Title: The relaxin receptor RXFP1 signals through a mechanism of autoinhibition

Authors: Sarah C. Erlandson[1], Shaun Rawson[1], James Osei-Owusu[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, Université de Montpellier, CNRS, INSERM 34094 Montpellier Cedex 5, France

[*]Correspondence 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. RXFP1 is a multi-domain G protein-coupled receptor (GPCR) with an ectodomain consisting of an LDLa module and leucine-rich repeats. The mechanism of RXFP1 signal transduction is clearly distinct from that of other GPCRs, but remains very poorly understood. Here, we present the cryo-electron microscopy structure of active-state human RXFP1, bound to a single-chain version of the endogenous agonist relaxin-2 and to the heterotrimeric G_s protein. 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. Our results explain how an unusual GPCR family functions, providing a path to rational drug development targeting the relaxin receptors.

30 **Main**

31 RXFP1 is a member of the leucine-rich repeat-containing GPCRs (LGR), a subset of
32 family A GPCRs that have remained a notable exception to our understanding of GPCR
33 signaling, despite substantial progress in studies of other GPCR families. In LGRs, the leucine-
34 rich repeats (LRRs) function as the extracellular ligand-binding domain for three types of protein
35 agonists: glycoprotein hormones, R-spondins, and relaxins[1]. The LGRs are involved in a variety
36 of physiological processes across reproductive and developmental biology. RXFP1, the receptor
37 for the relaxin-2 hormone in humans[2], plays an important role during pregnancy. In this setting,
38 it is responsible for physiological changes including increasing cardiac output and remodeling
39 reproductive tissues to facilitate parturition[3–5]. RXFP1 signaling also regulates the physiology of
40 numerous organs in both males and females, particularly the heart, lungs, liver, and kidneys.
41 Activation of RXFP1 by relaxin-2 in these organs leads to pleiotropic cellular effects, including
42 vasodilation, angiogenesis, anti-inflammatory responses, and extracellular matrix remodeling
43 through collagen degradation[6,7]. Accordingly, the RXFP1 receptor has emerged as a promising
44 therapeutic target for the treatment of cardiovascular and fibrotic diseases[8–10].

45 The relaxin receptors RXFP1 and RXFP2 are unique members of the LGR family, and
46 are classified as type C LGRs due to the presence of an additional domain called an LDLa
47 module at the receptors' distal N-termini, before the LRRs in sequence[11,12]. These receptors are
48 the only two mammalian GPCRs to contain an LDLa module, and the role of this domain in
49 relaxin receptor signaling is poorly understood. The LDLa module is dispensable for relaxin-2
50 binding to the LRRs but is essential for activation of RXFP1 signaling in response to relaxin-2[13].
51 The mechanisms that couple ligand binding in the LRRs to conformational changes within the
52 7TM domain required for G protein signaling remain undefined, largely due to an absence of
53 structural data. Recent structures of the luteinizing hormone-choriogonadotropin receptor
54 (LHCGR), one of the glycoprotein hormone receptors or type A LGRs, revealed conformational
55 changes of the LRRs between inactive and active states. These studies proposed that large
56 glycoprotein hormones signal through a steric "push-pull" mechanism that activates the receptor
57 by driving changes in LRR conformation[14]. However, the small 6 kDa size of the relaxin-2
58 peptide precludes such a mechanism, requiring an alternative explanation.

59 In order to elucidate the basis for RXFP1 signal transduction, we set out to determine the
60 active-state structure of human RXFP1 bound to an engineered relaxin-2 and the heterotrimeric

61 G protein G_s. We optimized receptor expression using fusions to a minimal G_α protein[15], then
62 formed a larger complex with the addition of G protein β_1 and γ_2 subunits, and determined the
63 structure using cryo-electron microscopy. Unexpectedly, the structure revealed that RXFP1's
64 extracellular loop 2 (ECL2) occupies the GPCR orthosteric ligand-binding pocket in the active
65 state. Results from structural and functional studies define a mechanism in which ECL2
66 conformation is regulated by the receptor's LRRs and hinge region, a short segment between the
67 LRRs and 7TMs. Collectively, these studies identify several conformational switches in both the
68 receptor ectodomain and 7TMs, showing that the concerted action of multiple receptor domains
69 controls the transduction of RXFP1 signaling by its agonist, relaxin-2.
70
71 **Results**
72 **Cryo-EM structure of the RXFP1–G protein complex**
73 Wild type (WT) full-length RXFP1 receptor could be expressed only at very low levels in
74 mammalian cells. To enable structural studies, we cloned a fusion of RXFP1 to the engineered
75 G_α protein minimal G_s (mini-G_s)[15]. Truncations of RXFP1's flexible C-terminus further
76 increased expression levels, with the optimal expression construct having a C-terminal truncation
77 of 20 amino acids (**Fig. S1**). The binding of the catalytically inactive mini-G_s protein blocks
78 RXFP1 signaling through endogenous G_s proteins and likely stabilizes the receptor, leading to
79 higher purification yields. The fusion protein of RXFP1 and mini-G_s was purified in complex
80 with human G protein β_1 and γ_2 subunits, the camelid antibody VHH fragment nanobody 35
81 (Nb35)[16], and an engineered version of the agonist relaxin-2 (SE001)[17] to form an agonist–
82 GPCR–G protein complex (hereafter referred to as RXFP1–G_s).
83 Cryo-electron microscopy (cryo-EM) was used for structural studies of RXFP1–G_s.
84 Initial two-dimensional classification analysis revealed averages that showed clear density for
85 RXFP1's 7TM domain and the heterotrimeric G protein. In contrast, the density for RXFP1's
86 ectodomain was weak and poorly defined, indicating flexibility of the ectodomain with respect to
87 the transmembrane domain. Due to the conformational heterogeneity of RXFP1–G_s, we analyzed
88 the cryo-EM data using two different approaches. The first utilized masking of RXFP1's 7TM
89 and G proteins to obtain a high-resolution cryo-EM map of these domains at 3.2 Å, allowing us
90 to build an atomic model (**Fig. S2**). The second approach used focused classifications of
91 RXFP1's ectodomain to obtain a cryo-EM map of the entire complex (**Fig. S3**). As a result of the

92 ectodomain's flexibility, the cryo-EM map of the full-length receptor is lower resolution, with an
93 overall resolution of 4.2 Å and local resolution of the ectodomain between 5-8 Å.
94 RXFP1's 7TM domain displays characteristic hallmarks of the active state for family A
95 GPCRs (**Fig. 1**). Most notable is the outward conformation of the intracellular end of
96 transmembrane helix 6 (TM6), which creates the binding site for the α5 helix of $G_{\alpha s}$[18,19].
97 Additional active-state features include an open conformation of the "ionic lock" between
98 Glu623$^{6.30}$ and Lys510$^{3.50}$ and the hydrogen bond between Tyr681$^{7.53}$ of the conserved NPxxY
99 motif and Tyr599$^{5.58}$ (superscript indicates Ballesteros-Weinstein numbering system)[20].
100 Hydrogen bonds between Tyr681$^{7.53}$ and Tyr599$^{5.58}$ in active-state family A GPCRs are
101 coordinated by a bridging water molecule, not visible at the resolution of our map[19]. In RXFP1–
102 G_s, the active state of the 7TMs displays a canonical interaction with heterotrimeric G_s, similar to
103 previously reported GPCR–G protein structures[21].
104

