Check for updates

OPEN

Ligand recognition and G-protein coupling selectivity of cholecystokinin A receptor

Qiufeng Liu[1,10], Dehua Yang [1,2,3,10], Youwen Zhuang[1,10], Tristan I. Croll[4], Xiaoqing Cai[3], Antao Dai[3], Xinheng He [1,2], Jia Duan[1,2], Wanchao Yin [1], Chenyu Ye[5], Fulai Zhou [1], Beili Wu [1,2,6,7], Qiang Zhao [2,7,8], H. Eric Xu [1,2,6 ✉], Ming-Wei Wang[1,2,3,5,6,9 ✉] and Yi Jiang [1,2 ✉]

Cholecystokinin A receptor (CCK$_A$R) belongs to family A G-protein-coupled receptors and regulates nutrient homeostasis upon stimulation by cholecystokinin (CCK). It is an attractive drug target for gastrointestinal and metabolic diseases. One distinguishing feature of CCK$_A$R is its ability to interact with a sulfated ligand and to couple with divergent G-protein subtypes, including G$_s$, G$_i$ and G$_q$. However, the basis for G-protein coupling promiscuity and ligand recognition by CCK$_A$R remains unknown. Here, we present three cryo-electron microscopy structures of sulfated CCK-8-activated CCK$_A$R in complex with G$_s$, G$_i$ and G$_q$ heterotrimers, respectively. CCK$_A$R presents a similar conformation in the three structures, whereas conformational differences in the 'wavy hook' of the Gα subunits and ICL3 of the receptor serve as determinants in G-protein coupling selectivity. Our findings provide a framework for understanding G-protein coupling promiscuity by CCK$_A$R and uncover the mechanism of receptor recognition by sulfated CCK-8.

C holecystokinin (CCK), one of the earliest discovered gastrointestinal hormones, participates in gallbladder contraction and pancreatic enzyme secretion. It also acts as a neurotransmitter and is extensively distributed throughout the nervous system[1]. Selective cleavage of the CCK precursor produces a series of bioactive isoforms of different lengths, with CCK-58, CCK-33, CCK-22 and CCK-8 comprising the major peptide fragments in humans. However, the carboxy-terminal octapeptide CCK-8 (DYMGWMDF) is well conserved across species and is the smallest form that retains the full range of biological actions[2], mediated by two CCK receptor subtypes (CCK$_A$R and CCK$_B$R), which are present throughout the central nervous system and the gut. CCK$_A$R is primarily expressed in the alimentary tract, while CCK$_B$R is mainly found in the brain and the stomach[3]. CCK$_A$R has an ~500-fold higher affinity to CCK with a sulfated tyrosine, whereas CCK$_B$R discriminates poorly between sulfated and non-sulfated CCK[4].

CCK regulates appetite and food intake primarily through CCK$_A$R on the vagal afferent neurons[5-8], making CCK$_A$R an attractive therapeutic target for obesity. However, the development of drugs against CCK$_A$R is challenging, partly due to limited efficacy and safety concerns. Although several drug candidates are undergoing clinical trials, none has been approved so far[9,10]. Extensive efforts have been made to elucidate the mechanism of agonism at CCK$_A$R through mutagenesis studies based on modeled receptor structures[11-15], but the lack of precise structural information largely impedes our understanding of the molecular details of ligand recognition and receptor activation, and thus the drug discovery targeting CCK$_A$R.

Most G-protein-coupled receptors (GPCRs) are known to couple with a specific subtype of G protein to elicit intracellular signal transduction[16-21]. There are four G-protein subtypes—stimulatory G protein (G$_s$), inhibitory G proteins (G$_i$), G$_q$ and G$_{12/13}$—participating in signaling pathways involving cyclic adenosine monophosphate (G$_s$ and G$_i$), calcium (G$_q$) and small G protein (G$_{12/13}$). A number of GPCR–G protein complex structures reported recently reveal that the primary determinants of G-protein coupling selectivity reside in the C-terminal α5-helix of the Gα subunit and relative outward movement of transmembrane helix 6 (TM6)[22,23]. However, CCK$_A$R is different from most GPCRs as a result of its ability to couple with several G-protein subtypes. Activation of CCK$_A$R elicits a diversified G-protein coupling pattern[24], predominantly G$_q$ (ref. [25]), but G$_s$ (ref. [26]), G$_i$ (refs. [25,27]) and G$_{13}$ (refs. [28,29]) all play roles in CCK$_A$R signaling. This unique feature makes CCK$_A$R an ideal model to study G-protein selectivity and promiscuity (Fig. 1a). Here, we report three cryo-EM structures of sulfated CCK-8-activated CCK$_A$R in complex with heterotrimeric G$_q$, G$_s$ or G$_i$ protein, respectively. These structures reveal the unique binding mode in ligand recognition and the structural determinants responsible for the G-protein selectivity and promiscuity of CCK$_A$R.

Overall structures of CCK$_A$R coupled to different G proteins

The structures of sulfated CCK-8-bound CCK$_A$R in complex with G$_q$, G$_s$ or G$_i$ heterotrimers were determined by single-particle cryo-EM at global resolutions of 2.9 Å, 3.1 Å and 3.2 Å, respectively (Fig. 1b–g, Supplementary Figs. 1–3 and Supplementary Table 1). Sulfated CCK-8 (DY^{SO3H}MGMWDF-NH$_2$), the highest-affinity natural ligand of CCK$_A$R[4], was used to assemble the CCK$_A$R–G protein

