Ala mutation that greatly reduced full-length RXFP1 signaling in
239 response to relaxin-2. The Phe564^{ECL2} to Ala mutation was able to completely ablate the 7TM
240 domain's high basal activity, confirming that ECL2 constitutively activates the 7TMs in the
241 absence of RXFP1's LRR domain. Likewise, Ala mutations of the hinge region residues Leu402
242 and Leu403 (which are included in the 7TM domain fusion construct) were also able to reduce
243 the high basal signaling. These data suggested a model of signaling in which the LRRs play an
244 autoinhibitory role in regulating the active state of ECL2 and the hinge region. As a result,
245 deletions of the LRRs allow constitutive activation the receptor, leading to high basal signaling.

246 **Inhibitory interactions revealed by evolutionary coupling analysis**

247 An additional insight into the regulation of ECL2 by the ectodomain arose from
248 evolutionary coupling (EC) analysis of RXFP1[32,33]. The strongest ECs, or evolutionary coupled
249 residues, are derived from applying a global probability model to multiple sequence alignments
250 and typically indicate residue pairs that are in contact in 3D[34], including residues involved in
251 conformational changes[35,36]. Residues of ECL2 had multiple strong ECs pairing them with
252 residues within the 7TM helices, supporting our active-state structure. However, ECL2 also
253 showed ECs with residues from RXFP1's hinge region not present in our model (**Fig. S7**). In
254 total, two residues from the hinge region, Ile396 and Ser397, had ECs with three residues of
255 ECL2, including the critical Phe564^{ECL2}. ECL2 and the hinge region are likely involved in other
256 interactions than those observed in our structure of the active state. If ECL2 contacts Ile396 and
257 Ser397 in an inactive-state conformation, we predicted that those interactions may contribute to
258 the inhibition of ECL2 in the absence of relaxin-2 binding to the LRRs.

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

270

271 **Relaxin-2 binding to the leucine-rich repeats**

272 The cryo-EM map of the full-length RXFP1–G_s complex is limited by lower resolution
273 due to the dynamic nature of the receptor's ectodomain. Continuous heterogeneity present in the
274 final particle stack for RXFP1–G_s was visualized through three-dimensional variability analysis
275 in cryoSPARC[37]. The resulting movies showed that the LRRs are very flexible in the relaxin-2–
276 bound state, moving at the hinge region between the LRRs and 7TM domain. Despite these

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

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

287 The low-resolution map shows relaxin-2 bound to the concave side of the LRRs. To aid
288 with modeling the relaxin-2–LRR interaction, we used crosslinking mass spectrometry (CLMS)
289 with the RXFP1–G_s complex. Using the Extended-EDC approach with an EDDA crosslinker[40],
290 $Glu14^{B\text{-}chain}$ of relaxin-2 crosslinked with three residues on the LRRs, Glu206, Glu299, and
291 Glu351 (**Fig. 3c,d**). Several residues involved in relaxin-2 binding have been previously
292 characterized through mutational analysis in radioligand binding assays. In those experiments, a
293 series of conserved residues on the B-chain of relaxin-2, Arg13, Arg17, and Ile/Val20, a motif
294 known as the relaxin binding cassette, were proposed to be part of the interface[41,42]. Multiple
295 residues on the concave side of the LRRs were also found to be important for relaxin-2
296 binding[31]. Interestingly, the $Glu14^{B\text{-}chain}$ residue of relaxin-2 identified by CLMS is adjacent to
297 the $Arg13^{B\text{-}chain}$ of the relaxin binding cassette. Additionally, the crosslinked Glu299 residue on
298 the LRRs had been previously identified by mutational studies, highlighting the close agreement
299 between our CLMS data and prior functional analyses.

300 The residues from our CLMS experiment were used in combination with the mutational
301 data as restraints for docking the relaxin-2–LRR interaction in HADDOCK[43] (**Fig. 3c,d**). The
302 highest scoring HADDOCK models agreed well with our low resolution cryo-EM map and
303 showed the B-chain of relaxin-2 bound to the concave side of the LRRs, while the A-chain made
304 limited contacts (**Fig. 3e**). Additional density in the cryo-EM map is present next to the A-chain,
305 which may belong to the ectodomain's linker region (**Fig. S8f**), which the A-chain has been
306 proposed to bind[30]. However, the low resolution and absence of crosslinks for these domains
307 prevented further characterization of A-chain interactions.

308 The docked model of relaxin-2 bound to the LRRs predicted that one of the relaxin
309 binding cassette residues, Arg17$^{B\text{-}chain}$, interacts with Glu206, a residue on the LRRs identified by
310 CLMS that was not previously known to be involved in the binding site. To verify this
311 interaction, we mutated Glu206 to Ala and tested the effect of the mutation on binding of an Fc-
312 tagged relaxin-2 protein (SE301; see Methods). The Glu206 to Ala mutant expressed at
313 equivalent levels to wild type receptor, but the single mutant reduced relaxin-2 binding to 42% of
314 wild type levels (**Fig. 3f,g**), validating the proposed interaction.
315
316