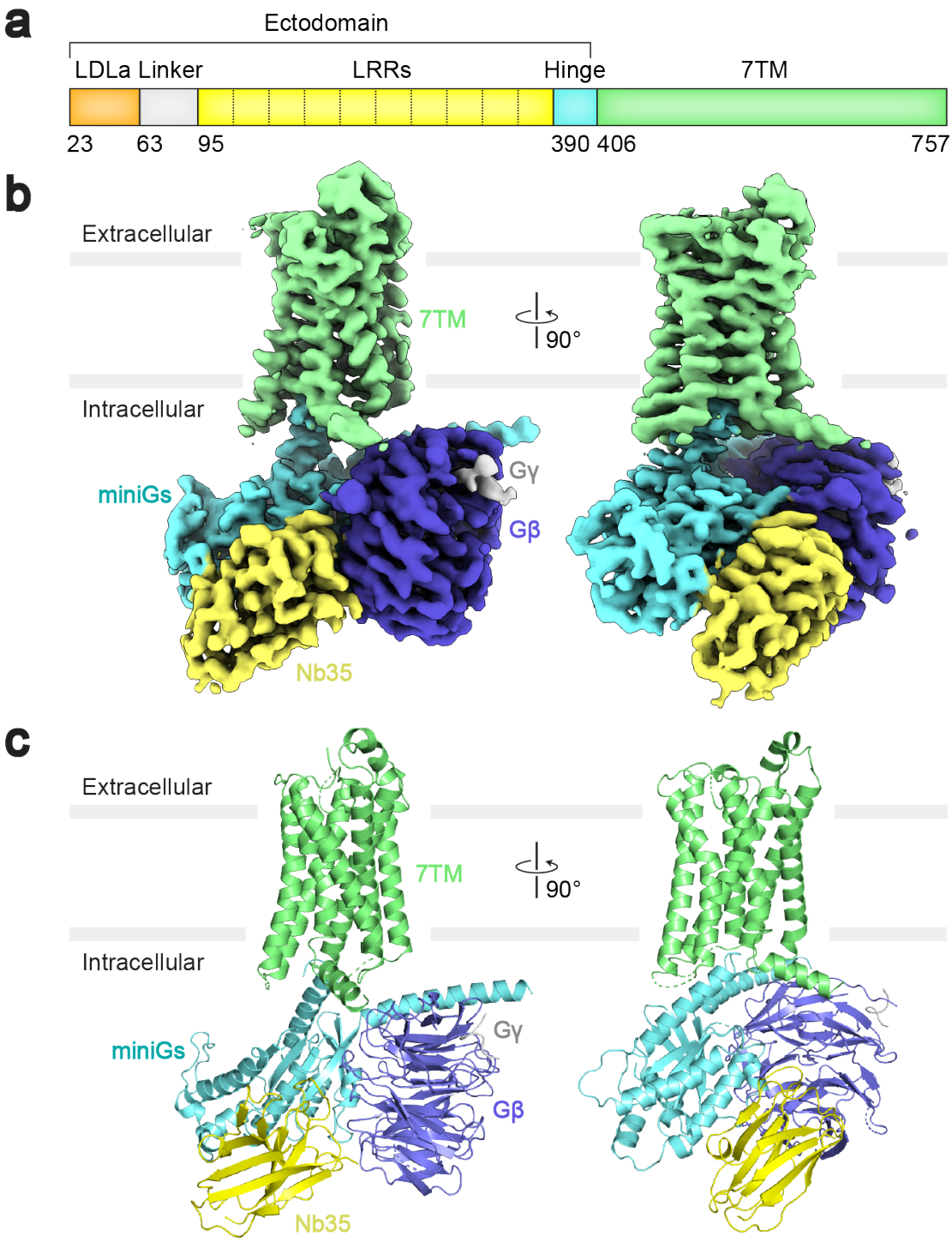


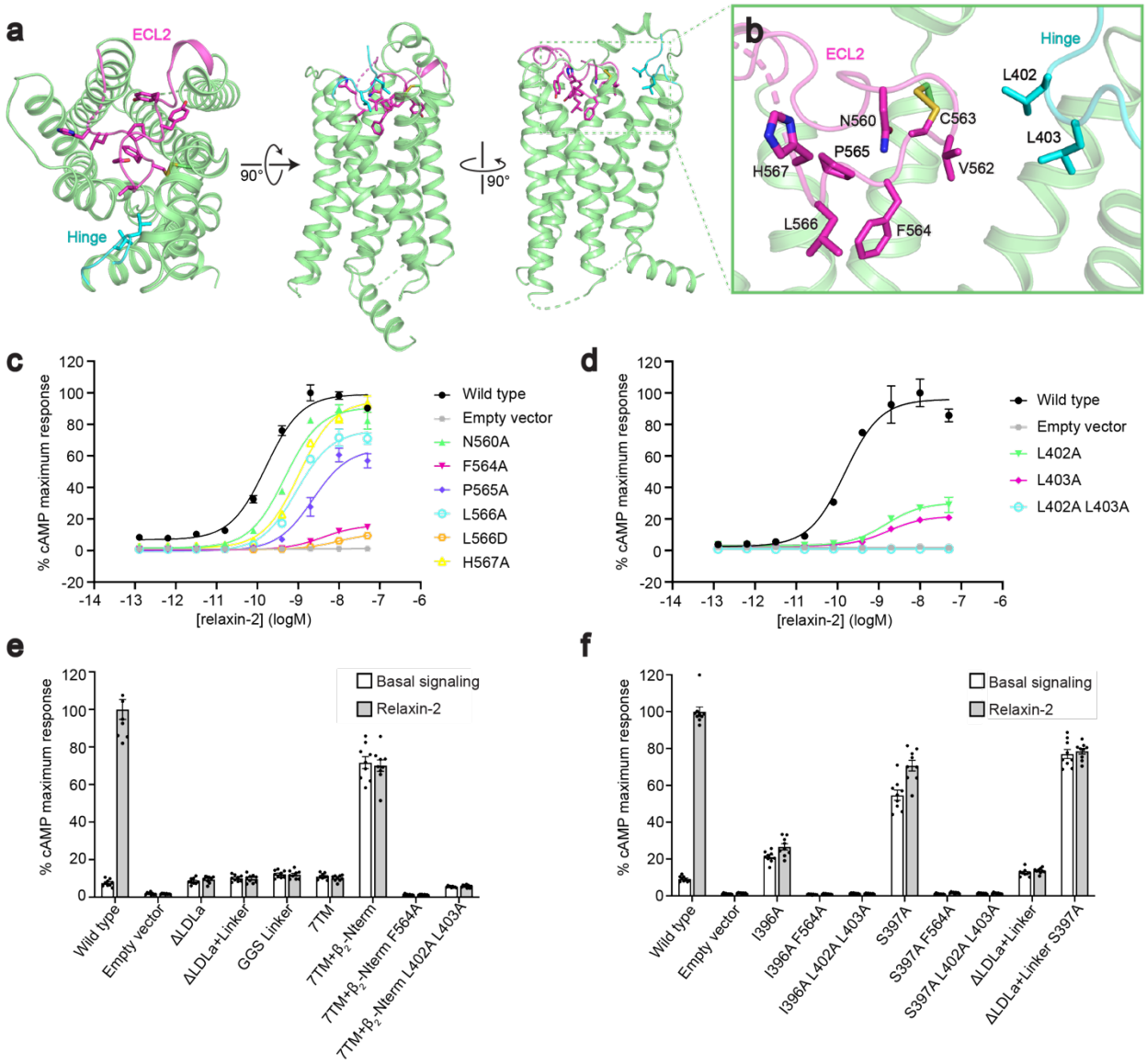
Fig. 1. Cryo-EM map and model of the RXFP1–G_s complex. a, Diagram of the primary structure of RXFP1 domains. **b-c**, Cryo-EM map (**b**) and model (**c**) of the RXFP1–G_s complex 7TM domain with heterotrimeric G_s proteins and Nb35.

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[22]. 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[23] and angiotensin II analogs to the angiotensin
131 II type I receptor[24] (**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**). Mutation of the
137 Pro565^{ECL2} residue between Phe564^{ECL2} and Leu566^{ECL2} also decreases relaxin-2 signaling,
138 likely by disrupting the loop structure of these ECL2 residues within the orthosteric site (**Fig.
139 2c**)[25].

140 The discovery of the importance of ECL2 to RXFP1 signaling is reminiscent of the
141 orphan receptor GPR52, which is directly activated by its own ECL2 as a tethered agonist[26].
142 While ECL2 shows no conservation of sequence or detailed structure between RXFP1 and
143 GPR52, the binding sites for ECL2 are in analogous positions within the 7TM domain for both
144 GPCRs (**Fig. S4a**). These structural parallels, along with RXPF1 mutagenesis in cell signaling
145 assays, are consistent with ECL2 serving a critical role in activating RXFP1.
146



147

148 **Fig. 2. Regulation of receptor signaling by ECL2 and the ectodomain. a**, The conformation
149 of ECL2 and the hinge region in active-state RXFP1. **b**, Details of ECL2 in the 7TM orthosteric
150 site and interactions between ECL2 and the hinge region. **c-d**, The effect of ECL2 (**c**) and

151 Leu402 and Leu403 hinge region (**d**) mutations in an assay for G_s signaling by RXFP1. Data are
152 mean ± s.e.m. from technical triplicates. **e-f**, Basal signaling and signaling in response to 50 nM
153 relaxin-2 for RXFP1 ectodomain truncation constructs (**e**) and Ile396 and Ser397 hinge region
154 mutations (**f**) in assay for G_s signaling. Data are mean ± s.e.m. from nine technical replicates.
155

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

157 Our structure of the 7TM domain also includes 6 residues from the hinge region of
158 RXFP1's ectodomain. The hinge region is composed of residues between the end of the LRRs
159 and the beginning of TM1, and the six residues adjacent to TM1 are included in our refined
160 model. We observed that the hinge region curves into the top of the 7TM domain, with residues
161 Leu402 and Leu403 in close proximity to ECL2 in RXFP1's orthosteric site. Additionally,
162 residues from both ECL2 and the hinge region pack against the top of TM7 (**Fig. 2b**).

163 Functional analyses of type A LGRs, the glycoprotein hormone receptors (GPHRs), have
164 previously established that the hinge region of those LGRs is critical for receptor signaling[27,28].
165 The unstructured loop in the hinge region of GPHRs is approximately 60-125 residues long and
166 plays a role in binding the receptors' glycoprotein hormone agonists[29]. For GPHRs, a 10-residue
167 section of their hinge region near the 7TM domain, P10, is also critical for receptor signaling[27,28]
168 and has been recently shown to adopt different conformations in inactive and active receptor
169 states of LHCGR[14]. In contrast, the hinge loop of type C LGRs, such as RXFP1, is predicted to
170 be about 15 residues long, has no established function, and shows no sequence conservation in
171 comparison to type A LGRs[1]. Despite the clear differences between GPHRs and RXFP1, the
172 interactions between the hinge region and orthosteric site in RXFP1's active-state structure
173 unexpectedly suggested that these residues may also play a role in signaling.

174 To investigate the role of the hinge region residues Leu402 and Leu403, we cloned single
175 mutations to Ala and constructed a double mutant with Ala substitutions for both residues. These
176 constructs expressed at low levels, so wild type receptor expression was reduced in order to
177 compare receptor signaling (**Fig. S5b**). While the single Ala mutants each decreased the efficacy
178 of relaxin-2, the double Ala mutant completely ablated RXFP1 signaling, despite maintaining the
179 ability to bind relaxin-2 and expression levels at roughly 50% of wild type (**Fig. 2d, S5f**). These
180 results indicated that Leu402 and Leu403 of the hinge region are essential for RXFP1 activation,
181 likely functioning to stabilize ECL2 into its active-state conformation in the orthosteric site.

182 **Autoinhibition of RXFP1 signaling**

183 While structures of RXFP1 and GPR52 revealed that ECL2 is critical for the activation of
184 both receptors, other aspects of their signaling suggest differing mechanisms. GPR52 is a self-
185 activated orphan GPCR with very high basal activity, signaling at 90% of its E_{max} without any
186 agonist bound[26,30]. The mechanism governing the intrinsic activity of GPR52 is clear, as the
187 ECL2 tethered agonist is a component of the 7TM structure itself. In contrast, RXFP1 does not
188 have high basal activity, but signals in response to the binding of relaxin-2 to its LRRs. These
189 differences suggest that while both GPCR structures show ECL2 binding in the orthosteric site,
190 RXFP1 likely uses additional mechanisms to prevent continuous self-activation of the 7TM
191 domain. The major structural difference between these two receptors is that GPR52 is a
192 conventional family A GPCR with an unstructured N-terminus, while RXFP1 is a type C LGR,
193 containing a structured ectodomain of LRRs and an LDLa module. For this reason, we set out to
194 investigate the role of RXFP1's ectodomain in modulating the activity of ECL2.