[1]The CAS Key Laboratory of Receptor Research, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai, China. [2]University of Chinese Academy of Sciences, Beijing, China. [3]The National Center for Drug Screening, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai, China. [4]Department of Haematology, Cambridge Institute for Medical Research, University of Cambridge, Cambridge, UK. [5]School of Pharmacy, Fudan University, Shanghai, China. [6]School of Life Science and Technology, ShanghaiTech University, Shanghai, China. [7]CAS Center for Excellence in Biomacromolecules, Chinese Academy of Sciences, Beijing, China. [8]State Key Laboratory of Drug Research, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai, China. [9]School of Basic Medical Sciences, Fudan University, Shanghai, China. [10]These authors contributed equally: Qiufeng Liu, Dehua Yang, Youwen Zhuang. ✉e-mail: eric.xu@simm.ac.cn; mwwang@simm.ac.cn; yijiang@simm.ac.cn
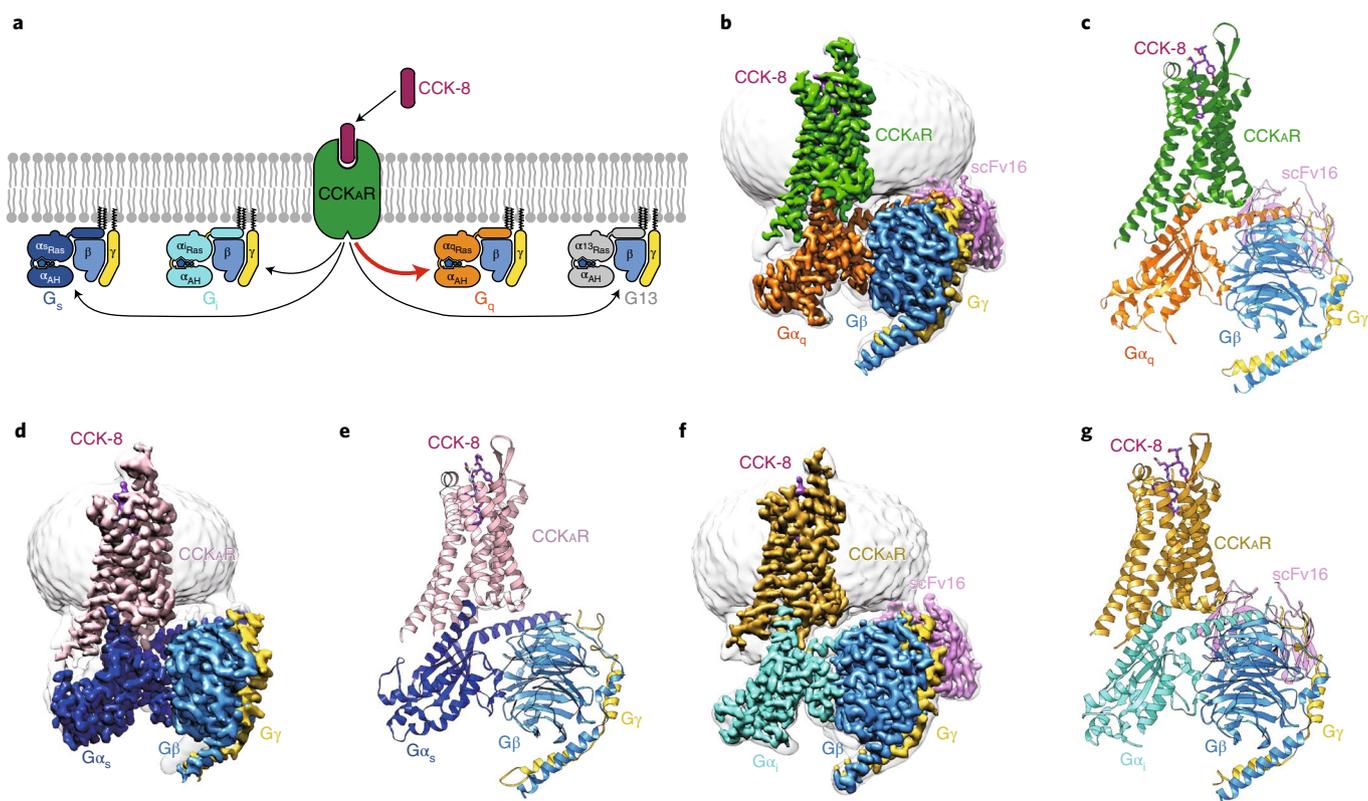
Fig. 1 | Cryo-EM structures of CCK$_A$R–G protein complexes. a, Schematic of G-protein coupling promiscuity of CCK$_A$R. **b–g**, Three-dimensional maps and models of the CCK-8–CCK$_A$R–G$_q$–scFv16 (**b,c**), CCK-8–CCK$_A$R–G$_s$ (**d,e**) and CCK-8–CCK$_A$R–G$_i$–scFv16 (**f,g**) complexes. CCK-8, magenta; G$_q$-coupled CCK$_A$R, green (**b,c**); G$_s$-coupled CCK$_A$R, pink (**d,e**); G$_i$-coupled CCK$_A$R, dark yellow (**f,g**); Gα$_q$, orange; Gα$_s$, blue; Gα$_i$, cyan; Gβ, light blue; Gγ, yellow; scFv16, light purple.

complexes. Three G-protein subtypes were engineered to stabilize the CCK$_A$R–G protein complexes (Supplementary Fig. 4). Gα$_q$ is chimerized by replacing its αN-helix with the equivalent region of Gα$_{i1}$ to facilitate scFv16 binding[30]. Gα$_s$ was modified based on the mini-Gα$_s$ that was used in the crystal structure determination of the G$_s$-coupled adenosine A$_{2A}$ receptor (A$_{2A}$R)[31]. Two dominant-negative (DN) mutations (G203A and A326S[32]) were introduced to Gα$_{i1}$, and corresponding DN mutations at equivalent sites of Gα$_s$ and Gα$_q$ were also introduced (Supplementary Fig. 4b). Unless otherwise specified, G$_q$, G$_s$ and G$_i$ refer to the respective engineered G proteins that are used in the CCK$_A$R structure determination.

The final structures of the CCK-8–CCK$_A$R–G protein complexes contain sulfated CCK-8 (residues D^{1P}–F^{8P}), Gα Ras-like domain, Gβγ subunits, scFv16 and the CCK$_A$R residues (E38$^{N\_term}$–F385$^{8.58}$, superscripts refer to Ballesteros–Weinstein numbering[33]). The majority of amino-acid side chains, including CCK-8, transmembrane domain (TMD), intracellular loops (ICLs 1–3) and extracellular loops (ECLs 1–3) were well resolved in the final models (Supplementary Fig. 5). Thus, the complex structures provide reliable details to study mechanisms of ligand recognition and G-protein coupling.

Globally, CCK$_A$R adopts similar overall conformations in all three structures, with an all-atom root-mean-square deviation (r.m.s.d.) of 0.84 for G$_q$/G$_s$-coupled receptors and 1.03 for G$_q$/G$_i$-coupled receptors. The structure of the CCK-8–CCK$_A$R–G$_q$ complex, which has the highest resolution at 2.9 Å, was used for detailed analysis and mechanistic evaluation of ligand recognition and receptor activation. The inactive and active structures of the closed homolog receptors (inactive: ghrelin receptor, PDB 6KO5[34]; active: neurotensin receptor 1 (NTSR1), PDB 6OS9[18]), all belonging to the

β-branch of the rhodopsin family, are applied for structural comparison. CCK$_A$R presents a fully active conformation, resembling the G$_i$-coupled NTSR1, displaying an ~9-Å outward movement of TM6 (measured at the Cα of the residue at position 6.27 in CCK$_A$R and the ghrelin receptor) and an ~4-Å inward shift of TM7 (Cα carbons of Y7.53) compared with the inactive ghrelin receptor (Extended Data Fig. 1a,b). Similar to the active NTSR1 complex, the conserved residues in the 'micro-switches' (PIF, ERY, CWxP and NPxxY) of CCK$_A$R display the conserved conformations observed in active GPCRs (Extended Data Fig. 1c).

Recognition of sulfated cholecystokinin

Sulfated CCK-8 occupies the orthosteric binding pocket composed of TM3, TM4, TM5–7 and ECL1–3 (Fig. 2 and Extended Data Figs. 2 and 3), with its C terminus inserting into the TMD bundle and the N terminus facing the extracellular vestibule (Fig. 2a). The binding pocket of CCK-8 is largely overlapped with that of other reported endogenous neuropeptides, such as neurotensin (NTS$_{8-13}$, PDB 6OS9[18]), angiotensin II (Ang II, PDB 6OS0[35]), orexin B (OXB, PDB 7L1U[36]) and arginine vasopressin (AVP, PDB 7DW9[37]). It is noteworthy that the extracellular sides of these neuropeptides undergo remarkable conformational shifts, while their intracellular parts converge in an approximately overlapped position at the bottom of the binding pocket (Extended Data Fig. 2).