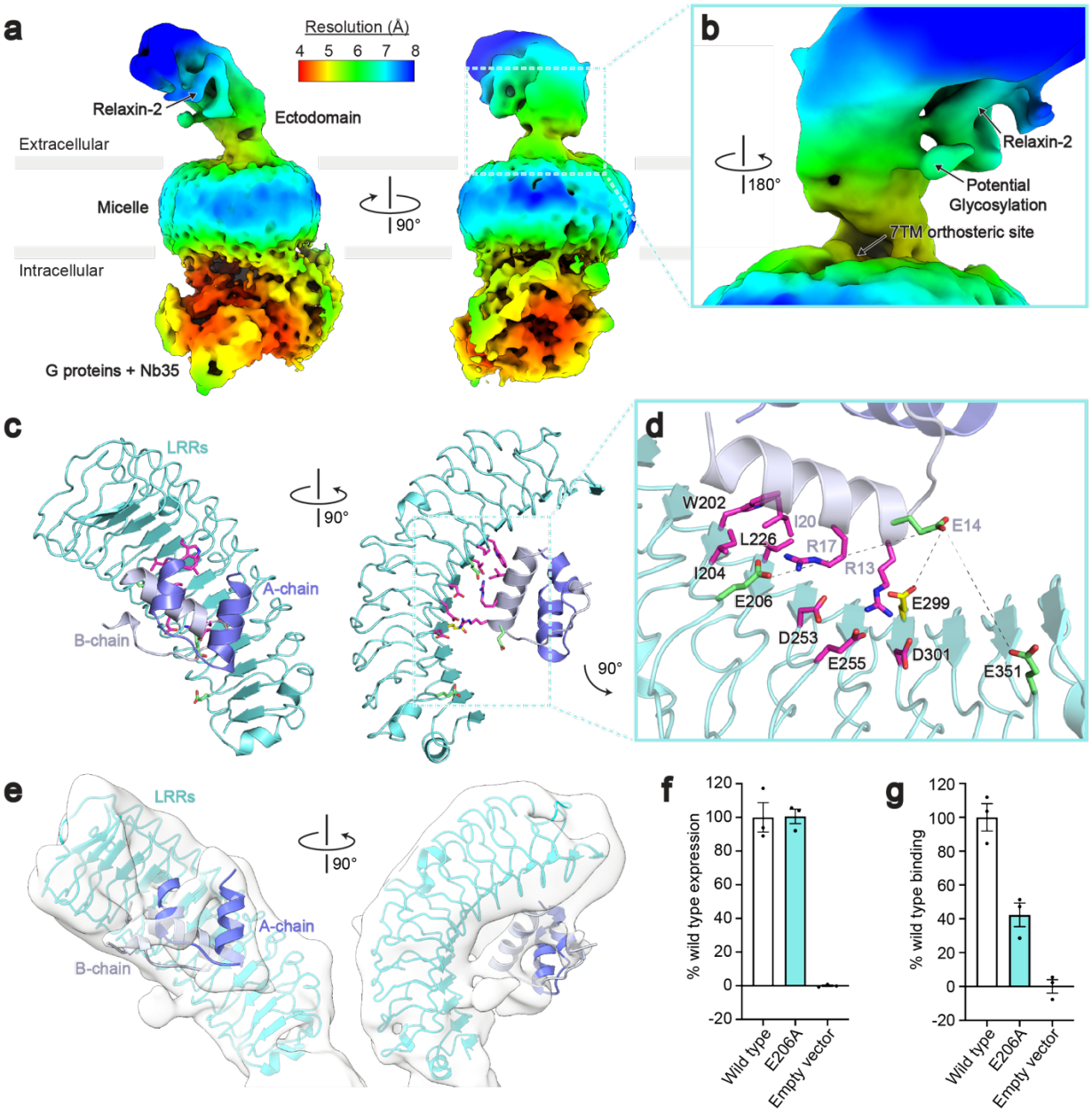

Fig. 3 | Cryo-EM and crosslinking mass spectrometry reveal interactions between relaxin-2 and the leucine-rich repeats. a, Local resolution cryo-EM map of the full-length RXFP1–G_s complex. **b**, The relaxin-2 binding site is above and rotated away from the 7TM orthosteric site. **c**, Model of the 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: $Glu14^{B\text{-}chain}$–Glu206 = 14.6 Å, $Glu14^{B\text{-}chain}$–Glu299 = 10 Å, $Glu14^{B\text{-}chain}$–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.

328 **Discussion**

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

338 Based on these results, it appears that relaxin-2's mechanism of RXFP1 signaling
339 involves indirectly releasing the autoinhibition of ECL2 (**Fig. 4**). As the LDLa module of the
340 ectodomain is essential for RXFP1 signaling by relaxin-2[13], it must play a role in signal
341 transduction between relaxin-2 binding to the LRRs and transition of the ECL2/hinge interface to
342 the active state. NMR studies with soluble constructs of RXFP1's ECLs showed that the LDLa
343 module can interact with ECL2[30,39,44]. However, the LDLa module was not resolved in our map,
344 indicating that it is likely mobile in the state that was captured in our cryo-EM studies.

345 An inactive-state structure will be necessary to define structural changes between
346 receptor states, such as conformational changes of the LRR domain. Nonetheless, our results
347 show that RXFP1 is not controlled by the steric "push-pull" mechanism proposed for GPHRs[14],
348 since the small size of relaxin-2 compared to glycoprotein hormones is incompatible with direct
349 steric activation of the receptor. Our better understanding of RXFP1's activation mechanism
350 suggests small molecule or antibody agonists could be created to either mimic ECL2-induced
351 activation or relieve inhibitory interactions between the ectodomain and ECL2. Such molecules
352 could be useful therapeutics for the treatment of numerous cardiovascular and fibrotic diseases.