195 To address this question, we first cloned constructs of RXFP1 with deletions of the
196 receptor's ectodomain. Basal signaling of RXFP1 was not increased by deletion of the LDLa
197 module, deletion of the LDLa module and the 32-residue linker that connects it to the LRRs, or
198 replacement of the linker with a 32 residue Gly-Gly-Ser linker (**Fig. 2e**). These results indicated
199 that the LDLa module and linker of RXFP1's ectodomain do not play an inhibitory role in
200 signaling. In contrast, these constructs are unable to signal in response to relaxin-2, consistent
201 with previous studies showing that the LDLa module and linker region are essential for receptor
202 activation and that the linker is involved in relaxin-2 binding[13,31] (**Fig. 2e, S5e**). We next focused
203 on testing the role of RXFP1's LRRs in the receptor signaling mechanism. To remove the
204 ectodomain, including the LRRs, constructs were designed to express RXFP1's 7TM domain
205 alone, which also included several residues from the hinge region immediately preceding TM1
206 (**Table S8**). However, these constructs showed very low cell-surface expression, at 12% of wild
207 type. Although the 7TM domain expressed very poorly, it showed similar basal signaling to wild
208 type RXFP1 in a G_s signaling assay, at 11% of the wild type relaxin-2 E_{max}. To increase
209 expression of the 7TMs, we cloned a fusion of RXFP1's 7TM domain to the unstructured N-
210 terminus of the high-expressing β_2 adrenergic receptor. This fusion rescued 7TM domain
211 expression to essentially wild type levels (**Fig. S5c**). When tested in a G_s signaling assay, the
212 fusion showed a high level of basal activity. In the absence of ligand, the 7TM domain signaled

213 at 70% of the maximum level of agonist-induced signaling for wild-type RXFP1 (**Fig. 2e**). As
214 expected, RXFP1's 7TM domain alone does not show any change in activity in response to
215 relaxin-2, since the relaxin-2 binding sites in the ectodomain are deleted in this construct[32] (**Fig.**
216 **2e, S5f**).
217 To establish whether the high basal activity of RXFP1's 7TM alone is due to ECL2, we
218 introduced the Phe564^{ECL2} to Ala mutation that greatly reduced full-length RXFP1 signaling in
219 response to relaxin-2. The Phe564^{ECL2} to Ala mutation was able to completely ablate the 7TM
220 domain's high basal activity, confirming that ECL2 constitutively activates the 7TMs in the
221 absence of RXFP1's LRR domain. Likewise, Ala mutations of the hinge region residues Leu402
222 and Leu403 (which are included in the 7TM domain fusion construct) were also able to reduce
223 the high basal signaling. These data suggested a model of signaling in which the LRRs play an
224 autoinhibitory role in regulating the active state of ECL2 and the hinge region. As a result,
225 deletions of the LRRs allow constitutive activation the receptor, leading to high basal signaling.
226
227 **Inhibitory interactions revealed by evolutionary coupling analysis**
228 An additional insight into the regulation of ECL2 by the ectodomain arose from
229 evolutionary coupling (EC) analysis of RXFP1[33,34]. The strongest ECs, or evolutionary coupled
230 residues, are derived from applying a global probability model to multiple sequence alignments
231 and typically indicate residue pairs that are in contact in 3D[35], including residues involved in
232 conformational changes[36,37]. Residues of ECL2 had strong ECs pairing them with residues within
233 the 7TM helices, supporting our active-state structure. However, ECL2 also showed ECs with
234 residues from RXFP1's hinge region not present in our model (**Fig. S6**). In total, two residues
235 from the hinge region, Ile396 and Ser397, had ECs with three residues of ECL2, including the
236 critical Phe564^{ECL2}. ECL2 and the hinge region are likely involved in other interactions than
237 those observed in our structure of the active state. If ECL2 contacts Ile396 and Ser397 in an
238 inactive-state conformation, we predicted that those interactions may contribute to the inhibition
239 of ECL2 in the absence of relaxin-2 binding to the LRRs.
240 Ile396 and Ser397 were each mutated to Ala to test the effects on the activation state of
241 ECL2. Despite having low expression levels, the Ile396 and Ser397 mutations each showed a
242 significant increase in basal signaling, at 21% and 55%, respectively, of wild type RXFP1's E_{max}
243 (**Fig. 2f**, **S5d**). The Ser397 to Ala mutant similarly increased the basal signaling of RXFP1 with a

244 deletion of the LDLa module and linker, confirming that those domains are not involved in the
245 mechanism of signaling inhibition. Interestingly, the hinge mutants showed a reduced signaling
246 response to relaxin-2 binding, suggesting that these residues involved in a potential inactive-state
247 ECL2 interface are also important for allosteric communication between the ectodomain and
248 7TM domain (**Fig. 2f, S5f**). Addition of the Phe564^{ECL2} or Leu402 and Leu403 substitutions to
249 the Ile396 and Ser397 mutants was able to abolish the increase in basal signaling, confirming
250 that RXFP1 activation is dependent on these residues (**Fig. 2f**).
251
252 **Mechanism of RXFP1 7TM autoactivation**
253 We performed enhanced-sampling molecular dynamics (MD) simulations (see Methods
254 for details) to study the role of ECL2 and the hinge region in the basal activity of RXFP1's 7TM
255 domain. An inactive-state model of the 7TM domain was obtained by MD simulations of
256 deactivation, starting from the cryo-EM structure truncated before the hinge region (residues
257 395–699). A sodium ion was placed in the sodium-binding site to favor the sampling of inactive
258 states during the simulations. We obtained an inactive state that was strikingly similar to the
259 inactive-state AlphaFold2 model[38], in which the intracellular end of TM6 bent toward TM3 to
260 form the classic Lys510$^{3.50}$–Glu623$^{6.30}$ ionic lock (**Fig. S7a-c**). The second segment of ECL2
261 maintained the same conformation throughout the simulations, whereas Ser568^{ECL2} to Ser573^{ECL2}
262 in the third segment were highly mobile, consistent with absence of clear density in the cryo-EM
263 map. Starting from the inactive state, we first simulated autoactivation of the RXFP1 7TM by
264 removing the sodium ion and protonating the sodium anchor, Asp451$^{2.50}$. As negative controls,
265 we performed the same simulations for two mutants with low basal activity, F564^{ECL2}A and
266 L566^{ECL2}D. The WT RXFP1 7TM exhibited autoactivation with outward movements of TM6 on
267 the intracellular side, destabilization of the ionic lock, and frequent side-chain flips of the toggle
268 switch W641$^{6.48}$ (**Fig. 3a-c**). In contrast, the two mutants remained in inactive conformations.
269 The WT 7TM exhibited distinct shapes of the orthosteric pocket compared to the two mutants,
270 owing to ECL2–TM7 interactions. Namely, Phe564^{ECL2} in the WT 7TM domain "pushed" TM7
271 toward TM2, which likely altered W641$^{6.48}$ conformations and triggered autoactivation (**Fig. 3a-**
272 **c**). The F564^{ECL2}A mutation directly eliminates this steric effect, whereas substituting
273 Leu566^{ECL2} with Asp reorients the charged side chain away from the pocket, leaving space for
274 F564^{ECL2} and diminishing its impact on TM7 conformation (**Fig. 3a,b**).

275 The truncated 7TM model is unsuited for studying the hinge region, which partially
276 unfolds during the simulations. Therefore, we used a truncated ^{half}LRRs-7TM form (residues
277 242–699) to investigate the role of the hinge. The ^{half}LRRs-7TM model was truncated before the
278 Cys243-Cys279 disulfide bond in the middle of the LRRs to prevent unfolding, while reducing
279 the system size to enable sufficient MD sampling. MD simulations were performed for the
280 constitutively active mutant S397A, in comparison with the WT and the triple mutant
281 S397A/L402A/L403A, starting from the inactive state. We found that the S397A mutation
282 disrupts H-bonds in the hinge and increases the mobility of the LRRs, which diminishes the
283 autoinhibition of ECL2 and promotes 7TM activation (**Fig. 3d and Table S6**). L402A/L403A
284 attenuates the effect of S397A by stabilizing the receptor in a different conformation (**Fig. 3d**).
285

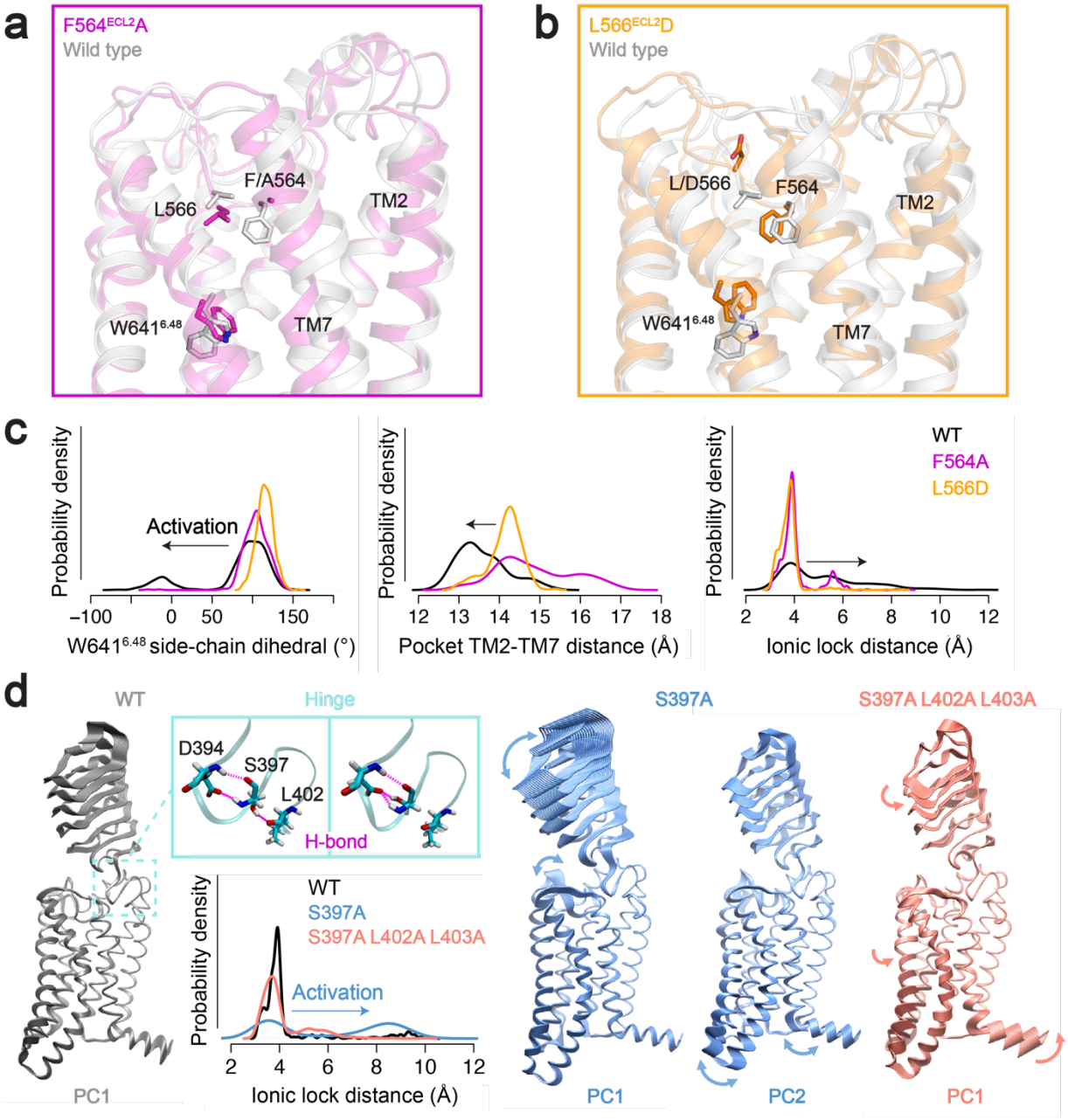


Fig. 3. Molecular dynamics of RXFP1 starting from the inactive-state AlphaFold2 model. **a-b,** The truncated RXFP1 7TM domain alone shows autoactivation. Autoactivation in these simulations is impaired by the addition of the F564A (**a**) or the L566D mutations (**b**). **c**, Histograms describing activation-related conformational differences between WT, F564A, and L566D RXFP1 7TM models, including the distance between TM2 and TM7 in the orthosteric site, side-chain flips of the toggle switch residue W641$^{6.48}$, and the ionic lock distance. **d,** Molecular dynamics of truncated RXFP1 ^{half}LRRs-7TM. Projection of the trajectories on the first and second principal components (PC) illustrates the mechanism of S397A-induced basal