It is of interest that the octapeptide CCK-8 almost completely occupies the polypeptide-binding pocket, structurally supporting the fact that it is the smallest active form of CCK isoforms. The binding modes of CCK-8 are highly conserved in all three CCK$_A$R–G protein complexes (all-atom r.m.s.d. of 0.71 for CCK-8
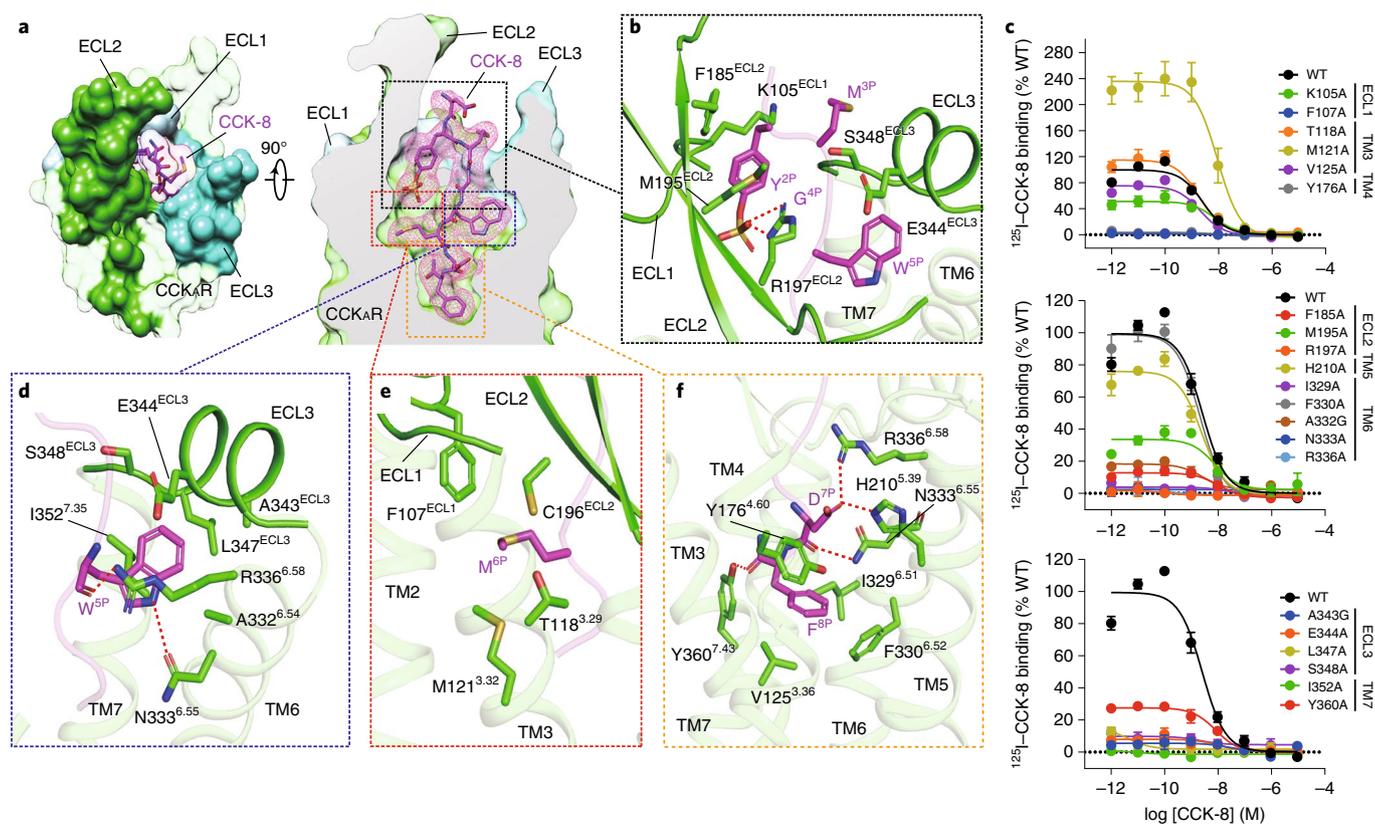
Fig. 2 | Recognition of sulfated CCK-8 by CCK$_A$R. a, CCK-8 sits in the orthosteric binding pocket of CCK$_A$R, as shown in an extracellular view (left) and side view (right). The density map of CCK-8 is shown as a magenta mesh, and CCK-8 is displayed as magenta sticks. CCK$_A$R is shown in green as a cutaway surface (right). ECL1 (light blue), ECL2 (lime green) and ECL3 (turquoise) are highlighted as solid surfaces. **b**, Detailed interactions between sulfated CCK-8 and three extracellular loops of CCK$_A$R. **c**, Effects of mutations in the receptor ligand-binding pocket on CCK-8 binding activity assessed by a radiolabeled ligand-binding assay. Data are presented as mean ± s.e.m. of three independent experiments ($n = 3$), except for the WT ($n = 4$), and conducted in triplicate. Competition curves of mutants from ECL1, TM3 and TM4 (upper), ECL2, TM5 and TM6 (middle), and ECL3 and TM7 (bottom) compared to WT CCK$_A$R are shown. **d**, Recognition of CCK-8 by the deep hydrophobic cavity beneath ECL3 of CCK$_A$R. **e**, Recognition of CCK-8 by the shallow hydrophobic cavity beneath ECL1 and ECL2 of CCK$_A$R. **f**, Recognition of CCK-8 by the bottom TMD region of CCK$_A$R. Key interaction residues from CCK$_A$R are shown as green sticks, and the receptor is shown in cartoon presentation. Polar interactions are indicated as red dashed lines.

in G$_q$/G$_s$-coupled complexes, and 1.18 for CCK-8 in G$_q$/G$_i$-coupled complexes), supported by clear EM density maps (Fig. 2a and Supplementary Fig. 5). The region of ligand recognition by CCK$_A$R can be divided into three major parts: (1) the extracellular loops, (2) hydrophobic cavities beneath ECLs and (3) the bottom of the TMD pocket (Fig. 2a).

At the extracellular side, three ECLs are folded to embrace the N-terminal amino acids of CCK-8 (Fig. 2a). The sulfate group of ionic Y^{2P} interacts with the side chain of R197^{ECL2}. This polar interaction prompts the aromatic ring of Y^{2P} to form hydrophobic contacts with F185^{ECL2}, M195^{ECL2} and the main chain of K105^{ECL1}, thus connecting CCK-8 to ECL1 and ECL2 (Fig. 2b and Extended Data Fig. 3). These structural observations are consistent with the previous finding that the R197^{ECL2}M mutation was 1,470-fold less potent than the wild-type (WT) CCK$_A$R[11]. The alanine mutation of R197^{ECL2} completely abolishes the binding of CCK-8, thus strongly supporting the contention that R197^{ECL2} serves as a determinant to discriminate between sulfated and non-sulfated CCK (Fig. 2c and Supplementary Table 2). Likewise, poor ligand selectivity of CCK$_B$R may be attributed to a substitution of arginine for valine at the corresponding position (Extended Data Fig. 4). Meanwhile, M^{3P}, G^{4P} and W^{5P} clamp the interior surface of ECL3 (Fig. 2b).

Two hydrophobic cavities exist below the ECLs to accommodate W^{5P} and M^{6P} (Fig. 2d,e). The side chain of W^{5P} is sandwiched by

the side chains of I352$^{7.35}$ and R336$^{6.58}$ and buries in a deep hydrophobic pocket composed of TM6, ECL3 and TM7 (Fig. 2d). The backbone CO group of W^{5P} forms a hydrogen bond with R336$^{6.58}$, and its indole nitrogen atom makes another hydrogen bond with N333$^{6.55}$ (Fig. 2d and Extended Data Fig. 3), which is reported to be critical to CCK$_A$R activation[38]. Alanine mutations in residues N333$^{6.55}$, R336$^{6.58}$, A343^{ECL3}, E344^{ECL3}, L347^{ECL3} and S348^{ECL3} completely abolish the binding of CCK-8, suggesting the key roles of these residues in CCK-8 recognition (Fig. 2c and Supplementary Table 2). In contrast to the W^{5P}-occupied hydrophobic pocket, M^{6P} sits in a relatively shallow hydrophobic cavity in the opposite direction, constituted by F107^{ECL1}, C196^{ECL2}, T118$^{3.29}$ and M121$^{3.32}$ (Fig. 2e and Extended Data Fig. 3). Mutating F107^{ECL1} and residues in ECL2 and ECL3 to alanine eliminated the binding ability of CCK-8 entirely, highlighting an essential function of the three ECLs in peptide recognition (Fig. 2c and Supplementary Table 2).