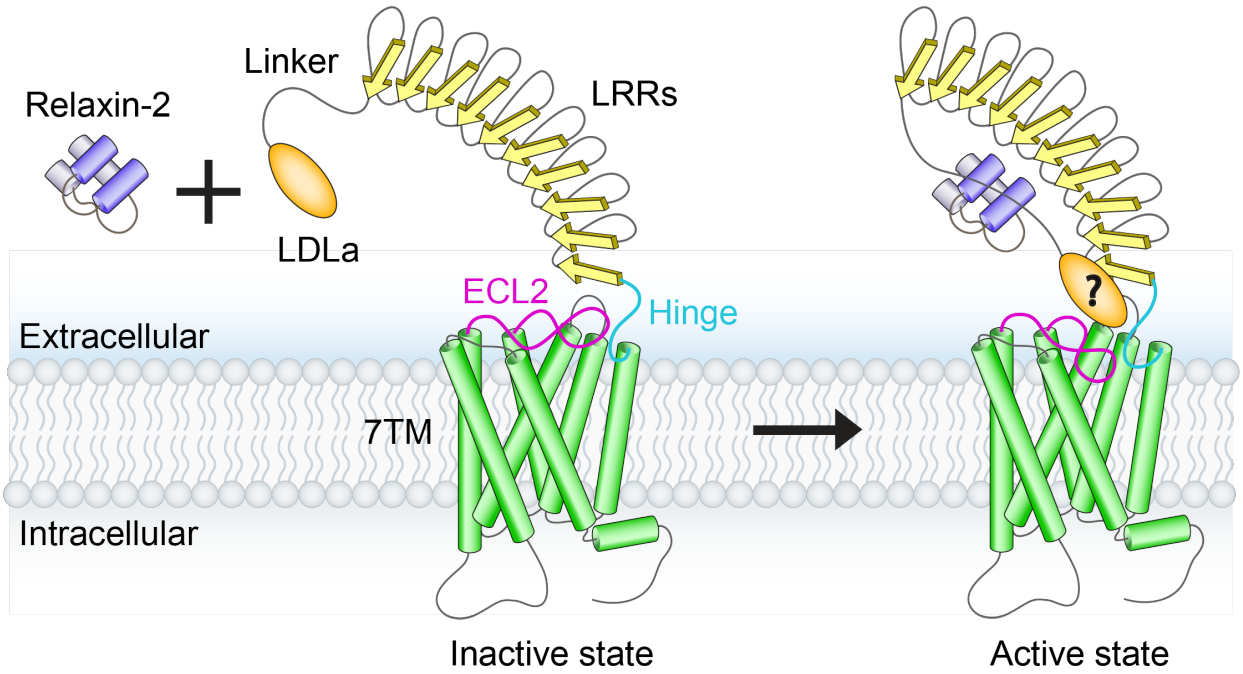

Fig. 4 | Model of RXFP1 activation by relaxin-2. The inactive state of RXFP1 is characterized by inhibitory interactions between ECL2, the hinge region, and the LRRs that prevent ECL2 from binding in the orthosteric pocket. Relaxin-2 binds to the concave side of the LRRs, away from the 7TM domain, also 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 the 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.

393 **Methods**

394 Cloning of RXFP1 constructs

395 Residues 23-757 of human RXFP1 were cloned with an N-terminal hemagglutinin signal
396 sequence, FLAG tag, and 3C protease site into the pcDNA-Zeo-tetO vector[45]. Fusions to the
397 miniG_s-399 protein with C-terminal truncations to RXFP1 were constructed using PCR followed
398 by NEBuilder HiFi DNA Assembly (New England Biolabs). Truncations removed 10, 15, 20,
399 25, 30, or 35 residues from the receptor C-terminus. For signaling assays, human RXFP1
400 residues 23-757 with an N-terminal hemagglutinin signal sequence and FLAG tag were cloned
401 into pcDNA-Zeo-tetO. Mutations to ECL2 or hinge region residues were introduced using
402 Quikchange Lightning PCR (Agilent). Ectodomain truncations were constructed using PCR and
403 NEBuilder HiFi DNA Assembly. Residue numbering is based on the canonical RXFP1 sequence
404 beginning at the initiating Met residue (UniProt ID Q9HBX9).

405

406 Cell surface expression tests

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

424 sample were collected. Data was normalized using the wild type and empty vector mean
425 fluorescence intensities as 100% and 0%, respectively, and plotted using GraphPad Prism.
426
427 cAMP signaling assay
428 The GloSensor assay from Promega, a live-cell signaling assay that detects cellular
429 cAMP levels, was used to measure activation of G_s signaling through RXFP1. White, clear-
430 bottom 96-well plates (Thermo Fisher Scientific) were coated with 30 µL of 10 µg/mL poly-D-
431 lysine (Sigma-Aldrich), and washed once with PBS. HEK293T cells were then plated at $2x10^4$
432 cells/well in DMEM with 10% (v/v) FBS. The following day, cells were transfected with 20 ng
433 of GloSensor reporter plasmid and 20 ng (unless otherwise stated) of human RXFP1 or empty
434 vector DNA per well using FuGENE, according to the manufacturer's instructions. For **Fig. 2d**,
435 0.4 ng of wild type RXFP1 DNA was used per well, plus 19.6 ng of empty vector DNA,
436 according to ratios determined during cell surface expression tests. For **Fig. 2e**, 10 ng of 7TM +
437 β_2-Nterm DNA was used per well, plus 10 ng of empty vector DNA, according to ratios
438 determined during cell surface expression tests. Twenty-four hours later, the media was replaced
439 with 40 µL of CO_2-independent media (Thermo Fisher Scientific) with 10% (v/v) FBS and 2
440 mg/mL D-luciferin (Goldbio) and incubated for 2 hours at room temperature (RT) in the dark.
441 Measurements of luminescence with 1 second integration times were taken before ligand
442 addition using a SpectraMax M5 microplate reader. For signaling curves, dilution series of native
443 relaxin-2 (R&D Systems), were added to the cells and luminescence measurements were taken
444 30 minutes after relaxin-2 addition. For measurements of RXFP1 basal signaling versus relaxin-2
445 signaling, vehicle control (PBS + 0.1% bovine serum albumin) or relaxin-2 at 50 nM final
446 concentration were added to the cells and the luminescence measurement was taken after 30
447 minutes. The maximum signaling response of relaxin-2 at RXFP1 for each experiment was
448 normalized to 100%, and the percentages were plotted using GraphPad Prism.
449
450 Expression of RXFP1-miniG_s
451 RXFP1-miniG_s constructs for protein purification were expressed in inducible Expi293F
452 tetR cells (Thermo Fischer Scientific). Stable cells lines were generated for RXFP1-miniG_s399-
453 20res by plating Expi293F tetR cells adherently in DMEM with 10% (v/v) FBS. Cells were
454 transfected with linearized RXFP1-miniG_s399-20res pcDNA-Zeo-tetO DNA using