295 activity. The S397A mutation disrupts the H-bonds with L402 (backbone) and D394 (side chain)
296 present in the WT (Table S6). The hinge and LRRs become more mobile in the S397A mutant,
297 which triggers activation through ECL2. Addition of the L402A/L403A mutations reduces the
298 steric hindrance of the hinge and leads to an overall twist of the receptor, which attenuates the
299 activation effect of S397A.
300
301 **Relaxin-2 binding to the leucine-rich repeats**
302 The cryo-EM map of the full-length RXFP1–G_s complex is limited by lower resolution
303 due to the dynamic nature of the receptor's ectodomain. Continuous heterogeneity present in the
304 final particle stack for RXFP1–G_s was visualized through three-dimensional variability analysis
305 in cryoSPARC[39]. The resulting movies showed that the LRRs are very flexible in the relaxin-2–
306 bound state, moving at the hinge region between the LRRs and 7TM domain (**Movie S1**).
307 Despite these limitations, the cryo-EM map offered several new insights into the overall
308 ectodomain architecture and relaxin-2 binding.
309 In the refined cryo-EM map of active-state RXFP1, the LRRs are positioned above
310 ECL1, likely giving this region a role in stabilizing ectodomain conformations. Mutation of the
311 ECL1 residue $W479^{ECL1}$ has been previously shown to reduce both relaxin-2 binding and
312 signaling[25]. In the active-state structure, W479 projects from ECL1 to interact extensively with
313 residues in the 7TM domain. Most or all of these interactions would be abrogated by the W479A
314 substitution, accounting for lack of RXFP1 function due to structural changes that would
315 destabilize the receptor.
316 The LRRs are extended away from the transmembrane domain in the active state, at an
317 angle of 40° from the membrane plane (**Fig. 4a**). The orientation of the ectodomain also rotates
318 the concave ligand-binding side of the LRRs away from the extracellular side of the 7TM
319 domain (**Fig. 4b**). Each of these features physically separates the relaxin-2 binding site on the
320 LRRs from the 7TMs. While a secondary binding site between relaxin-2 and the extracellular
321 loops (ECLs) has been proposed[25,40], the active state that we captured through cryo-EM does not
322 show any direct interaction between relaxin-2 and the ECLs.
323 The low-resolution map shows relaxin-2 bound to the concave side of the LRRs. To aid
324 with modeling the relaxin-2–LRR interaction, we used crosslinking mass spectrometry (CLMS)
325 with the RXFP1–G_s complex. Using the Extended-EDC approach with an EDDA crosslinker[41],
326 $Glu14^{B\text{-}chain}$ of relaxin-2 crosslinked with three residues on the LRRs, Glu206, Glu299, and

327 Glu351 (**Fig. 4c,d**). Several residues involved in relaxin-2 binding have been previously
328 characterized through mutational analysis in radioligand binding assays. In those experiments, a
329 series of conserved residues on the B-chain of relaxin-2, Arg13, Arg17, and Ile/Val20, a motif
330 known as the relaxin binding cassette, were proposed to be part of the interface[42,43]. Multiple
331 residues on the concave side of the LRRs were also found to be important for relaxin-2
332 binding[32]. Interestingly, the Glu14$^{\text{B-chain}}$ residue of relaxin-2 crosslinked to the LRRs is adjacent
333 to the Arg13$^{\text{B-chain}}$ of the relaxin binding cassette. Additionally, Glu299 on the LRRs had been
334 previously identified as a residue involved in relaxin binding by mutational studies, highlighting
335 the close agreement between our CLMS data and prior functional analyses.
336 The residues from our CLMS experiment were used in combination with the mutational
337 data as restraints for docking the relaxin-2–LRR interaction in HADDOCK[44] (**Fig. 4c,d**). The
338 highest scoring HADDOCK models agreed well with our low resolution cryo-EM map and
339 showed the B-chain of relaxin-2 bound to the concave side of the LRRs, while the A-chain made
340 limited contacts (**Fig. 4e**). Additional density in the cryo-EM map is present at multiple sites of
341 potential N-linked glycosylation and next to the A-chain **(Fig. S8e,f)**. Density near the A-chain
342 of relaxin-2 may belong to the ectodomain's linker region, which the A-chain has been proposed
343 to bind[31]. However, the low resolution and absence of crosslinks for these domains prevented
344 further characterization of A-chain interactions.
345 In the docked model of relaxin-2 bound to the LRRs, Glu14$^{\text{B-chain}}$ falls within the E-EDC
346 crosslink distance of 14 Å from Glu206, Glu299, and Glu351 and is not directly involved in the
347 relaxin-2 binding interface, in agreement with previous mutational analysis[43]. The model also
348 predicted that one of the relaxin binding cassette residues, Arg17$^{\text{B-chain}}$, interacts with Glu206, a
349 residue on the LRRs identified by CLMS that was not previously known to be involved in the
350 binding site. To verify this interaction, we mutated Glu206 to Ala and tested the effect of the
351 mutation on binding of an Fc-tagged relaxin-2 protein (SE301)[17]. The Glu206 to Ala mutant
352 expressed at equivalent levels to wild type receptor, but the single mutant reduced relaxin-2
353 binding to 42% of wild type levels (**Fig. 4f,g**), validating the proposed interaction.
354

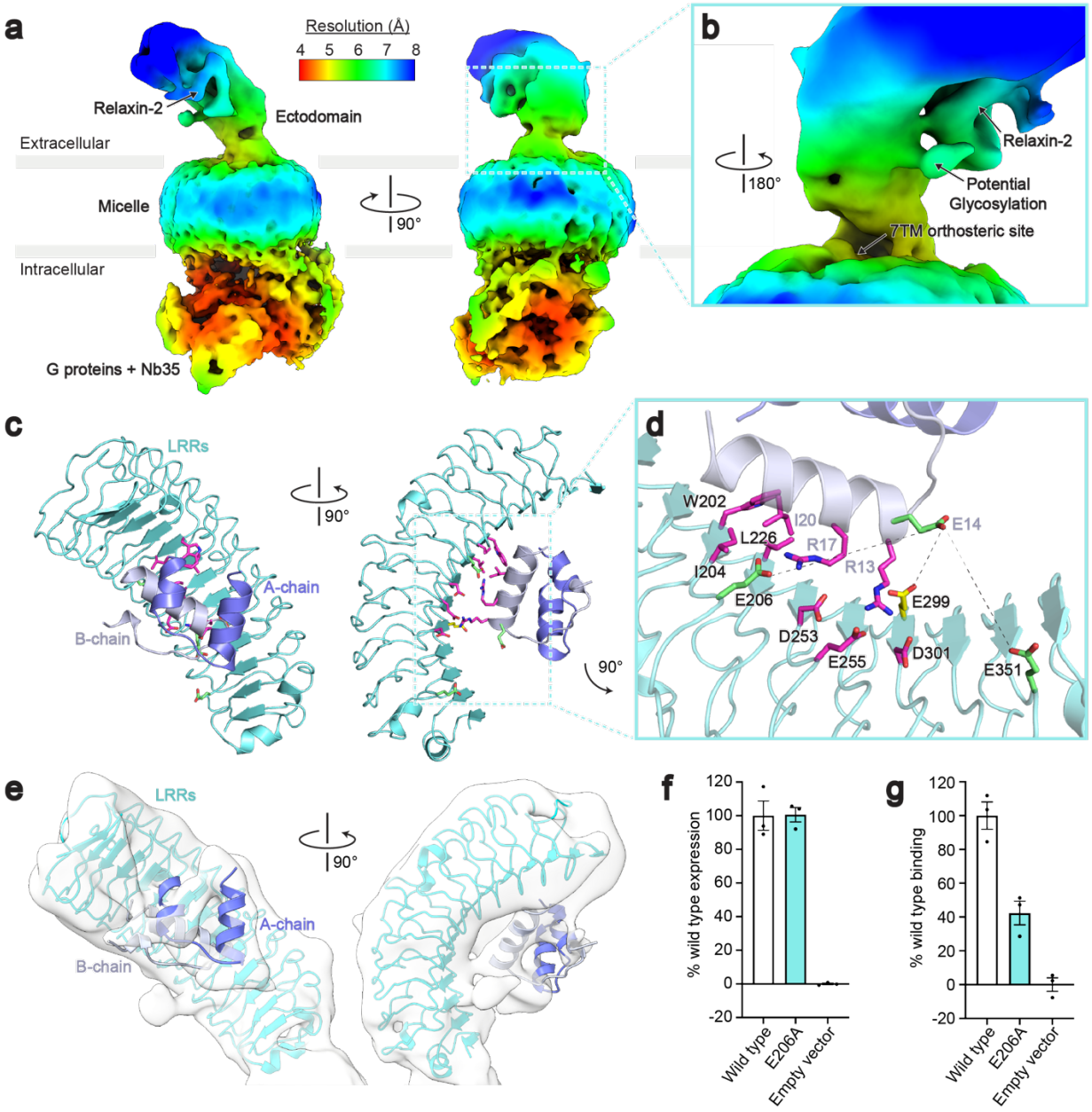


Fig. 4. Cryo-EM and crosslinking mass spectrometry reveal interactions between relaxin-2 and the leucine-rich repeats. a, Local resolution cryo-EM map of the full-length RXFP1–G_s complex. **b**, The relaxin-2 binding site is above and rotated away from the 7TM orthosteric site. **c**, Model of the relaxin-2–LRR 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.

366 **Discussion**

367 Our active-state structure of RXFP1–G_s revealed unexpected features involved in the
368 activation of RXFP1. Most strikingly, RXFP1's ECL2 was observed to occupy the GPCR
369 orthosteric ligand binding site, with the key residues Phe564^{ECL2} and Leu566^{ECL2} playing an
370 essential role in receptor activation. Residues Leu402 and Leu403 of the hinge region were also
371 required for RXFP1 signaling. These two signaling motifs were shown to be autoinhibited in the
372 absence of relaxin-2 binding by both the LRRs and two residues of the receptor hinge region,
373 Ile396 and Ser397, with the minimal inhibitory ectodomain construct requiring both of these
374 features. We also defined the binding site of relaxin-2 on the LRRs, showing that it is physically
375 separated from the 7TM domain in the active state.