At the bottom of the binding pocket, D^{7P} and main chain CO group of CCK-8 form a stabilizing polar interaction network with TM5 (H210$^{5.39}$), TM6 (N333$^{6.55}$ and R336$^{6.58}$) and TM7 (Y360$^{7.43}$) (Fig. 2f and Extended Data Fig. 3). The phenyl ring of F^{8P} makes a polar hydrogen–π interaction with Y176$^{4.60}$, and inserts into a large hydrophobic crevice composed of residues from TM3, TM4, TM5 and TM6 (Fig. 2f and Extended Data Fig. 3). Besides N333$^{6.55}$ and R336$^{6.58}$, which also interact in a polar
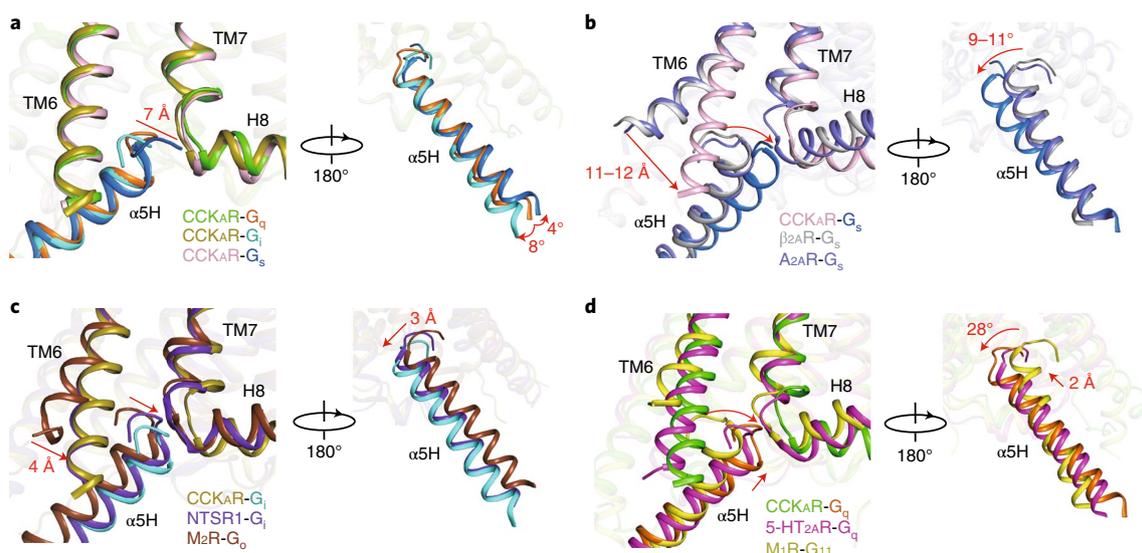
**Fig. 3 | Structural comparison of TM6 and the α5-helix between CCK$_A$R–G protein complexes and representative G$_s$-, G$_q$- and G$_i$-coupled GPCR structures, in two different views. a**, Structural comparison of CCK$_A$R–G$_q$, CCK$_A$R–G$_s$ and CCK$_A$R–G$_i$ complexes. A 7-Å movement of the distal end of the Gα$_s$ α5-helix (α5H) relative to that of Gα$_q$ and a swing of the Gα α5-helix are highlighted as red arrows. **b**, Structural comparison of CCK$_A$R–G$_s$ with β$_{2A}$R–G$_s$ and A$_{2A}$R–G$_s$ complexes. Red arrows indicate an 11–12-Å displacement of TM6 and a 9–11° swing of the Gα α5-helix of G$_s$-coupled CCK$_A$R relative to G$_s$-coupled β$_{2A}$R and A$_{2A}$R. **c**, Structural comparison of CCK$_A$R–G$_i$ with the NTSR1–G$_i$ and M$_2$R–G$_o$ complexes. A 4-Å inward displacement of TM6 and a 3-Å Gα$_i$ α5-helix shift of G$_q$-coupled CCK$_A$R in comparison to G$_o$-coupled M$_2$R are indicated as red arrows. **d**, Structural comparison of CCK$_A$R–G$_q$ with 5-HT$_{2A}$R–G$_q$ and M$_1$R–G$_{11}$ complexes. A 2-Å upward movement of Gα$_q$ of G$_q$-coupled CCK$_A$R compared to G$_q$-coupled 5-HT$_{2A}$R and a 28° rotation relative to G$_{11}$-coupled M$_1$R are highlighted as red arrows. The complex structures are aligned based on TM2–TM4 of the receptors. β$_{2A}$R–G$_s$, A$_{2A}$R–G$_s$, NTSR1–G$_i$, M$_2$R–G$_o$, 5-HT$_{2A}$R–G$_q$ and M$_1$R–G$_{11}$ structures (PDB 3SN6, 5G53, 6OS9, 6OIK, 6WHA and 6OIJ) are colored in gray, marine, purple blue, dark brown, magenta and yellow, respectively.

manner with W^{5P}, I329$^{6.51}$ is closely related to CCK-8 binding (Fig. 2c and Supplementary Table 2).

Elucidation of the recognition mechanism of CCK-8 provides clues for therapeutic development against CCK$_A$R. GW-5823, CE-326597 and Glaxo-11p (Extended Data Fig. 5a) are small-molecule agonists for CCK$_A$R with moderate activities[10,39,40]. Docking of these agonists to the CCK$_A$R shows that they only occupy the bottom half of the TMD binding pocket, thus lacking essential interactions with ECLs 1–3 of CCK$_A$R (Extended Data Fig. 5b). This structural feature may lead to a weaker activity of these small-molecule agonists relative to CCK-8. Together, our data provide a framework for understanding the mechanism of small-molecule agonist recognition and offer a template for guiding drug design targeting CCK$_A$R.

Overall coupling mode of CCK$_A$R–G protein complexes

Although all four G-protein subtypes were reported to interact with CCK$_A$R[24], only three of the CCK$_A$R–G protein samples (CCK$_A$R–G$_q$, CCK$_A$R–G$_s$ and CCK$_A$R–G$_i$ protein complexes) were obtained for high-resolution cryo-EM structure determination (Fig. 1). Structural comparison indicated that TM6 and ICL2 in CCK$_A$R adopt nearly identical conformations in G$_q$-, G$_i$- and G$_s$-coupled structures (Fig. 3a and Extended Data Fig. 6a–c). However, slightly different tilts of the Gα α5-helix were seen among the three heterotrimeric G proteins (4° for Gα$_q$/Gα$_s$ and 8° for Gα$_q$/Gα$_i$; Fig. 3a). Meanwhile, the distal end of the Gα$_s$ α5-helix moves 7 Å outward, away from the TMD core relative to the equivalent Gα$_q$ residue (measured at the Cα atom of L$^{H5.25}$, where superscripts refer to the common Gα numbering (CGN) system[41]; Fig. 3a). G$_q$ presents the largest solvent-accessible surface area (SASA, 1,492 Å^2) with the receptor, compared to a G$_s$ value of 1,293 Å^2 and G$_i$ value of 1,167 Å^2, consistent with a 6.6- to 20.3-fold increased potency of G$_q$ coupling to CCK$_A$R in comparison to coupling with G$_s$ and G$_i$ (Supplementary Table 3). This finding supports the hypothesis that the size of the

G-protein coupling interface may correlate with the ability of a receptor to link with different G proteins[21,22]. In addition, coupling of different G-protein subtypes exhibits distinct effects on CCK-8 binding. Compared to G$_s$ or G$_i$ proteins, G$_q$ coupling increases the binding affinity of CCK-8 (Supplementary Table 3), consistent with the increased binding activity of isoproterenol against β$_{2A}$R in the presence of G$_s$ protein[16]. This finding indicates an allosteric modulation effect of G$_q$ protein on CCK-8 binding, supporting the positive cooperativity between agonists and G proteins[42].