455 Lipofectamine, according to the manufacturer's protocols (Thermo Fischer Scientific). Stable
456 integrations were selected using 200 μg/mL Zeocin. After selection, polyclonal adherent RXFP1-
457 miniG$_s$399-20res cells were readapted to suspension culture in antibiotic-free Expi293 media
458 (Thermo Fischer Scientific), then maintained in Expi293 media with 10 μg/mL Zeocin. Stable
459 RXFP1-miniG$_s$399-20res Expi293F tetR cells were expanded for protein expression in
460 antibiotic-free Expi293 media and induced with 4 μg/mL doxycycline, 0.4% glucose, and 5 mM
461 sodium butyrate. After 48 hours of induction, cells were harvested by spinning at 4,000 xg for 30
462 minutes at 4°C. The pellet was flash frozen in liquid nitrogen and stored at -80°C.
463
464 Purification of relaxin proteins:
465 To generate relaxin-2 proteins for RXFP1 complex purifications and flow cytometry
466 binding assays, single-polypeptide relaxin-2 constructs were cloned with linkers connecting the
467 B-chain and A-chain. Sequences for these proteins can be found in **Table S5**. Briefly, a His-
468 tagged single-chain relaxin-2 (SE001) was expressed as a secreted protein from inducible
469 Expi293F tetR cells. SE001 was purified from cell supernatants after 5 days of induction using
470 Nickel Excel resin (GE Healthcare), followed by size exclusion chromatography with a
471 Sephadex S200 column (GE Healthcare) in 30 mM MES pH 6.5 and 300 mM sodium chloride.
472 A single-chain relaxin-2 fused to an antibody IgG1 Fc fragment (SE301) was expressed as
473 described above in Expi293F tetR cells. SE301 was purified using Protein G resin (GE
474 Healthcare) in 20 mM HEPES pH 7.5, 150 mM sodium chloride.
475
476 Purification of the RXFP1–G$_s$ complex
477 Purification of the RXFP1–G$_s$ complex used a cell pellet from a 2 L induction of the
478 RXFP1-miniG$_s$399-20res Expi293F tetR stable cell line, and the purified proteins SE001 (single-
479 chain relaxin-2), human G$\beta_1\gamma_2$, and Nb35-His-PrC. G$\beta_1\gamma_2$ and Nb35-His-PrC were expressed and
480 purified according to previously published protocols[16]. The RXFP1-miniG$_s$399-20res cell pellet
481 was lysed through osmotic shock by stirring in 250 mL cold Lysis buffer containing 20 mM
482 HEPES pH 7.5, 2 mM magnesium chloride, 1 μL benzonase (Sigma Aldrich), 1 protease
483 inhibitor tablet (Thermo Fisher Scientific), and 100 nM SE001. Once stirring, iodoacetamide was
484 added at a final concentration of 2 mg/mL. Lysed cells were centrifuged at 50,000 xg for 30
485 minutes at 4°C and the supernatant was decanted. Membranes were homogenized with a glass

486 dounce (Thermo Fisher Scientific) in 270 mL Complexation buffer containing 20 mM HEPES
487 pH 7.5, 350 mM sodium chloride, 20% (v/v) glycerol, 1 μL benzonase, and 1 protease inhibitor
488 tablet. After homogenization, 960 μg $G\beta_1\gamma_2$, 740 μg Nb35 (1:2:4 ratio of RXFP1:$G\beta_1\gamma_2$:Nb35),
489 100 nM SE001, and 0.2 U apyrase (New England Biolabs) were added to the membranes and the
490 solution was stirred for 1 hour at 4°C. Next, 30 mL of 10X Detergent buffer [10% (w/v) lauryl
491 maltose neopentyl glycol (L-MNG; Anatrace) and 1% (w/v) cholesterol hemisuccinate (CHS;
492 Anatrace)] were added to the solution, dounce homogenization was repeated, and the solution
493 was stirred for 2 hours at 4°C to extract the RXFP1–G_s complex from the membrane. After 2
494 hours, the solution was centrifuged at 50,000 xg for 30 minutes at 4°C and the supernatant was
495 filtered with a glass fiber prefilter (Millipore). Calcium was added to the solubilized membranes
496 at a final concentration of 2 mM and the solution was loaded by gravity flow over 3 mL M1 anti-
497 FLAG resin (In house). This step of the purification was based on the N-terminal FLAG tag of
498 RXFP1-miniG_s399-20res. The M1 resin was equilibrated with Wash buffer 1 [0.1% (w/v) L-
499 MNG, 0.01% (w/v) CHS, 350 mM sodium chloride, 20 mM HEPES pH 7.5, 2 mM calcium
500 chloride, and 100 nM SE001]. After loading the solubilized membrane fraction, the column was
501 subsequently washed with 50 mL each of Wash buffer 1, Wash buffer 2 [0.1% (w/v) L-MNG,
502 0.01% (w/v) CHS, 350 mM sodium chloride, 20 mM HEPES pH 7.5, 2 mM calcium chloride, 2
503 mM magnesium chloride, 5 mM adenosine 5'-triphosphate magnesium salt, and 100 nM SE001],
504 and Wash buffer 3 [0.01% (w/v) L-MNG, 0.001% (w/v) CHS, 350 mM sodium chloride, 20 mM
505 HEPES pH 7.5, 2 mM calcium chloride, and 100 nM SE001]. M1 resin was eluted with 0.01%
506 (w/v) L-MNG, 0.001% (w/v) CHS, 350 mM sodium chloride, 20 mM HEPES pH 7.5, 0.5
507 mg/mL FLAG peptide (GenScript), and 100 nM SE001, then 2 mM calcium chloride was added
508 to the eluted protein. Next, 0.5 mL anti-Protein C resin (In house) was equilibrated with Protein
509 C wash buffer [0.005% (w/v) L-MNG, 0.0005% (w/v) CHS, 350 mM sodium chloride, 20 mM
510 HEPES pH 7.5, 2 mM calcium chloride, and 100 nM SE001]. The M1 elution with 2 mM
511 calcium chloride was loaded onto the anti-Protein C resin by gravity flow, washed with 10 mL
512 Protein C wash buffer, and eluted with 0.005% (w/v) L-MNG, 0.0005% (w/v) CHS, 350 mM
513 sodium chloride, 20 mM HEPES pH 7.5, 0.5 mg/mL Protein C peptide (GenScript), and 100 nM
514 SE001. This step of the purification was based on the Protein C tag of Nb35-His-PrC. The
515 elution was concentrated with a 3 kDa molecular weight cutoff centrifugal concentrator
516 (Millipore) and loaded onto a Sephadex S200 column (GE Healthcare) in SEC buffer [0.005%