376 Our observations indicate that RXFP1 cannot be controlled by the steric occlusion
377 mechanism proposed for other GPHRs[14,45]. In fact, owing to its small size (6 kDa), relaxin-2
378 binding to the RXFP1 ectodomain is fully compatible with a membrane-proximal inactive
379 conformation like that observed for the inactive-state of LHCGR (**Fig. 5a**). As a result, purely
380 steric effects cannot drive receptor activation in RXFP1, necessitating an alternative mechanism.
381 Moreover, RXFP1 and LHCGR have opposing LRR orientations in their active-state structures,
382 highlighting the divergence of mechanisms between GPHRs and RXFP1 (**Fig. S9**). A possible
383 mechanism for RXFP1 activation is suggested by the fact that, unlike other GPHRs, RXFP1
384 contains an LDLa domain which is strictly required for signaling (**Fig. 2e**)[13]. NMR studies with
385 soluble constructs of RXFP1's ECLs concluded that the LDLa module and residues from the
386 adjacent linker region may interact with ECL2[25,31,46], consistent with our observation that ECL2
387 serves as a key activation switch. The LDLa module was not resolved in our maps, suggesting
388 that it is mobile in the active state of the receptor (**Fig. 5b**).

389 Notably, the structure of LHCGR in complex with G proteins[14] shows a similar ECL2
390 conformation to that observed for the relaxin receptor, with a Phe515/Met517 pair positioned
391 similarly to Phe564^{ECL2} and Leu566^{ECL2} in RXFP1. In fact, all LGRs share a CΦPΦ sequence
392 motif in ECL2 (where "Φ" denotes a hydrophobic amino acid), and AlphaFold2 models of all
393 LGRs show similar ECL2 conformations to those of LHCGR and RXFP1. This suggests that
394 ECL2-triggered activation may be a general feature of the LGR family as a whole, inducing
395 receptor activation in response to conformational changes of the LRRs and hinge region by either
396 steric "push" of the ligand or, in the case of RXFP1, indirect rearrangement of the LDLa module.

397 The unusual features of RXFP1 signaling by relaxin-2 raises the question of whether
398 other ligands, such as the small molecule agonist ML290[47], activate the receptor through similar
399 mechanisms. To answer this question, we assayed ML290's signaling activity at wild type
400 RXFP1 versus mutants of ECL2 and the hinge region, finding a similar dependence on these
401 residues to relaxin-2 (**Fig. S10**). Our results indicate that small molecules are likely able to
402 exploit the ECL2/hinge region conformational switch, although more work will be required to
403 understand ML290's mechanism in detail. Additional small molecule or biologic agonists could
404 be created to exploit allostery in the receptor and either mimic ECL2-induced activation or
405 relieve inhibitory interactions between the ectodomain and ECL2. Such molecules could be
406 useful therapeutics for the treatment of numerous cardiovascular and fibrotic diseases, and
407 similar approaches may be applicable to other members of the LGR family.

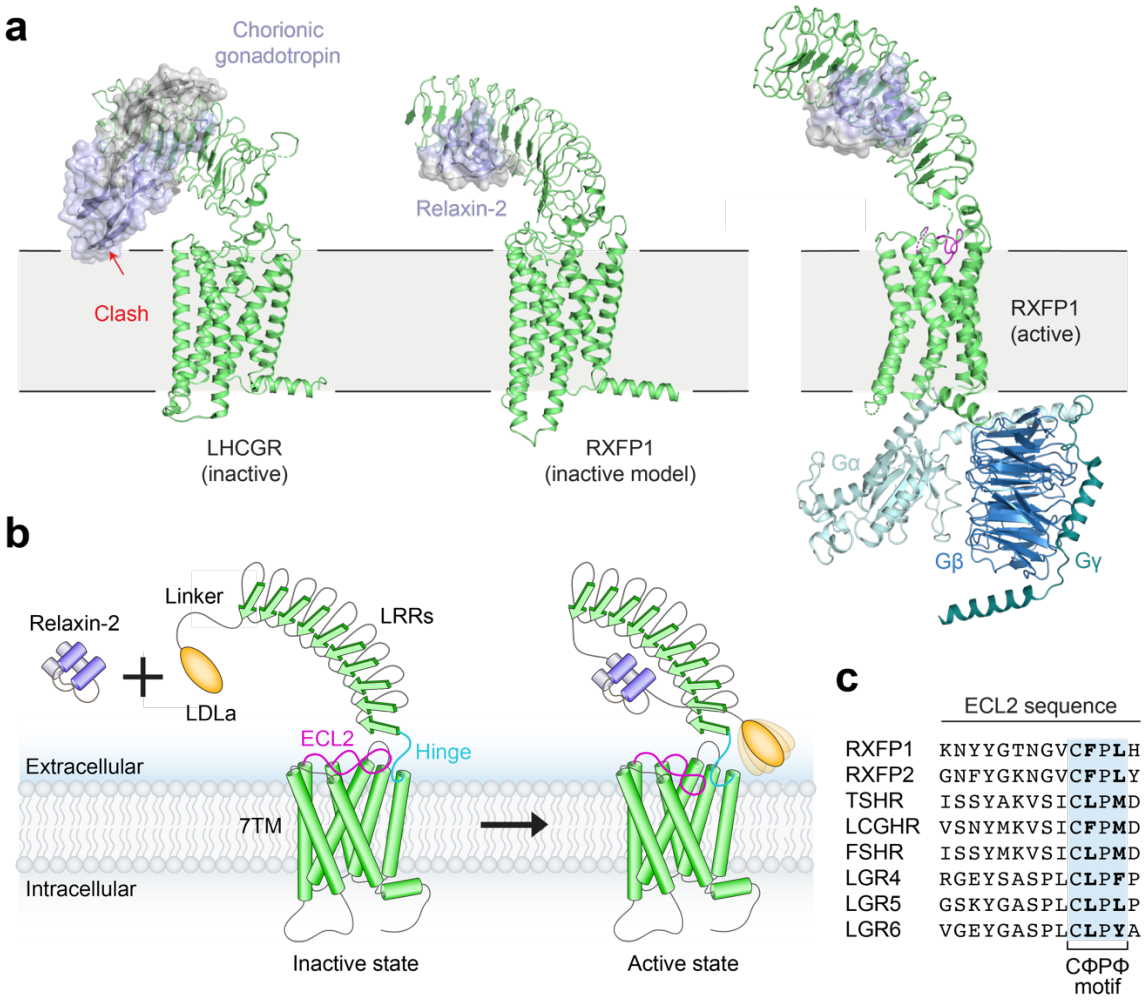


Fig. 5. Model of RXFP1 activation by relaxin-2. a, Model of LHCGR in the inactive conformation (PDB ID: 7FIJ) bound to the hormone chorionic gonadotropin, showing steric clash with the membrane that induces activation[14]. In contrast, a similar potential conformation of RXFP1 would easily accommodate bound relaxin-2 (generated from alignments of inactive-state AlphaFold2 models of RXFP1's LRRs and 7TM to PDB 7FIJ). A hybrid model of active-state RXFP1 is also shown for comparison, based on our 7TM domain structure, cryo-EM maps, and AlphaFold2 model of the LRRs with docked relaxin-2 hormone. The hybrid model uses $G\gamma_2$ from PDB 3SN6[16]. **b,** The inactive state of RXFP1 is characterized by inhibitory interactions between ECL2, the hinge region, and the LRRs that prevent receptor activation. Relaxin-2 binds to the concave side of the LRRs, away from the 7TM domain. Relaxin-2 binding leads to reorganization of the LDLa/hinge/ECL2 interface, allowing residues in both the hinge region and ECL2 to activate the receptor. **c,** Sequence alignment of human LGRs showing the conserved CΦPΦ motif in ECL2.

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. Molecular dynamics simulations were performed using HPC resources from GENCI-TGCC France (grants 2021-2022 Spe00015 and A0100712461).

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, and cryo-EM grid preparation were performed by S.C.E. with supervision from A.C.K. Flow cytometry and cell signaling assays were performed by S.C.E and J.O.-O. with supervision by 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., X.C., and A.C.K. with input from all authors.

452 **Methods**

453 Cloning of RXFP1 constructs

454 Residues 23-757 of human RXFP1 were cloned with an N-terminal hemagglutinin signal
455 sequence, FLAG tag, and 3C protease site into the pcDNA-Zeo-tetO vector[48]. Fusions to the
456 miniG$_s$-399 protein with C-terminal truncations to RXFP1 were constructed using PCR followed
457 by NEBuilder HiFi DNA Assembly (New England Biolabs). Truncations removed 10, 15, 20,
458 25, 30, or 35 residues from the receptor C-terminus. For signaling assays, human RXFP1
459 residues 23-757 with an N-terminal hemagglutinin signal sequence and FLAG tag were cloned
460 into pcDNA-Zeo-tetO. Mutations to ECL2, hinge region, or LRR residues were introduced using
461 Quikchange Lightning PCR (Agilent). Ectodomain truncations were constructed using PCR and
462 NEBuilder HiFi DNA Assembly. Residue numbering is based on the canonical RXFP1 sequence
463 beginning at the initiating Met residue (UniProt ID Q9HBX9).

464

465 Cell surface expression tests

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

483 sample were collected and analyzed using FlowJo (**Fig. S11**). Data was normalized using the
484 wild type and empty vector mean fluorescence intensities as 100% and 0%, respectively, and
485 plotted using GraphPad Prism.
486
487 cAMP signaling assay
488 The GloSensor assay from Promega, a live-cell signaling assay that detects cellular
489 cAMP levels, was used to measure activation of G_s signaling through RXFP1. White, clear-
490 bottom 96-well plates (Thermo Fisher Scientific) were coated with 30 µL of 10 µg/mL poly-D-
491 lysine (Sigma-Aldrich) and washed once with PBS. HEK293T cells were then plated at $2x10^4$
492 cells/well in DMEM with 10% (v/v) FBS. The following day, cells were transfected with 20 ng
493 of GloSensor reporter plasmid and 20 ng (unless otherwise stated) of human RXFP1 or empty
494 vector DNA per well using FuGENE, according to the manufacturer's instructions. For **Fig. 2d**,
495 0.4 ng of wild type RXFP1 DNA was used per well, plus 19.6 ng of empty vector DNA,
496 according to ratios determined during cell surface expression tests. For **Fig. 2e**, 10 ng of
497 7TM+β_2-Nterm DNA was used per well, plus 10 ng of empty vector DNA, according to ratios
498 determined during cell surface expression tests. Twenty-four hours later, the media was replaced
499 with 40 µL of CO_2-independent media (Thermo Fisher Scientific) with 10% (v/v) FBS and 2
500 mg/mL D-luciferin (Goldbio) and incubated for 2 hours at room temperature (RT) in the dark.
501 Measurements of luminescence with 1 second integration times were taken before ligand
502 addition using a SpectraMax M5 microplate reader. For signaling curves, dilution series of native
503 relaxin-2 (R&D Systems), were added to the cells and luminescence measurements were taken
504 30 minutes after relaxin-2 addition. For measurements of RXFP1 basal signaling versus agonist-
505 induced signaling, vehicle control (PBS + 0.1% bovine serum albumin), relaxin-2 at 50 nM final
506 concentration, or ML-290 at 490 nM final concentration were added to the cells and the
507 luminescence measurement was taken after 30 minutes. The maximum signaling response of
508 relaxin-2 or ML-290 at wild type RXFP1 for each experiment was normalized to 100%, and the
509 percentages were plotted using GraphPad Prism.
510
511 Expression of RXFP1-miniG_s
512 RXFP1-miniG_s constructs for protein purification were expressed in inducible Expi293F
513 tetR cells[48] (Thermo Fischer Scientific). Stable cells lines were generated for RXFP1-miniG_s399-