In addition, comparisons of these three complex structures to previously reported G-protein-coupled class A GPCRs reveal the different extent of TM6 displacement and the concomitant shift of the Gα α5-helix (Fig. 3b–d). TM6 of CCK$_A$R in all three G-protein complexes displays an 11–12-Å (measured at the Cα atom of the residue at position 6.27) smaller outward displacement compared to G$_s$-coupled GPCRs, which translates into a notable swing of the Gα α5-helix in the same direction (9–11° relative to G$_s$-coupled β$_{2A}$R and A$_{2A}$R as measured at the Cα atom of Y$^{H5.23}$). This smaller displacement of TM6 is contrary to the previous assumption that TM6 of G$_s$-coupled GPCRs undergoes a notable outward movement, thus opening a larger cytoplasmic pocket to accommodate bulkier residues at the distal end of the Gα α5-helix relative to G$_{i/o}$-coupled receptors[21,43]. To avoid a potential clash with TM6, the distal end of the Gα α5-helix in the CCK$_A$R–G$_s$ complex stretches away from the TMD core and inserts into the crevice between the TM6 and TM7–helix 8 joint. This featured conformation of the Gα α5-helix in the CCK$_A$R–G$_s$ complex is unique compared to that in structures of the G$_s$-coupled β$_{2A}$R and A$_{2A}$R, supporting the complexity of the GPCR–G protein coupling mechanism (Fig. 3b).

TM6 and the Gα α5-helix of CCK$_A$R–G protein complexes display similar conformational changes to other G$_i$- and G$_q$-coupled GPCRs, such as the G$_i$-coupled NTSR1 and the G$_q$-coupled 5-HT$_{2A}$R (Fig. 3c,d). TM6 of the CCKA$_R$–G$_i$ protein complex is highly overlaid
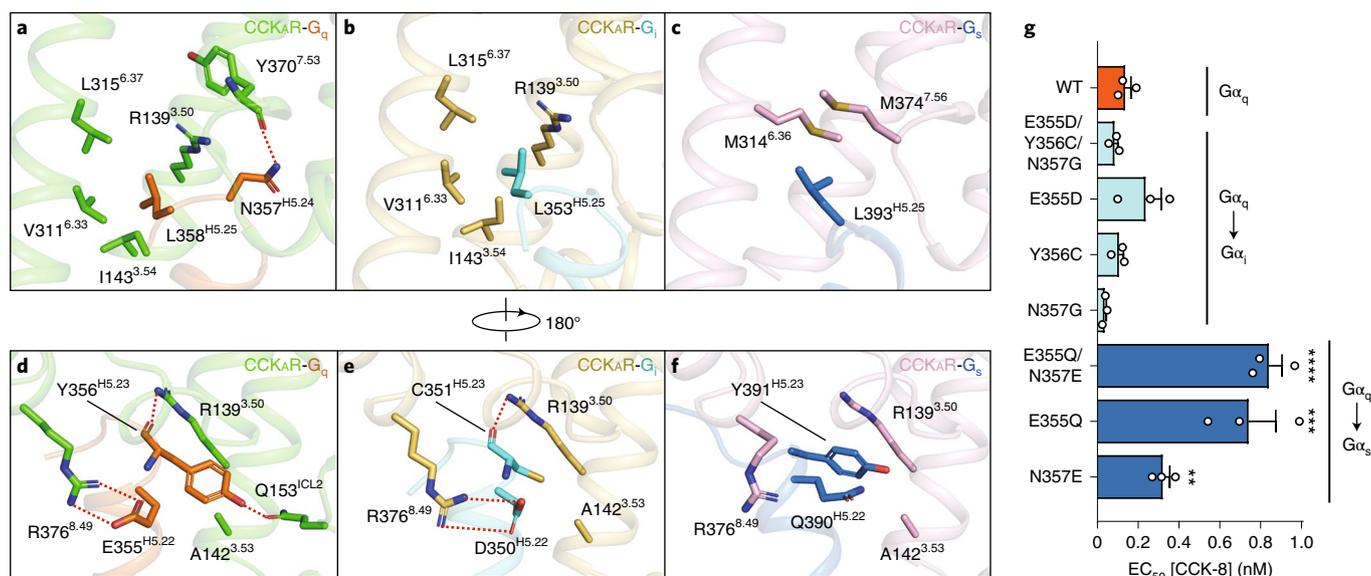
Fig. 4 | Distinct interaction patterns of residues from the 'wavy hook' motif. a–c, Details of the interaction between CCK$_A$R and L358$^{H5.25}$ and N357$^{H5.24}$ of the Gα$_q$ subunit (**a**), L353$^{H5.25}$ of the Gα$_i$ subunit (**b**) and L393$^{H5.25}$ of the Gα$_s$ subunit (**c**). **d–f**, Details of the interaction between CCK$_A$R and Y356$^{H5.23}$ and E355$^{H5.22}$ of the Gα$_q$ subunit (**d**), C351$^{H5.23}$ and D350$^{H5.22}$ of the Gα$_i$ subunit (**e**) and Y391$^{H5.23}$ and Q390$^{H5.22}$ of the Gα$_s$ subunit (**f**). Hydrogen bonds and salt bridges are indicated as red dashed lines. **g**, BRET assay evaluating the effects of 'wavy hook' substitutions on CCK$_A$R–G protein coupling. The 'wavy hook' residues of the Gα$_q$ subunit were displaced by the corresponding residues in the Gα$_s$ and Gα$_i$ subunits. Data are shown as mean ± s.e.m. of three independent experiments ($n = 3$), conducted in triplicate. All data were analyzed by one-way analysis of variance Dunnett multiple comparisons test ($P = 1$, $P = 0.9827$, $P = 0.7521$, $P = 0.9994$, $P = 0.7421$, $P < 0.0001$, $P < 0.0001$ and $P = 0.2078$ from top to bottom, ***$P < 0.001$, ****$P < 0.0001$ versus WT).

with that of G$_i$-coupled NTSR1, while the cytoplasmic end of TM6 shows a 4-Å smaller outward displacement compared to that of G$_o$-coupled M$_2$R (Fig. 3c). On the G-protein side, the α5-helix of Gα$_i$ in the CCK$_A$R–G$_i$ complex shows a nearly overlapped conformation with that of the NTSR1–G$_i$ complex. In contrast, it exhibits a 3-Å (measured at the Cα atom of Y$^{H5.23}$) shift away from TM6 relative to that of G$_o$-coupled M$_2$R (Fig. 3c). Structural comparison of G$_q$-coupled CCK$_A$R with G$_q$/G$_{11}$-coupled GPCRs demonstrates a 2-Å (measured at the Cα atom of Y$^{H5.23}$) upward shift toward the cytoplasmic cavity in comparison to the G$_q$-coupled 5-HT$_{2A}$R and a 28° rotation away from TM6 relative to G$_{11}$-coupled M$_1$R (Fig. 3d).

Interactions of the 'wavy hook' of CCK$_A$R–G protein complexes

The 'wavy hook' at the extreme C terminus of the Gα α5-helix is thought to be one of the coupling specificity determinants for G proteins[44,45], and undergoes distinct conformational rearrangements among the three CCK$_A$R–G protein complexes (Fig. 3a).