517 (w/v) L-MNG, 0.0005% (w/v) CHS, 350 mM sodium chloride, 20 mM HEPES pH 7.5, and 100
518 nM SE001]. Peak fractions from size exclusion were concentrated with a 3 kDa molecular
519 weight cutoff centrifugal concentrator and either stored overnight at 4°C prior to grid freezing or
520 immediately used to freeze cryo-EM grids.
521
522 Cryo-EM grid preparation and data collection
523 Cryo-EM grids were prepared using QUANTIFOIL® holey carbon grids (400-mesh,
524 copper, R1.2/1.3; Electron Microscopy Sciences). Grids were washed with ethyl acetate (Sigma
525 Aldrich) and then glow-discharged at -20 mA for 60 seconds with a Pelco Easiglow. 3 µL of
526 sample at 0.2 to 0.4 mg/mL was applied to the grids. Grids were plunge-frozen in liquid ethane
527 using a Mark IV Vitrobot (Thermo Fisher Scientific) at 10°C and 100% humidity with a 10
528 second wait time, 3 to 4 second blot time, and blot force of 15.
529 Cryo-EM data for the RXFP1–G_s complex were collected in four separate sessions. Grids
530 were imaged with a Titan Krios microscope (Thermo Fisher Scientific) at 300 kV with a Gatan
531 BioQuantum GIF/K3 direct electron detection camera in counting mode. Movies were collected
532 with 50 frames each at 81,000x, corresponding to 1.06 Å per pixel, and a total dose of around 52
533 electrons per $Å^2$. Defocus values ranged from -0.8 to -2.3 µm.
534
535 Cryo-EM data processing
536 Cryo-EM data was collected in four separate session and initially processed individually,
537 then particle stacks were merged to generate the final maps. Motion correction was carried out
538 with MOTIONCOR2[46] and CTF parameters were estimated with CTFFIND-4.1[47]. Particles were
539 picked in RELION 3.1[48] using Laplacian of Gaussian autopicking. A map from a previous small
540 dataset collected on a Talos Arctica microscope and processed in RELION was used as an initial
541 model. Particles were extracted with a box size large enough to include the entire complex (288
542 pixels). Multiple rounds of 2D classifications in RELION used a circular mask of 150 Å around
543 only the micelle and G proteins as the first step of particle sorting for both 3D reconstructions.
544 Masking of the micelle and G proteins was used to initially sort high quality particles from
545 contamination because 2D classifications attempted for the entire particle showed weak signal
546 for the ectodomain, likely indicating a region of higher flexibility. It became evident at this stage
547 of data processing that the datasets showed a preferred orientation for one of the side views of

548 the RXFP1–G_s complex. Processing 2D classifications of the preferred view separately from
549 other views of the complex resulted in a larger number of particles from different orientations in
550 the final particle stacks.
551 For the map of the 7TMs, masked 3D classifications of the 7TM domain bound to G
552 proteins were performed in RELION. Following these classifications, iterative 3D refinements
553 (using a mask of the micelle and G proteins) and Bayesian particle polishing steps were carried
554 out in RELION. At this stage, particle stacks from different datasets were joined together,
555 followed by additional 2D classifications with a 150 Å circular mask and masked 3D
556 classifications. Finally, 3D focused classifications without alignments were carried out using
557 masks for either the ECLs, TM helices, or helix8. These particle stacks were imported into
558 cryoSPARC v3.1.0[37] and used with Non-uniform Refinement (New). The cryoSPARC refined
559 maps were post-processed in DeepEMhancer[49] using the half-maps as input. This data processing
560 workflow is described in **Fig. S2**.
561 For the map of the full-length RXFP1–G_s complex, rounds of unmasked 3D
562 classifications were used in RELION after the 2D classifications described above. Next, the
563 micelle and G proteins were subtracted from the particles and 3D focused classifications without
564 alignments were carried out using a mask on the entire ectodomain of RXFP1. 3D refinements
565 with a mask of the entire complex and iterative Bayesian particle polishing steps were then
566 performed in RELION. Following particle polishing, particle stacks from different datasets were
567 combined and 3D classifications with a mask around the entire complex were performed in
568 RELION, followed by 3D refinement in RELION with a mask around the entire complex.
569 Finally, local resolution estimation and filtering were performed in RELION on the final maps.
570 This data processing workflow is described in **Fig. S3**.
571
572 Model building and refinement
573 A combined focused map for the RXFP1 7TM domain bound to G proteins was
574 generated using the DeepEMhancer post-processed maps as input in Phenix[50]. Model building
575 for the RXFP1 7TM domain bound to heterotrimeric G_s and Nb35 was performed using
576 DeepEMhancer focused and combined focused cryo-EM maps in *Coot*[51]. An initial model for the
577 RXFP1 7TM domain (Gly395 to Gly709) was generated using trRosetta[52] and manually fit into
578 the DeepEMhancer focused cryo-EM map (for ECLs) in Chimera[53]. Intracellular and