514 20res by plating Expi293F tetR cells adherently in DMEM with 10% (v/v) FBS. Cells were
515 transfected with linearized RXFP1-miniG_s399-20res pcDNA-Zeo-tetO DNA using
516 Lipofectamine, according to the manufacturer's protocols (Thermo Fischer Scientific). Stable
517 integrations were selected using 200 μg/mL Zeocin. After selection, polyclonal adherent RXFP1-
518 miniG_s399-20res cells were readapted to suspension culture in antibiotic-free Expi293 media
519 (Thermo Fischer Scientific), then maintained in Expi293 media with 10 μg/mL Zeocin. Stable
520 RXFP1-miniG_s399-20res Expi293F tetR cells were expanded for protein expression in
521 antibiotic-free Expi293 media and induced with 4 μg/mL doxycycline, 0.4% glucose, and 5 mM
522 sodium butyrate. After 48 hours of induction, cells were harvested by spinning at 4,000 xg for 30
523 minutes at 4°C. The pellet was flash frozen in liquid nitrogen and stored at -80°C.
524
525 Purification of relaxin proteins
526 Single-polypeptide versions of relaxin-2 were used for RXFP1 complex purifications and
527 flow cytometry binding assays. These single-polypeptide relaxin-2 constructs utilize linkers to
528 connect relaxin-2's B-chain and A-chain. Sequences for these proteins can be found in **Table S8**
529 and their design and characterization are described in detail elsewhere[17].
530 Single-chain relaxin proteins were purified as previously described[17]. Briefly, His-tagged
531 single-chain relaxin-2 (SE001) was expressed as a secreted protein from inducible Expi293F tetR
532 cells. SE001 was purified from cell supernatants after 5 days of induction using Nickel Excel
533 resin (GE Healthcare), followed by size exclusion chromatography with a Sephadex S200
534 column (GE Healthcare). Purified SE001, in 30 mM MES pH 6.5 and 300 mM sodium chloride,
535 was aliquoted, flash frozen in liquid nitrogen, and stored at -80°C until purifications of the
536 RXPF1 complex.
537 Single-chain relaxin-2 fused to an antibody IgG1 Fc fragment (SE301) was expressed as
538 described above in Expi293F tetR cells. SE301 was purified using Protein G resin (GE
539 Healthcare) in 20 mM HEPES pH 7.5, 150 mM sodium chloride. Purified SE301 was aliquoted,
540 flash frozen in liquid nitrogen, and stored at -80°C until flow cytometry binding assays.
541
542 Purification of the RXFP1–G_s complex
543 Purification of the RXFP1–G_s complex used a cell pellet from a 2 L induction of the
544 RXFP1-miniG_s399-20res Expi293F tetR stable cell line, and the purified proteins SE001 (single-

545 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
546 purified according to previously published protocols[16]. The RXFP1-miniG_s399-20res cell pellet
547 was lysed through osmotic shock by stirring in 250 mL cold Lysis buffer containing 20 mM
548 HEPES pH 7.5, 2 mM magnesium chloride, 1 μL benzonase (Sigma Aldrich), 1 protease
549 inhibitor tablet (Thermo Fisher Scientific), and 100 nM SE001. Once stirring, iodoacetamide was
550 added at a final concentration of 2 mg/mL. Lysed cells were centrifuged at 50,000 xg for 30
551 minutes at 4°C and the supernatant was decanted. Membranes were homogenized with a glass
552 dounce (Thermo Fisher Scientific) in 270 mL Complexation buffer containing 20 mM HEPES
553 pH 7.5, 350 mM sodium chloride, 20% (v/v) glycerol, 1 μL benzonase, and 1 protease inhibitor
554 tablet. After homogenization, 960 μg $G\beta_1\gamma_2$, 740 μg Nb35-His-PrC (1:2:4 ratio of
555 RXFP1:$G\beta_1\gamma_2$:Nb35-His-PrC), 100 nM SE001, and 0.2 U apyrase (New England Biolabs) were
556 added to the membranes and the solution was stirred for 1 hour at 4°C. Next, 30 mL of 10X
557 Detergent buffer [10% (w/v) lauryl maltose neopentyl glycol (L-MNG; Anatrace) and 1% (w/v)
558 cholesterol hemisuccinate (CHS; Anatrace)] were added to the solution, dounce homogenization
559 was repeated, and the solution was stirred for 2 hours at 4°C to extract the RXFP1–G_s complex
560 from the membrane. After 2 hours, the solution was centrifuged at 50,000 xg for 30 minutes at
561 4°C and the supernatant was filtered with a glass fiber prefilter (Millipore). Calcium was added
562 to the solubilized membranes at a final concentration of 2 mM and the solution was loaded by
563 gravity flow over 3 mL M1 anti-FLAG resin (In house). This step of the purification was based
564 on the N-terminal FLAG tag of RXFP1-miniG_s399-20res. The M1 resin was equilibrated with
565 Wash buffer 1 [0.1% (w/v) L-MNG, 0.01% (w/v) CHS, 350 mM sodium chloride, 20 mM
566 HEPES pH 7.5, 2 mM calcium chloride, and 100 nM SE001]. After loading the solubilized
567 membrane fraction, the column was subsequently washed with 50 mL each of Wash buffer 1,
568 Wash buffer 2 [0.1% (w/v) L-MNG, 0.01% (w/v) CHS, 350 mM sodium chloride, 20 mM
569 HEPES pH 7.5, 2 mM calcium chloride, 2 mM magnesium chloride, 5 mM adenosine 5'-
570 triphosphate magnesium salt, and 100 nM SE001], and Wash buffer 3 [0.01% (w/v) L-MNG,
571 0.001% (w/v) CHS, 350 mM sodium chloride, 20 mM HEPES pH 7.5, 2 mM calcium chloride,
572 and 100 nM SE001]. M1 resin was eluted with 0.01% (w/v) L-MNG, 0.001% (w/v) CHS, 350
573 mM sodium chloride, 20 mM HEPES pH 7.5, 0.5 mg/mL FLAG peptide (GenScript), and 100
574 nM SE001, then 2 mM calcium chloride was added to the eluted protein. Next, 0.5 mL anti-
575 Protein C resin (In house) was equilibrated with Protein C wash buffer [0.005% (w/v) L-MNG,

576 0.0005% (w/v) CHS, 350 mM sodium chloride, 20 mM HEPES pH 7.5, 2 mM calcium chloride,
577 and 100 nM SE001]. The M1 elution with 2 mM calcium chloride was loaded onto the anti-
578 Protein C resin by gravity flow, washed with 10 mL Protein C wash buffer, and eluted with
579 0.005% (w/v) L-MNG, 0.0005% (w/v) CHS, 350 mM sodium chloride, 20 mM HEPES pH 7.5,
580 0.5 mg/mL Protein C peptide (GenScript), and 100 nM SE001. This step of the purification was
581 based on the Protein C tag of Nb35-His-PrC. The elution was concentrated with a 3 kDa
582 molecular weight cutoff centrifugal concentrator (Millipore) and loaded onto a Sephadex S200
583 column (GE Healthcare) in SEC buffer [0.005% (w/v) L-MNG, 0.0005% (w/v) CHS, 350 mM
584 sodium chloride, 20 mM HEPES pH 7.5, and 100 nM SE001]. Peak fractions from size
585 exclusion were concentrated with a 3 kDa molecular weight cutoff centrifugal concentrator and
586 either stored overnight at 4°C prior to grid freezing or immediately used to freeze cryo-EM grids.
587
588 Cryo-EM grid preparation and data collection
589 Cryo-EM grids were prepared using QUANTIFOIL® holey carbon grids (400-mesh,
590 copper, R1.2/1.3; Electron Microscopy Sciences). Grids were washed with ethyl acetate (Sigma
591 Aldrich) and then glow-discharged at -20 mA for 60 seconds with a Pelco Easiglow. 3 μL of
592 sample at 0.2 to 0.4 mg/mL was applied to the grids. Grids were plunge-frozen in liquid ethane
593 using a Mark IV Vitrobot (Thermo Fisher Scientific) at 10°C and 100% humidity with a 10
594 second wait time, 3 to 4 second blot time, and blot force of 15.
595 Cryo-EM data for the RXFP1–G_s complex were collected in four separate sessions. Grids
596 were imaged with a Titan Krios microscope (Thermo Fisher Scientific) at 300 kV with a Gatan
597 BioQuantum GIF/K3 direct electron detection camera in counting mode. Movies were collected
598 with 50 frames each at 81,000x, corresponding to 1.06 Å per pixel, and a total dose of around 52
599 electrons per $Å^2$. Defocus values ranged from -0.8 to -2.3 μm. A total of 13,457 movies were
600 collected across the four data collection sessions.
601
602 Cryo-EM data processing
603 Cryo-EM data was collected in four separate session and initially processed individually,
604 then particle stacks were merged to generate the final maps. Motion correction was carried out
605 with MOTIONCOR2[49] and CTF parameters were estimated with CTFFIND-4.1[50]. Particles were
606 picked in RELION 3.1[51] using Laplacian of Gaussian autopicking. A map from a previous small

607 dataset collected on a Talos Arctica microscope and processed in RELION was used as an initial
608 model. Particles were extracted with a box size large enough to include the entire complex (288
609 pixels). Multiple rounds of 2D classifications in RELION used a circular mask of 150 Å around
610 only the micelle and G proteins as the first step of particle sorting for both 3D reconstructions.
611 Masking of the micelle and G proteins was used to initially sort high quality particles from
612 contamination because 2D classifications attempted for the entire particle showed weak signal
613 for the ectodomain, likely indicating a region of higher flexibility. It became evident at this stage
614 of data processing that the datasets showed a preferred orientation for one of the side views of
615 the RXFP1–G_s complex. Processing 2D classifications of the preferred view separately from
616 other views of the complex resulted in a larger number of particles from different orientations in
617 the final particle stacks.
618 For the map of the 7TMs, masked 3D classifications of the 7TM domain bound to G
619 proteins were performed in RELION. Following these classifications, iterative 3D refinements
620 (using a mask of the micelle and G proteins) and Bayesian particle polishing steps were carried
621 out in RELION. At this stage, particle stacks from different datasets were joined together,
622 followed by additional 2D classifications with a 150 Å circular mask and masked 3D
623 classifications. Finally, 3D focused classifications without alignments were carried out using
624 masks for either the ECLs, TM helices, or helix8. These particle stacks were imported into
625 cryoSPARC v3.1.0[39] and used with Non-uniform Refinement (New). The cryoSPARC refined
626 maps were post-processed in DeepEMhancer[52] using the half-maps as input. This data processing
627 workflow is described in **Fig. S2**.
628 For the map of the full-length RXFP1–G_s complex, rounds of unmasked 3D
629 classifications were used in RELION after the 2D classifications described above. Next, the
630 micelle and G proteins were subtracted from the particles and 3D focused classifications without
631 alignments were carried out using a mask on the entire ectodomain of RXFP1. 3D refinements
632 with a mask of the entire complex and iterative Bayesian particle polishing steps were then
633 performed in RELION. Following particle polishing, particle stacks from different datasets were
634 combined and 3D classifications with a mask around the entire complex were performed in
635 RELION, followed by 3D refinement in RELION with a mask around the entire complex.
636 Finally, local resolution estimation and filtering were performed in RELION on the final maps.
637 This data processing workflow is described in **Fig. S3**. Continuous heterogeneity in the final