A structural comparison of the interaction interface between the receptor cytoplasmic cavity and the Gα 'wavy hook' reveals distinct features of CCK$_A$R–G protein coupling. Well-defined densities of Gα-protein 'wavy hook' residues allow for detailed structural analyses except for residues at the −1 position. The L(−2)$^{H5.25}$ in the α5-helix is highly conserved across the G-protein families and plays a pivotal role in G-protein coupling. Both L358$^{H5.25}$ in Gα$_q$ and L353$^{H5.25}$ in Gα$_i$ hydrophobically interact with residues in TM3 and TM6 (R139$^{3.50}$, I143$^{3.54}$, V311$^{6.33}$ and L315$^{6.37}$; Fig. 4a,b). Owing to the notable displacement of the Gα$_s$ C terminus, L393$^{H5.25}$ in Gα$_s$ moves 7 Å outward away from the TMD core relative to the equivalent Gα$_q$ residue (Fig. 3a), repositioning it in a hydrophobic subpocket formed by M314$^{6.36}$ and M373$^{7.56}$ (Fig. 4c). In contrast to L(−2)$^{H5.25}$, residues at positions H(−3)$^{5.24}$, H(−4)$^{5.23}$ and H(−5)$^{5.22}$ are less conserved. N357(−3)$^{H5.24}$ in Gα$_q$ makes a hydrogen bond with the backbone CO group of Y370$^{7.53}$ (Fig. 4a). As a result of the replacement of Gα$_i$ G352(−3)$^{H5.24}$ and the repositioning of Gα$_s$ E392(−3)$^{H5.24}$,

the corresponding hydrogen bond is absent in CCK$_A$R–G$_i$ and CCK$_A$R–G$_s$ complex structures. Additionally, Y356(−4)$^{H5.23}$ in Gα$_q$ forms extensive interactions with the receptor cytoplasmic cavity by making hydrogen bonds with R139$^{3.50}$ and Q153^{ICL2} (Fig. 4d). By contrast, C351(−4)$^{H5.23}$ in Gα$_i$ only forms a weak hydrogen bond with R139$^{3.50}$ via its backbone CO group (Fig. 4e). Y391(−4)$^{H5.23}$ in Gα$_s$ exhibits limited hydrophobic and van der Waals interactions with residues in TM2 and TM3 (T76$^{2.39}$, R139$^{3.50}$ and A142$^{3.53}$; Fig. 4f). Furthermore, both E355(−5)$^{H5.22}$ in Gα$_q$ and D350(−5)$^{H5.22}$ in Gα$_i$ form salt bridges with R376$^{8.49}$ in CCK$_A$R, while Q390(−5)$^{H5.22}$ in Gα$_s$ disfavors the formation of the corresponding electrostatic interaction (Fig. 4d–f). To understand the 'wavy hook'-mediated G-protein selectivity, we displaced the amino acids (H5.22–H5.25) in the Gα$_q$ subunit with the corresponding ones in the Gα$_s$ and Gα$_i$ subunits. Bioluminescence resonance energy transfer (BRET) assay results show that the Gα$_i$ displacement has no impact on CCK$_A$R–G protein coupling compared to the WT Gα$_q$ subunit. However, partial (E355Q or N357E) or complete Gα$_s$ substitution remarkably decreased the G-protein coupling activity of CCK$_A$R (Fig. 4g). These results indicate that the 'wavy hook' may play a crucial role in the coupling selectivity of CCK$_A$R with G$_q$ over the G$_s$ protein.

Contribution of CCK$_A$R ICL3 to G$_q$ coupling selectivity

In the CCK$_A$R–G$_q$ protein complex structure, CCK$_A$R displays a comparable length of TM5 relative to the M$_1$R–G$_{11}$ complex[19]. However, the cytoplasmic end of the CCK$_A$R TM5 exhibits an 8-Å outward bend (measured at the Cα atoms of A$^{5.73}$), which prevents it from interacting with the Gα$_q$ subunit (Fig. 5a). Instead, the ICL3 inserts into the cleft between TM5 of CCK$_A$R and the α5-helix of the Gα$_q$ subunit (Fig. 5a). Compared to L225$^{5.75}$ in M$_1$R, I296^{ICL3} in CCK$_A$R interacts with the same hydrophobic patch formed by the side chains of Y325$^{6.02}$, F339$^{H5.06}$ and A342$^{H5.09}$ in the Gα$_q$ subunit, but is buried deeper to create more closely packed hydrophobic contacts (Fig. 5a,b). These hydrophobic interactions are critical to CCK$_A$R–G$_q$ coupling, as evidenced by our BRET analysis
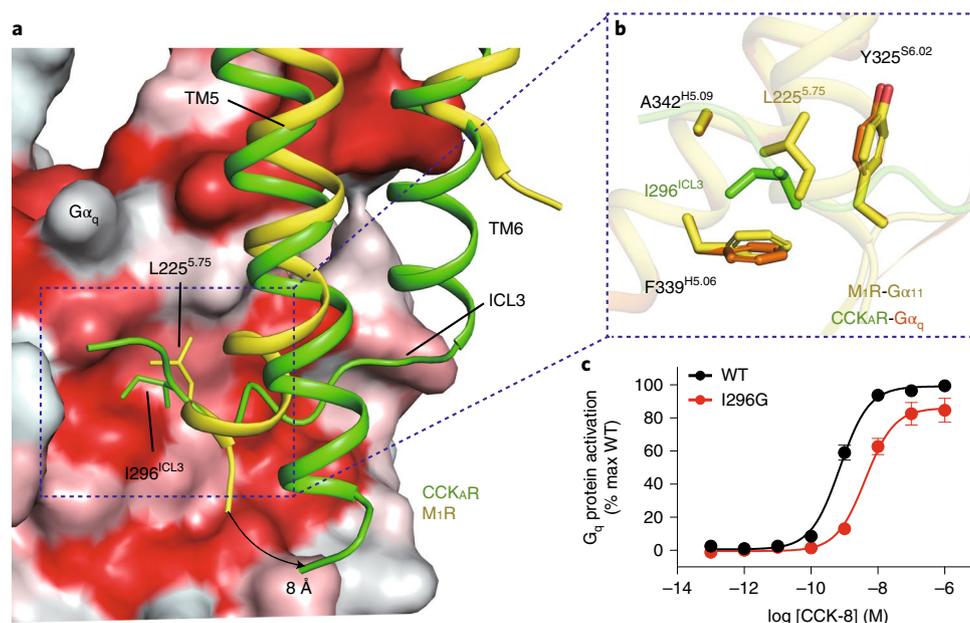
**Fig. 5 | Interaction between the ICL3 loop of CCK$_A$R and the Gα_q subunit. a**, I296^{ICL3} of CCK$_A$R and L225$^{5.75}$ of M$_1$R occupy the same hydrophobic subpocket of the Gα subunit. The Gα_q subunit is shown as a surface presentation by hydrophobicity (hydrophobic surface in red). An 8-Å outward bend of TM5 of CCK$_A$R relative to that of M$_1$R is highlighted by a black arrow. **b**, Detail of the interactions between I296^{ICL3}(CCK$_A$R)/L225$^{5.75}$(M$_1$R) and the hydrophobic patch comprising Y325$^{S6.02}$, F339$^{H5.06}$ and A342$^{H5.09}$ of the Gα_q and Gα_{11} subunits. **c**, BRET assay showing that the I296G mutation decreases the association rate of CCK$_A$R with the G$_q$ heterotrimer. Data are shown as mean ± s.e.m. of three independent experiments ($n = 3$), conducted in triplicate.

showing that the I296^{ICL3}G mutation significantly weakens G$_q$ coupling to CCK$_A$R but has no impact on G$_s$ and G$_i$ coupling (Fig. 5c and Supplementary Table 4). This hydrophobic patch, which lies on the outer surface, may be unique for the G$_{q/11}$ subunit. The equivalent residues in the Gα_s and Gα_i subunits are polar or charged residues, for which it would be energetically unfavorable to form hydrophobic interactions (Extended Data Fig. 7a–d). Indeed, this unconventional ICL3–G$_q$ interaction is not seen in the structures of G$_s$- and G$_{i/o}$-coupled CCK$_A$Rs (Fig. 3b,c). Together, our findings offer structural evidence on the possible role of ICL3 in CCK$_A$R–G$_q$ coupling preference. Hydrophobic residues on the inner surface of the ICL3 loop of CCK$_A$R or the extended TM5 of M$_1$R may represent a common feature of G$_{q/11}$-coupled GPCRs.