579 extracellular loops and the hinge region were removed from the trRosetta model and manually
580 rebuilt in *Coot*. ECL2 residues Lys554–Tyr557 and hinge region residues Glu400–Leu403 were
581 manually built with the coordinates of the human RXFP1 AlphaFold2 model[28] overlayed in
582 *Coot*. Initial models for miniG_s399, Gβγ, and Nb35 were generated in MODELLER[54] using
583 coordinates from PDB ID 6GDG[55]. All models were refined with Phenix real-space refinement[50]
584 using secondary structure restraints against the DeepEMhancer combined focused map. Statistics
585 for the final model were evaluated using MolProbity in a Phenix comprehensive validation
586 (cryo-EM) job that used the map from cryoSPARC v3.1.0 Non-uniform Refinement (New) and
587 the final model as inputs. Figures were prepared using PyMOL and ChimeraX[56]. Structural
588 biology programs used in this work, other than cryoSPARC, were compiled and configured by
589 SBGrid[57]. Refinement statistics are present in **Table S1** and representative images of cryo-EM
590 map and model quality are shown in **Fig. S8**.
591
592 Evolutionary coupling analysis
593 For comparing the constructed model to the strongest evolutionarily coupled pairs, we
594 first downloaded the Uniprot canonical sequence for Q9HBX9 for residues 405-689 and then
595 used the Jackhmmer software suite[58] to build multiple sequence alignments across multiple
596 bitscore thresholds based on the June 2019 download of the Uniref100 database[59]. We then chose
597 an alignment with 352,511 sequences, with 90% of columns consisting of fewer than 30% gaps,
598 to move forward with in analysis. We used the EVcouplings v0.0.5 software, available at
599 <github.com/debbiemarkslab>, to identify evolutionary couplings (ECs) for this alignment.
600 To incorporate the LRR region, we ran the Q9HBX9 sequence for residues 120-757 in
601 Jackhmmer, against the 02/2021 Uniref100 database for normalized bitscores between 0.1 and
602 0.9. We chose an alignment with 8.145 sequences with 86% of columns consisting of fewer than
603 30% gaps, and input it to v0.1.1 EVcouplings software. The strongest couplings were ranked
604 based on assigning probabilities from a logistic regression model new to v0.1.1. This version of
605 the pipeline was also used for plotting all contact maps.
606
607 Crosslinking mass spectrometry
608 The RXFP1–G_s complex used for CLMS was prepared as described above and
609 crosslinking reactions were carried out the following day, after storage overnight at 4°C. CLMS

610 was performed as previously described[40]. Briefly, crosslinking reactions were carried out for 1 h
611 at room temperature in 100 mM MOPS Buffer, pH 6.5 with 50 mM EDC, ~24 mM EDDA
612 linker, and 20 mM sulfo-NHS. Reactions were quenched with hydroxylamine to a final
613 concentration of 100 mM. Samples were reduced for 1 h in 2% SDS and 5 mM TCEP followed
614 by alkylation with 10 mM iodoacetamide in the dark for 30 min and quenching with 5 mM DTT
615 for 15 min. Samples were then processed with the SP3[60] method and digested with trypsin
616 (Promega) at 1:25 enzyme:substrate ratio overnight at 37°C. Digested peptides were acidified
617 with 10% formic acid to pH ~2 and desalted using stage tips with Empore C18 SPE Extraction
618 Disks (3M) and dried under vacuum.
619 Sample was reconstituted in 5% formic acid (FA)/5% acetonitrile and analyzed in the
620 Orbitrap Eclipse Mass Spectrometer (Thermo Fischer Scientific) coupled to an EASY-nLC 1200
621 (Thermo Fisher Scientific) ultra-high pressure liquid chromatography (UHPLC) pump, as well as
622 a high-Field Asymmetric waveform Ion Mobility Spectrometry (FAIMS) interface. Peptides
623 were separated on an in-house packed 100 μm inner diameter column packed with 35 cm of
624 Accucore C18 resin (2.6 um, 150 Å, ThermoFisher), using a gradient consisting of 5–35%
625 (ACN, 0.125% FA) over 135 min at ∼500 nL/min. The instrument was operated in data-
626 dependent mode. FTMS1 spectra were collected at a resolution of 120K, with an automated gain
627 control (AGC) target of 5×10^5, and a max injection time of 50 ms. The most intense ions were
628 selected for MS/MS for 1s in top-speed mode, while switching among three FAIMS
629 compensation voltages (CV): −40, –60, and –80 V in the same method. Precursors were filtered
630 according to charge state (allowed $3 <= z <= 7$), and monoisotopic peak assignment was turned
631 on. Previously interrogated precursors were excluded using a dynamic exclusion window (60 s ±
632 7 ppm). MS2 precursors were isolated with a quadrupole mass filter set to a width of 0.7 m/z and
633 analyzed by FTMS2, with the Orbitrap operating at 30K resolution, an AGC target of 100K, and
634 a maximum injection time of 150 ms. Precursors were then fragmented by high-energy collision
635 dissociation (HCD) at a 30% normalized collision energy.
636 Mass spectra were processed and searched using the PIXL search engine[40]. The sequence
637 database contained proteins identified at 1% FDR in non-cross-linked Comet[61] search. For PIXL
638 search, precursor tolerance was set to 15 ppm and fragment ion tolerance to 10 ppm. Methionine
639 oxidation was set as a variable modification in addition to mono-linked mass of +130.110613 for
640 EDDA. Crosslinked peptides were searched assuming zero-length (-18.010565) and EDDA