638 particle stack was visualized using 3D variability analysis in cryoSPARC and shown in **Movie**
639 **S1**.
640
641 Model building and refinement
642 A combined focused map for the RXFP1 7TM domain bound to G proteins was
643 generated using the DeepEMhancer post-processed maps as inputs in Phenix[53]. Model building
644 for the RXFP1 7TM domain bound to heterotrimeric G_s and Nb35 was performed using
645 DeepEMhancer focused and combined focused cryo-EM maps in *Coot*[54]. An initial model for the
646 RXFP1 7TM domain (Gly395 to Gly709) was generated using trRosetta[55] and manually fit into
647 the DeepEMhancer focused cryo-EM map (for ECLs) in Chimera[56]. Intracellular and
648 extracellular loops and the hinge region were removed from the trRosetta model and manually
649 rebuilt in *Coot*. ECL2 residues Lys554–Tyr557 and hinge region residues Glu400–Leu403 were
650 manually built with the coordinates of the human RXFP1 AlphaFold2 model[38] overlayed in
651 *Coot*. Initial models for miniG_s399, Gβγ, and Nb35 were generated in MODELLER[57] using
652 coordinates from PDB ID 6GDG[58]. All models were refined with Phenix real-space refinement[53]
653 using secondary structure restraints against the DeepEMhancer combined focused map. Statistics
654 for the final model were evaluated using MolProbity in a Phenix comprehensive validation
655 (cryo-EM) job that used the map from cryoSPARC v3.1.0 Non-uniform Refinement (New) and
656 the final model as inputs. Figures were prepared using PyMOL and ChimeraX[59]. Structural
657 biology programs used in this work, other than cryoSPARC, were compiled and configured by
658 SBGrid[60]. Refinement statistics are present in **Table S1** and representative images of cryo-EM
659 map and model quality are shown in **Fig. S8**.
660
661 Evolutionary coupling analysis
662 For comparing the constructed model to the strongest evolutionarily coupled pairs, we
663 first downloaded the Uniprot canonical sequence for Q9HBX9 for residues 405-689 and then
664 used the Jackhmmer software suite[61] to build multiple sequence alignments across multiple
665 bitscore thresholds based on the June 2019 download of the Uniref100 database[62]. We then chose
666 an alignment with 352,511 sequences, with 90% of columns consisting of fewer than 30% gaps,
667 to move forward with in analysis. We used the EVcouplings v0.0.5 software, available at
668 <github.com/debbiemarkslab>, to identify evolutionary couplings (ECs) for this alignment.

669 To incorporate the LRR region, we ran the Q9HBX9 sequence for residues 120-757 in
670 Jackhmmer, against the 02/2021 Uniref100 database for normalized bitscores between 0.1 and
671 0.9. We chose an alignment with 8.145 sequences with 86% of columns consisting of fewer than
672 30% gaps, and input it to v0.1.1 EVcouplings software. The strongest couplings were ranked
673 based on assigning probabilities from a logistic regression model new to v0.1.1. This version of
674 the pipeline was also used for plotting all contact maps.
675
676 Crosslinking mass spectrometry
677 The RXFP1–G_s complex used for CLMS was prepared as described above and
678 crosslinking reactions were carried out the following day, after storage overnight at 4°C. CLMS
679 was performed as previously described[41]. Briefly, crosslinking reactions were carried out for 1 h
680 at room temperature in 100 mM MOPS Buffer, pH 6.5 with 50 mM EDC, ~24 mM EDDA
681 linker, and 20 mM sulfo-NHS. Reactions were quenched with hydroxylamine to a final
682 concentration of 100 mM. Samples were reduced for 1 h in 2% SDS and 5 mM TCEP followed
683 by alkylation with 10 mM iodoacetamide in the dark for 30 min and quenching with 5 mM DTT
684 for 15 min. Samples were then processed with the SP3[63] method and digested with trypsin
685 (Promega) at 1:25 enzyme:substrate ratio overnight at 37°C. Digested peptides were acidified
686 with 10% formic acid to pH ~2 and desalted using stage tips with Empore C18 SPE Extraction
687 Disks (3M) and dried under vacuum.
688 Sample was reconstituted in 5% formic acid (FA)/5% acetonitrile and analyzed in the
689 Orbitrap Eclipse Mass Spectrometer (Thermo Fischer Scientific) coupled to an EASY-nLC 1200
690 (Thermo Fisher Scientific) ultra-high pressure liquid chromatography (UHPLC) pump, as well as
691 a high-Field Asymmetric waveform Ion Mobility Spectrometry (FAIMS) interface. Peptides
692 were separated on an in-house packed 100 μm inner diameter column packed with 35 cm of
693 Accucore C18 resin (2.6 um, 150 Å, ThermoFisher), using a gradient consisting of 5–35%
694 (ACN, 0.125% FA) over 135 min at ∼500 nL/min. The instrument was operated in data-
695 dependent mode. FTMS1 spectra were collected at a resolution of 120K, with an automated gain
696 control (AGC) target of 5×10^5, and a max injection time of 50 ms. The most intense ions were
697 selected for MS/MS for 1s in top-speed mode, while switching among three FAIMS
698 compensation voltages (CV): −40, –60, and –80 V in the same method. Precursors were filtered
699 according to charge state (allowed $3 <= z <= 7$), and monoisotopic peak assignment was turned

700 on. Previously interrogated precursors were excluded using a dynamic exclusion window (60 s ±
701 7 ppm). MS2 precursors were isolated with a quadrupole mass filter set to a width of 0.7 m/z and
702 analyzed by FTMS2, with the Orbitrap operating at 30K resolution, an AGC target of 100K, and
703 a maximum injection time of 150 ms. Precursors were then fragmented by high-energy collision
704 dissociation (HCD) at a 30% normalized collision energy.
705 Mass spectra were processed and searched using the PIXL search engine[41]. The sequence
706 database contained proteins identified at 1% FDR in non-cross-linked Comet[64] search. For PIXL
707 search, precursor tolerance was set to 15 ppm and fragment ion tolerance to 10 ppm. Methionine
708 oxidation was set as a variable modification in addition to mono-linked mass of +130.110613 for
709 EDDA. Crosslinked peptides were searched assuming zero-length (-18.010565) and EDDA
710 crosslinker +112.100048. Crosslinked searches considered 60 protein sequences to ensure
711 sufficient statistics for FDR estimation. Matches were filtered to 1% FDR on the unique peptide
712 level using linear discriminant features as previously described[41].
713
714 Docking
715 HADDOCK[44] was used to dock the relaxin-2–LRR interaction. Docking used the human
716 relaxin-2 X-ray crystal structure (PDB ID: 6RLX)[65] and a model of the LRRs from residues 104-
717 391 of the AlphaFold2[38] prediction for human RXFP1. Residues from CLMS studies that were
718 part of crosslinks between relaxin-2 and RXFP1 were used as active restraints in the docking run.
719 These CLMS residues were Glu14$^{B\text{-chain}}$ of relaxin-2, and Glu206, Glu299, and Glu351 of
720 RXFP1. Residues identified to be important for relaxin-2 binding from published mutations in
721 radioligand binding assays were also used as active restraints[32,42,43]. These residues included
722 Arg13$^{B\text{-chain}}$, Arg17$^{B\text{-chain}}$, and Ile20$^{B\text{-chain}}$ of relaxin-2 and Trp202, Ile204, Leu226, Asp253,
723 Glu255, Glu299, and Asp301 of RXFP1. The resulting docking models were analyzed according
724 to the HADDOCK scoring function and fit into the low resolution cryo-EM map of the RXFP1
725 ectodomain in ChimeraX[59].
726
727 Flow cytometry binding assay
728 A flow cytometry assay was used to measure the binding of an Fc-tagged relaxin-2
729 protein, SE301[17], to Expi293F cells transfected with human RXFP1 or empty pcDNA-Zeo-tetO
730 vector. Expi293F tetR cells were grown in Expi293 media and transfected using FectoPRO

731 (Polyplus), according to the manufacturer's protocols. The cells were enhanced 24 hours post-
732 transfection with 0.4% glucose and induced 48 hours post-transfection with 4 μg/mL
733 doxycycline and 5 mM sodium butyrate. After 24 hours of induction, cells were harvested by
734 spinning at 200 xg for 5 minutes at 4°C and washed once with HBS with 1% (v/v) FBS and 2
735 mM calcium chloride (Buffer B). Cells were plated into a V-bottom 96-well plate (Corning) at
736 100,000 cells/well and blocked by incubation in Buffer B for 30 minutes at 4°C. After blocking,
737 cells were centrifuged at 200 xg for 5 minutes at 4°C, resuspended in 100 μL of Buffer B
738 containing 500 nM SE301, and incubated for 1 hour at 4°C. Cells were then centrifuged at 200
739 xg for 5 minutes at 4°C, washed twice with 200 μL Buffer B, and resuspended in 100 μL Buffer
740 B containing 100 nM M1 anti-FLAG antibody labeled with Alexa Fluor 488 (In house) and
741 Alexa Fluor 647 anti-human IgG Fc (BioLegend) diluted 1:100 (v/v). Cells were incubated in
742 secondary antibodies for 30 minutes at 4°C, washed once with 200 μL Buffer B, and
743 resuspended in 100 μL Buffer B for flow cytometry.
744 Samples were analyzed on a CytoFLEX flow cytometer (Beckman Coulter) and gated
745 according to plots of FSC-A/SCA-A, FSC-A/FSC-H, and receptor expression according to Alexa
746 Fluor 488 M1 anti-FLAG antibody binding (**Fig. S11**). The receptor expression gate was drawn
747 by comparing empty vector and wild type RXFP1-transfected cells. Approximately 500
748 events/sample were collected from cells expressing receptor for human RXFP1-transfected cells
749 or post-FSC-A/FSC-H gating for empty vector-transfected cells. The data were plotted and
750 analyzed in FlowJo and GraphPad Prism. For the binding assay in **Fig. 4**, the E206A and wild
751 type RXFP1 samples expressed very similarly, so cell surface expression and SE301 binding
752 were plotted separately. For comparing the binding of multiple constructs in **Fig. S5**, ratios of
753 SE301 binding to receptor expression were calculated in order to normalize for the differences in
754 RXFP1 construct expression levels.
755 A flow cytometry competition binding assay was used to measure SE001 binding to
756 human RXFP1-expressing Expi293F cells. Expi293F cells were transfected, harvested, and
757 blocked as stated above. After blocking in Buffer B, cells were incubated with 200 nM SE301
758 (Fc-tag) and increasing concentrations of SE001 (His-tag). After 1 hour of incubation at 4°C, the
759 reaction was terminated by centrifugation at 200 x g for 5 minutes at 4°C and cells were washed
760 twice with 200 μL Buffer B. In order to detect SE301 and measure receptor expression, cells