Discussion

As the largest family of cell surface receptors, GPCRs have more than 800 members but only couple to four G-protein subtypes. Specific GPCR signaling requires the receptor to couple with either a single or multiple G-protein subtypes[45–47]. Thus, one of the main questions is how does a given GPCR select a G-protein subtype for downstream signal transduction. The critical G-protein determinants of selectivity vary widely for different receptors that couple to specific G proteins. It is thought that G$_s$- or G$_q$-coupled receptors are relatively promiscuous and to some extent couple to G$_{i1}$ (ref. [22]). However, G$_i$-coupled receptors are more selective[22]. The minor outward movement of TM6 contributes to such a superior G$_i$ coupling selection in comparison to that of G$_s$ (refs. [17,23,44,48,49]). Although proven to be promiscuous, G$_q$-coupled receptors tend to adopt an active conformation similar to that of G$_i$-coupled GPCRs, reflecting the complexity of the GPCR–G protein coupling mechanism[19,20]. Because CCK$_A$R has the ability to couple with different G-protein subtypes, it stands out as a suitable model for studying the promiscuity of G-protein coupling. In this article, we show that TM6 of CCK$_A$R undergoes a similar outward displacement relative to G$_{i/o}$-coupled (NTSR1 and M$_2$R) and G$_{q/11}$-coupled GPCRs

(5-HT$_{2A}$R and M$_1$R), but has a smaller shift relative to G$_s$-coupled GPCRs (β_{2A}R and A$_{2A}$R). CCK$_A$Rs share almost identical conformations, whereas G$_q$, G$_s$ and G$_i$ proteins vary in distinct orientations, producing different sizes of receptor–G protein interface. The predominant coupling to G$_q$ by CCK$_A$R can be explained by its largest interface of the three CCK$_A$R–G protein complexes. Structural comparison of the three CCK$_A$R–G protein complexes reveals that 'wavy hook' residues of the Gα α5-helix and ICL3 of the receptor are important for the coupling promiscuity. In addition, detailed inspections disclose structural clues relative to the recognition mechanism of sulfated CCK-8 by CCK$_A$R, in which R197^{ECL2} is a major determinant. Together, our structures provide a framework for better understanding ligand recognition as well as G-protein coupling selectivity and promiscuity by CCK$_A$R.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41589-021-00841-3.

References

1. Bowers, M. E., Choi, D. C. & Ressler, K. J. Neuropeptide regulation of fear and anxiety: implications of cholecystokinin, endogenous opioids and neuropeptide Y. *Physiol. Behav.* **107**, 699–710 (2012).
2. Rehfeld, J. F., Friis-Hansen, L., Goetze, J. P. & Hansen, T. V. The biology of cholecystokinin and gastrin peptides. *Curr. Top. Med. Chem.* **7**, 1154–1165 (2007).
3. Wank, S. A. Cholecystokinin receptors. *Am. J. Physiol.* **269**, G628–G646 (1995).
4. Dufresne, M., Seva, C. & Fourmy, D. Cholecystokinin and gastrin receptors. *Physiol. Rev.* **86**, 805–847 (2006).

5. Moran, T. H., Smith, G. P., Hostetler, A. M. & McHugh, P. R. Transport of cholecystokinin (CCK) binding sites in subdiaphragmatic vagal branches. *Brain Res.* **415**, 149–152 (1987).

6. Weatherford, S. C., Chiruzzo, F. Y. & Laughton, W. B. Satiety induced by endogenous and exogenous cholecystokinin is mediated by CCK-A receptors in mice. *Am. J. Physiol.* **262**, R574–R578 (1992).

7. Miller, L. J., Holicky, E. L., Ulrich, C. D. & Wieben, E. D. Abnormal processing of the human cholecystokinin receptor gene in association with gallstones and obesity. *Gastroenterology* **109**, 1375–1380 (1995).

8. Hewson, G., Leighton, G. E., Hill, R. G. & Hughes, J. The cholecystokinin receptor antagonist L364,718 increases food intake in the rat by attenuation of the action of endogenous cholecystokinin. *Br. J. Pharmacol.* **93**, 79–84 (1988).

9. Jordan, J. et al. Stimulation of cholecystokinin-A receptors with GI181771X does not cause weight loss in overweight or obese patients. *Clin. Pharmacol. Ther.* **83**, 281–287 (2008).

10. Elliott, R. L. et al. Discovery of *N*-benzyl-2-[(4*S*)-4-(1*H*-indol-3-ylmethyl)-5-oxo-1-phenyl-4,5-dihydro-6*H*-[1,2,4]triazolo[4,3-a][1,5]benzodiazepin-6-yl]-*N*-isopropylacetamide, an orally active, gut-selective CCK1 receptor agonist for the potential treatment of obesity. *Bioorg. Med. Chem. Lett.* **20**, 6797–6801 (2010).

11. Gigoux, V. et al. Arginine 197 of the cholecystokinin-A receptor binding site interacts with the sulfate of the peptide agonist cholecystokinin. *Protein Sci.* **8**, 2347–2354 (1999).

12. Escrieut, C. et al. The biologically crucial C terminus of cholecystokinin and the non-peptide agonist SR-146,131 share a common binding site in the human CCK1 receptor. Evidence for a crucial role of Met-121 in the activation process. *J. Biol. Chem.* **277**, 7546–7555 (2002).

13. Archer-Lahlou, E. et al. Molecular mechanism underlying partial and full agonism mediated by the human cholecystokinin-1 receptor. *J. Biol. Chem.* **280**, 10664–10674 (2005).

14. Archer-Lahlou, E. et al. Modeled structure of a G-protein-coupled receptor: the cholecystokinin-1 receptor. *J. Med. Chem.* **48**, 180–191 (2005).

15. Miller, L. J. & Gao, F. Structural basis of cholecystokinin receptor binding and regulation. *Pharmacol. Ther.* **119**, 83–95 (2008).

16. Rasmussen, S. G. et al. Crystal structure of the β₂ adrenergic receptor–Gs protein complex. *Nature* **477**, 549–555 (2011).

17. Kang, Y. et al. Cryo-EM structure of human rhodopsin bound to an inhibitory G protein. *Nature* **558**, 553–558 (2018).

18. Kato, H. E. et al. Conformational transitions of a neurotensin receptor 1–Gᵢ₁ complex. *Nature* **572**, 80–85 (2019).

19. Maeda, S., Qu, Q., Robertson, M. J., Skiniotis, G. & Kobilka, B. K. Structures of the M1 and M2 muscarinic acetylcholine receptor/G-protein complexes. *Science* **364**, 552–557 (2019).

20. Kim, K. et al. Structure of a hallucinogen-activated Gq-coupled 5-HT₂ₐ serotonin receptor. *Cell* **182**, 1574–1588 (2020).

21. Qiao, A. et al. Structural basis of Gₛ and Gᵢ recognition by the human glucagon receptor. *Science* **367**, 1346–1352 (2020).

22. Okashah, N. et al. Variable G protein determinants of GPCR coupling selectivity. *Proc. Natl Acad. Sci. USA* **116**, 12054–12059 (2019).

23. Rose, A. S. et al. Position of transmembrane helix 6 determines receptor G protein coupling specificity. *J. Am. Chem. Soc.* **136**, 11244–11247 (2014).

24. Hauser, A. S. et al. Pharmacogenomics of GPCR drug targets. *Cell* **172**, 41–54 (2018).

25. Paulssen, R. H., Fraeyman, N. & Florholmen, J. Activation of phospholipase C by cholecystokinin receptor subtypes with different G-protein-coupling specificities in hormone-secreting pancreatic cell lines. *Biochem. Pharmacol.* **60**, 865–875 (2000).

26. Wu, V. et al. First intracellular loop of the human cholecystokinin-A receptor is essential for cyclic AMP signaling in transfected HEK-293 cells. *J. Biol. Chem.* **272**, 9037–9042 (1997).