641 crosslinker +112.100048. Crosslinked searches considered 60 protein sequences to ensure
642 sufficient statistics for FDR estimation. Matches were filtered to 1% FDR on the unique peptide
643 level using linear discriminant features as previously described[40].
644
645 Docking
646 HADDOCK[43] was used to dock the relaxin-2–LRR interaction. Docking used the human
647 relaxin-2 X-ray crystal structure (PDB ID: 6RLX)[62] and a model of the LRRs from residues 104-
648 391 of the AlphaFold2[28] prediction for human RXFP1. Residues from CLMS studies that were
649 part of crosslinks between relaxin-2 and RXFP1 were used as active restraints in the docking run.
650 These CLMS residues were Glu14$^{B\text{-chain}}$ of relaxin-2, and Glu206, Glu299, and Glu351 of
651 RXFP1. Residues identified to be important for relaxin-2 binding from published mutations in
652 radioligand binding assays were also used as active restraints[31,41,42]. These residues included
653 Arg13$^{B\text{-chain}}$, Arg17$^{B\text{-chain}}$, and Ile20$^{B\text{-chain}}$ of relaxin-2 and Trp202, Ile204, Leu226, Asp253,
654 Glu255, Glu299, and Asp301 of RXFP1. The resulting docking models were analyzed according
655 to the HADDOCK scoring function and fit into the low resolution cryo-EM map of the RXFP1
656 ectodomain in ChimeraX[56].
657
658 Flow cytometry binding assay
659 A flow cytometry assay was used to measure the binding of an Fc-tagged relaxin-2
660 protein, SE301, to Expi293F cells transfected with human RXFP1 or empty pcDNA-Zeo-tetO
661 vector. Expi293F tetR cells were grown in Expi293 media and transfected using FectoPRO
662 (Polyplus), according to the manufacturer's protocols. The cells were enhanced 24 hours post-
663 transfection with 0.4% glucose and induced 48 hours post-transfection with 4 μg/mL
664 doxycycline and 5 mM sodium butyrate. After 24 hours of induction, cells were harvested by
665 spinning at 200 xg for 5 minutes at 4°C and washed once with HBS with 1% (v/v) FBS and 2
666 mM calcium chloride (Buffer B). Cells were plated into a V-bottom 96-well plate (Corning) at
667 100,000 cells/well and blocked by incubation in Buffer B for 30 minutes at 4°C. After blocking,
668 cells were centrifuged at 200 xg for 5 minutes at 4°C, resuspended in 100 μL of Buffer B
669 containing 500 nM SE301 or Fc N297Q, and incubated for 1 hour at 4°C. Cells were then
670 centrifuged at 200 xg for 5 minutes at 4°C, washed twice with 200 μL Buffer B, and resuspended
671 in 100 μL Buffer B containing 100 nM M1 anti-FLAG antibody labeled with Alexa Fluor 488

672 and Alexa Fluor 647 anti-human IgG Fc (BioLegend) diluted 1:100 (v/v). Cells were incubated
673 in secondary antibodies for 30 minutes at 4°C, washed once with 200 μL Buffer B, and
674 resuspended in 100 μL Buffer B for flow cytometry.
675 Samples were analyzed on a CytoFLEX flow cytometer (Beckman Coulter) and gated
676 according to plots of FSC-A/SCA-A, FSC-A/FSC-H, and receptor expression according to Alexa
677 Fluor 488 M1 anti-FLAG antibody binding. Approximately 500 events/sample were collected
678 from cells expressing receptor for human RXFP1-transfected cells or post-FSC-A/FSC-H gating
679 for empty vector-transfected cells. Receptor expression levels were analyzed using the Alexa
680 Fluor 488 M1 anti-FLAG mean fluorescence intensity (MFI) from post-FSC-A/FSC-H gated
681 cells. SE301 binding was analyzed by plotting the MFI for Alexa Fluor 647 anti-human IgG Fc
682 cells that had been gated for receptor expression. The data were plotted and analyzed in
683 GraphPad Prism.
684
685 Molecular dynamics
686 The initial model was built from the cryo-EM structure reported here. The missing
687 segments in ECL2, ICL3 and TM6 were generated by MODELLER v9.15[63]. PACKMOL-
688 Memgen[64] was used to assign the side-chain protonation states and embed the models in a
689 bilayer of POPC lipids. The system was solvated in a periodic 80 × 80 × 110 Å[64] box of explicit
690 water and neutralized with 0.15 M of Na^+ and Cl^- ions. We used the Amber ff14SB[65] and lipid
691 14[66] force fields, the TIP3P[67] water model and the Joung-Cheatham[68] parameters the ions. For
692 the simulations RXFP1 7TM deactivation, a Na^+ ion was placed at the conserved Na^+-binding
693 site (between Asp451$^{2.50}$ and Ser499$^{3.35}$). For the simulations of RXFP1 7TM autoactivation, the
694 truncated AlphaFold2[28] model was used to build the system and Asp451$^{2.50}$ was protonated.
695 After energy minimization, all-atom MD simulations were carried out using Gromacs
696 5.1[69] patched with the PLUMED 2.3 plugin[70]. The LINCS algorithm[71] was applied to constrain
697 bonds involving hydrogen atoms, allowing a time step of 2 fs. Each system was gradually heated
698 to 310 K and pre-equilibrated during 10 ns of brute-force MD in the *NPT*-ensemble. The replica
699 exchange with solute scaling (REST2)[72] technique was used to enhance the conformational
700 sampling. A total of 64 replicas of simulations were performed in the *NVT* ensemble. REST2 is a
701 type of Hamiltonian replica exchange simulation scheme. Besides the original simulation, many
702 replicas of the same system were simulated simultaneously. The additional replicas have