761 were then stained with Alexa Fluor 488 M1 anti-FLAG and Alexa Fluor 647 anti-human IgG Fc
762 and analyzed by flow cytometry as stated above. Data points were calculated as a percentage of
763 wild type RXFP1 SE301 binding and plotted in GraphPad Prism (**Fig. S1f**).
764
765 Molecular dynamics
766 The initial model was built from the cryo-EM structure reported here. The missing
767 segments in ECL2, ICL3 and TM6 were generated by MODELLER v9.15[66]. PACKMOL-
768 Memgen[67] was used to assign the side-chain protonation states and embed the models in a
769 bilayer of POPC lipids. The systems were solvated in a periodic box of explicit water and
770 neutralized with 0.15 M of Na^+ and Cl^- ions. We used the Amber ff14SB[68] and lipid 14[69] force
771 fields, the TIP3P water model[70] and the Joung-Cheatham ion parameters[71]. For the simulations of
772 7TM deactivation, a Na^+ ion was placed at the conserved Na^+-binding site (between Asp451$^{2.50}$
773 and Ser495$^{3.35}$). For the simulations of autoactivation, the truncated AlphaFold2[38] model was
774 used to build the truncated 7TM and ^{half}LRRs-7TM forms in which Asp451$^{2.50}$ was protonated.
775 After energy minimization, all-atom MD simulations were carried out using Gromacs
776 5.1[72] patched with the PLUMED 2.3 plugin[73]. The LINCS algorithm[74] was applied to constrain
777 bonds involving hydrogen atoms, allowing for a time step of 2 fs. Each system was gradually
778 heated to 310 K and pre-equilibrated during 10 ns of brute-force MD in the *NPT*-ensemble. The
779 replica exchange with solute scaling (REST2)[75] technique was used to enhance the
780 conformational sampling. A total of 64 replicas of simulations were performed in the *NVT*
781 ensemble. REST2 is a type of Hamiltonian replica exchange simulation scheme. Besides the
782 original simulation, many replicas of the same system were simulated simultaneously. The
783 additional replicas have modified free energy surfaces, in which the energy barriers are easier to
784 cross than in the original simulation system. By frequently swapping the replicas and the original
785 system during the MD, the simulations "travel" on different free energy surfaces and easily visit
786 various conformational zones. Finally, only the samples on the original free energy surface are
787 collected. The additional replicas are artificial to ease barrier crossing, which are discarded after
788 the simulations. REST2, in particular, modifies the free energy surfaces by scaling (reducing) the
789 force constants of the "solute" molecules in the simulation system. In this case, the protein was
790 considered as "solute"–the force constants of its van der Waals, electrostatic and dihedral terms
791 were subject to scaling–in order to facilitate the conformational changes. The scaling factors

792 were generated using the Patriksson-van der Spoel approach[76] and effective temperatures ranging
793 from 310 K to 1000 K. Exchange between replicas was attempted every 1000 simulation steps.
794 This setup resulted in an average exchange probability of ~40% for the 7TM and ~20% for the
795 ^{half}LRRs-7TM systems, respectively. We performed 80 ns × 64 replicas of REST2 MD in the
796 *NVT* ensemble for each system. The first 30 ns were discarded for equilibration. The simulation
797 trajectories on the original unmodified free energy surface was reassembled and analyzed.
798

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. Erlandson, S. C., Wang, J., Jiang, H., Rockman, H. A. & Kruse, A. C. Engineering and characterization of a long half-life relaxin receptor RXFP1 agonist. *bioRxiv* (2022).
18. Hilger, D., Masureel, M. & Kobilka, B. K. Structure and dynamics of GPCR signaling complex. *Nat. Struct. Mol. Biol.* **25**, 4–12 (2018).
19. Erlandson, S. C., McMahon, C. & Kruse, A. C. Structural basis for G protein-coupled receptor signaling. *Annu. Rev. Biophys.* **47**, 1–18 (2018).
20. Ballesteros, J. A. & Weinstein, H. Integrated methods for the construction of three-

845 dimensional models and computational probing of structure-function relations in G
846 protein-coupled receptors. *Methods Neurosci.* **25**, 366–428 (1995).
847 21. Wang, J., Hua, T. & Liu, Z.-J. Structural features of activated GPCR signaling complexes.
848 *Curr. Opin. Struct. Biol.* **63**, 82–89 (2020).
849 22. Venkatakrishnan, A. J. *et al.* Molecular signatures of G-protein-coupled receptors. *Nature*
850 **494**, 185–194 (2013).
851 23. Ring, A. M. *et al.* Adrenaline-activated structure of β2-adrenoceptor stabilized by an
852 engineered nanobody. *Nature* **502**, 575–579 (2013).
853 24. Wingler, L. M., McMahon, C., Staus, D. P., Lefkowitz, R. J. & Kruse, A. C. Distinctive
854 activation mechanism for angiotensin receptor revealed by a synthetic nanobody. *Cell*
855 **176**, 479-490.e12 (2019).
856 25. Diepenhorst, N. A. *et al.* Investigation of interactions at the extracellular loops of the
857 relaxin family peptide receptor 1 (RXFP1). *J. Biol. Chem.* **289**, 34938–34952 (2014).
858 26. Lin, X. *et al.* Structural basis of ligand recognition and self-activation of orphan GPR52.
859 *Nature* **579**, 152–157 (2020).
860 27. Brüser, A. *et al.* The activation mechanism of glycoprotein hormone receptors with
861 implications in the cause and therapy of endocrine diseases. *J. Biol. Chem.* **291**, 508–520
862 (2016).
863 28. Schulze, A. *et al.* The intramolecular agonist is obligate for activation of glycoprotein
864 hormone receptors. *FASEB J.* **34**, 11243–11256 (2020).
865 29. Jiang, X., Dias, J. A. & He, X. Structural biology of glycoprotein hormones and their
866 receptors: Insights to signaling. *Mol. Cell. Endocrinol.* **382**, 424–451 (2014).
867 30. Martin, A. L., Steurer, M. A. & Aronstam, R. S. Constitutive activity among orphan class-
868 A G protein coupled receptors. *PLoS One* **10**, e0138463 (2015).
869 31. Sethi, A. *et al.* The complex binding mode of the peptide hormone H2 relaxin to its
870 receptor RXFP1. *Nat. Commun.* **7**, 11344 (2016).
871 32. Büllesbach, E. E. & Schwabe, C. The trap-like relaxin-binding site of the leucine-rich G-
872 protein-coupled receptor 7. *J. Biol. Chem.* **280**, 14051–14056 (2005).
873 33. Marks, D. S. *et al.* Protein 3D structure computed from evolutionary sequence variation.
874 *PLoS One* **6**, (2011).
875 34. Marks, D. S., Hopf, T. A. & Sander, C. Protein structure prediction from sequence
876 variation. *Nat. Biotechnol.* **30**, 1072–1080 (2012).
877 35. Sjodt, M. *et al.* Structure of the peptidoglycan polymerase RodA resolved by evolutionary
878 coupling analysis. *Nature* **556**, 118–121 (2018).
879 36. Zheng, S. *et al.* Structure and mutagenic analysis of the lipid II flippase MurJ from
880 Escherichia coli. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 6709–6714 (2018).
881 37. Toth-Petroczy, A. *et al.* Structured states of disordered proteins from genomic sequences.
882 *Cell* **167**, 158-170.e12 (2016).
883 38. Tunyasuvunakool, K. *et al.* Highly accurate protein structure prediction for the human
884 proteome. *Nature* **596**, 590–596 (2021).
885 39. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: Algorithms for
886 rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).
887 40. Halls, M. L. *et al.* Multiple binding sites revealed by interaction of relaxin family peptides
888 with native and chimeric relaxin family peptide receptors 1 and 2 (LGR7 and LGR8). *J.*
889 *Pharmacol. Exp. Ther.* **313**, 677–687 (2005).
890 41. Mintseris, J. & Gygi, S. P. High-density chemical cross-linking for modeling protein

interactions. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 93–102 (2020).
42. 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).
43. 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).
44. 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).
45. Faust, B. *et al.* Autoantibody and hormone activation of the thyrotropin G protein-coupled receptor. *bioRxiv* (2022).
46. 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).
47. Hu, X. *et al.* Structural Insights into the Activation of Human Relaxin Family Peptide Receptor 1 by Small-Molecule Agonists. *Biochemistry* **55**, 1772–1783 (2016).
48. 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).
49. 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.
50. Rohou, A. & Grigorieff, N. CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015).
51. Scheres, S. H. W. RELION: Implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* **180**, 519–530 (2012).
52. Sanchez-Garcia, R. *et al.* DeepEMhancer: a deep learning solution for cryo-EM volume post-processing. *Commun. Biol.* **4**, 1–8 (2021).
53. 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).
54. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 486–501 (2010).
55. Du, Z. *et al.* The trRosetta server for fast and accurate protein structure prediction. *Nat. Protoc.* **16**, 5634–5651 (2021).
56. Pettersen, E. F. *et al.* UCSF Chimera - A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
57. Webb, B. & Sali, A. Comparative protein structure modeling using MODELLER. *Curr. Protoc. Bioinforma.* **54**, 5.6.1-5.6.37 (2016).
58. 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).
59. Pettersen, E. F. *et al.* UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci.* **30**, 70–82 (2021).
60. Morin, A. *et al.* Collaboration gets the most out of software. *Elife* **2**, e01456 (2013).
61. Potter, S. C. *et al.* HMMER web server: 2018 update. *Nucleic Acids Res.* **46**, W200–W204 (2018).
62. 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).
63. 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).
64. Eng, J. K., Jahan, T. A. & Hoopmann, M. R. Comet: An open-source MS/MS sequence database search tool. *Proteomics* **13**, 22–24 (2013).
65. 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).
66. 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).
67. 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).
68. 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).
69. Dickson, C. J. *et al.* Lipid14: The amber lipid force field. *J. Chem. Theory Comput.* **10**, 865–879 (2014).
70. 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).
71. 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).
72. Van Der Spoel, D. *et al.* GROMACS: Fast, flexible, and free. *J. Comput. Chem.* **26**, 1701–1718 (2005).
73. 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).
74. 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).
75. 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).
76. Patriksson, A. & van der Spoel, D. A temperature predictor for parallel tempering simulations. *Phys. Chem. Chem. Phys* **10**, 2073–2077 (2008).