703 modified free energy surfaces, in which the energy barriers are easier to cross than in the original
704 simulation system. By frequently swapping the replicas and the original system during the MD,
705 the simulations "travel" on different free energy surfaces and easily visit various conformational
706 zones. Finally, only the samples on the original free energy surface are collected. The additional
707 replicas are artificial to ease barrier crossing, which are discarded after the simulations. REST2,
708 in particular, modifies the free energy surfaces by scaling (reducing) the force constants of the
709 "solute" molecules in the simulation system. In this case, the protein was considered as "solute"–
710 the force constants of its van der Waals, electrostatic and dihedral terms were subject to scaling–
711 in order to facilitate the conformational changes. The scaling factors were generated using the
712 Patriksson-van der Spoel approach[73] and effective temperatures ranging from 310 K to 1000 K.
713 Exchange between replicas was attempted every 1000 simulation steps. This setup resulted in an
714 average exchange probability of ~40%. We performed 80 ns × 64 replicas of MD in the *NVT*
715 ensemble for each system. The first 30 ns were discarded for equilibration. The simulation
716 trajectory on the original unscaled free energy surface was reassembled and analyzed.
717

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 β2 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).

764 20. Wang, J., Hua, T. & Liu, Z.-J. Structural features of activated GPCR signaling complexes.
765 *Curr. Opin. Struct. Biol.* **63**, 82–89 (2020).
766 21. Venkatakrishnan, A. J. *et al.* Molecular signatures of G-protein-coupled receptors. *Nature*
767 **494**, 185–194 (2013).
768 22. Ring, A. M. *et al.* Adrenaline-activated structure of β2-adrenoceptor stabilized by an
769 engineered nanobody. *Nature* **502**, 575–579 (2013).
770 23. Wingler, L. M., McMahon, C., Staus, D. P., Lefkowitz, R. J. & Kruse, A. C. Distinctive
771 activation mechanism for angiotensin receptor revealed by a synthetic nanobody. *Cell*
772 **176**, 479-490.e12 (2019).
773 24. Lin, X. *et al.* Structural basis of ligand recognition and self-activation of orphan GPR52.
774 *Nature* **579**, 152–157 (2020).
775 25. Brüser, A. *et al.* The activation mechanism of glycoprotein hormone receptors with
776 implications in the cause and therapy of endocrine diseases. *J. Biol. Chem.* **291**, 508–520
777 (2016).
778 26. Schulze, A. *et al.* The intramolecular agonist is obligate for activation of glycoprotein
779 hormone receptors. *FASEB J.* **34**, 11243–11256 (2020).
780 27. Jiang, X., Dias, J. A. & He, X. Structural biology of glycoprotein hormones and their
781 receptors: Insights to signaling. *Mol. Cell. Endocrinol.* **382**, 424–451 (2014).
782 28. Tunyasuvunakool, K. *et al.* Highly accurate protein structure prediction for the human
783 proteome. *Nature* **596**, 590–596 (2021).
784 29. Martin, A. L., Steurer, M. A. & Aronstam, R. S. Constitutive activity among orphan class-
785 A G protein coupled receptors. *PLoS One* **10**, e0138463 (2015).
786 30. Sethi, A. *et al.* The complex binding mode of the peptide hormone H2 relaxin to its
787 receptor RXFP1. *Nat. Commun.* **7**, 11344 (2016).
788 31. Büllesbach, E. E. & Schwabe, C. The trap-like relaxin-binding site of the leucine-rich G-
789 protein-coupled receptor 7. *J. Biol. Chem.* **280**, 14051–14056 (2005).
790 32. Marks, D. S. *et al.* Protein 3D structure computed from evolutionary sequence variation.
791 *PLoS One* **6**, (2011).
792 33. Marks, D. S., Hopf, T. A. & Sander, C. Protein structure prediction from sequence
793 variation. *Nat. Biotechnol.* **30**, 1072–1080 (2012).
794 34. Sjodt, M. *et al.* Structure of the peptidoglycan polymerase RodA resolved by evolutionary
795 coupling analysis. *Nature* **556**, 118–121 (2018).
796 35. Zheng, S. *et al.* Structure and mutagenic analysis of the lipid II flippase MurJ from
797 Escherichia coli. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 6709–6714 (2018).
798 36. Toth-Petroczy, A. *et al.* Structured states of disordered proteins from genomic sequences.
799 *Cell* **167**, 158-170.e12 (2016).
800 37. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: Algorithms for
801 rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).
802 38. Halls, M. L. *et al.* Multiple binding sites revealed by interaction of relaxin family peptides
803 with native and chimeric relaxin family peptide receptors 1 and 2 (LGR7 and LGR8). *J.*
804 *Pharmacol. Exp. Ther.* **313**, 677–687 (2005).
805 39. Diepenhorst, N. A. *et al.* Investigation of interactions at the extracellular loops of the
806 relaxin family peptide receptor 1 (RXFP1). *J. Biol. Chem.* **289**, 34938–34952 (2014).
807 40. Mintseris, J. & Gygi, S. P. High-density chemical cross-linking for modeling protein
808 interactions. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 93–102 (2020).
809 41. Büllesbach, E. E., Yang, S. & Schwabe, C. The receptor-binding site of human relaxin II:

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