27. Scemama, J. L. et al. CCK and gastrin inhibit adenylate cyclase activity through a pertussis toxin-sensitive mechanism in the tumoral rat pancreatic acinar cell line AR 4-2J. *FEBS Lett.* **242**, 61–64 (1988).

28. Sabbatini, M. E., Bi, Y., Ji, B., Ernst, S. A. & Williams, J. A. CCK activates RhoA and Rac1 differentially through Gα₁₃ and Gαq in mouse pancreatic acini. *Am. J. Physiol. Cell Physiol.* **298**, C592–C601 (2010).

29. Le Page, S. L., Bi, Y. & Williams, J. A. CCK-A receptor activates RhoA through Gα₁₂/₁₃ in NIH3T3 cells. *Am. J. Physiol. Cell Physiol.* **285**, C1197–C1206 (2003).

30. Maeda, S. et al. Development of an antibody fragment that stabilizes GPCR/G-protein complexes. *Nat. Commun.* **9**, 3712 (2018).

31. Carpenter, B., Nehme, R., Warne, T., Leslie, A. G. & Tate, C. G. Structure of the adenosine A₂ₐ receptor bound to an engineered G protein. *Nature* **536**, 104–107 (2016).

32. Liu, P. et al. The structural basis of the dominant negative phenotype of the Gαᵢ₁β₁γ₂ G203A/A326S heterotrimer. *Acta Pharmacol. Sin.* **37**, 1259–1272 (2016).

33. 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).

34. Shiimura, Y. et al. Structure of an antagonist-bound ghrelin receptor reveals possible ghrelin recognition mode. *Nat. Commun.* **11**, 4160 (2020).

35. Wingler, L. M. et al. Angiotensin and biased analogs induce structurally distinct active conformations within a GPCR. *Science* **367**, 888–892 (2020).

36. Hong, C. et al. Structures of active-state orexin receptor 2 rationalize peptide and small-molecule agonist recognition and receptor activation. *Nat. Commun.* **12**, 815 (2021).

37. Zhou, F. et al. Molecular basis of ligand recognition and activation of human V2 vasopressin receptor. *Cell Res.* https://doi.org/10.1038/s41422-021-00480-2 (2021).

38. Gigoux, V. et al. Arginine 336 and asparagine 333 of the human cholecystokinin-A receptor binding site interact with the penultimate aspartic acid and the C-terminal amide of cholecystokinin. *J. Biol. Chem.* **274**, 20457–20464 (1999).

39. Henke, B. R. et al. Optimization of 3-(1*H*-indazol-3-ylmethyl)-1,5-benzodiazepines as potent, orally active CCK-A agonists. *J. Med. Chem.* **40**, 2706–2725 (1997).

40. Aquino, C. J. et al. Discovery of 1,5-benzodiazepines with peripheral cholecystokinin (CCK-A) receptor agonist activity. 1. Optimization of the agonist 'trigger'. *J. Med. Chem.* **39**, 562–569 (1996).

41. Flock, T. et al. Universal allosteric mechanism for Gα activation by GPCRs. *Nature* **524**, 173–179 (2015).

42. DeVree, B. T. et al. Allosteric coupling from G protein to the agonist-binding pocket in GPCRs. *Nature* **535**, 182–186 (2016).

43. Garcia-Nafria, J. & Tate, C. G. Cryo-EM structures of GPCRs coupled to Gₛ, Gᵢ and G₀. *Mol. Cell. Endocrinol.* **488**, 1–13 (2019).

44. Garcia-Nafria, J., Nehme, R., Edwards, P. C. & Tate, C. G. Cryo-EM structure of the serotonin 5-HT1B receptor coupled to heterotrimeric G₀. *Nature* **558**, 620–623 (2018).

45. Kim, H. R. et al. Structural mechanism underlying primary and secondary coupling between GPCRs and the Gi/o family. *Nat. Commun.* **11**, 3160 (2020).

46. Albert, P. R. & Robillard, L. G protein specificity: traffic direction required. *Cell. Signal.* **14**, 407–418 (2002).

47. Woehler, A. & Ponimaskin, E. G. G protein-mediated signaling: same receptor, multiple effectors. *Curr. Mol. Pharm.* **2**, 237–248 (2009).

48. Koehl, A. et al. Structure of the μ-opioid receptor–Gᵢ protein complex. *Nature* **558**, 547–552 (2018).

49. Draper-Joyce, C. J. et al. Structure of the adenosine-bound human adenosine A1 receptor–Gᵢ complex. *Nature* **558**, 559–563 (2018).

Methods

Expression and purification of CCK$_A$R–G protein complexes. The WT CCK$_A$R (residues 1–428) was used for cryo-EM studies. Full-length CCK$_A$R complementary DNA was cloned into a modified pFastBac vector (Invitrogen) containing a hemagglutinin (HA) signal sequence followed by an 8× histidine tag, a double-maltose binding protein tag and a tobacco etch virus (TEV) protease site before the receptor sequence using homologous recombination (using a CloneExpress One Step Cloning Kit, Vazyme; Supplementary Fig. 4a). The N-terminal 1–29 amino acids of Gα$_q$ were replaced by the equivalent residues of Gα$_i$ to facilitate scFv16 binding[19]. An engineered Gα$_s$ construct was generated based on mini-Gα$_s$[31]. The N-terminal 1–18 amino acids and the α-helical domain of Gα$_s$ were replaced by human Gα$_{i1}$, thus providing binding sites for scFv16 and Fab-G50, respectively[17,19]. Additionally, human Gα$_{i1}$ with two dominant-negative mutations (G203A and A326S[32]) was used to assemble a stable GPCR–G$_i$ protein complex. These two cognate mutations also exist in engineered Gα$_q$ and Gα$_s$ (Supplementary Fig. 4b). Receptor, rat H6-Gβ, bovine Gγ and the specific Gα subunit were co-expressed in *Spodoptera frugiperda* (*sf*9) insect cells (Invitrogen) as previously described[50]. In addition, GST-Ric-8A (a gift from B. Kobilka) was applied to improve the expression of Gα$_q$.

ScFv16 was applied to improve the protein stability of CCK$_A$R–G$_q$ and CCK$_A$R–G$_i$ complex samples. The monomeric scFv16 was prepared as previously reported[51]. Cell pellets of the co-expression culture were thawed and lysed in 20 mM HEPES, pH 7.4, 100 mM NaCl, 10% glycerol, 5 mM MgCl$_2$ and 10 mM CaCl$_2$ supplemented with EDTA-free protease inhibitor cocktail (TargetMol). CCK$_A$R–G protein complexes were assembled at room temperature (r.t.) for 1 h by the addition of 10 μM CCK-8 (GenScript) and 25 mU ml^{-1} apyrase. The lysate was then solubilized in 0.5% lauryl maltose neopentyl glycol (LMNG), 0.1% cholesteryl hemisuccinate TRIS salt (CHS), and the soluble fraction was purified by nickel affinity chromatography (Ni Smart Beads 6FF, SMART Lifesciences). In the case of CCK$_A$R–G$_i$ and CCK$_A$R–G$_q$ complexes, a three-molar excess of scFv16 was added to the protein elute. The mixture was incubated with amylose resin for 2 h at 4 °C. The excess G protein and scFv16 were washed with 20 column volumes of 20 mM HEPES, pH 7.4, 100 mM NaCl, 10% glycerol, 0.01% LMNG, 0.002% CHS and 2 μM CCK-8. TEV protease was then included to remove the N-terminal fusion tags of CCK$_A$R. After 1 h of incubation at r.t., the flow-through was collected, concentrated and injected onto a Superdex 200 10/300 column equilibrated in buffer containing 20 mM HEPES, pH 7.4, 100 mM NaCl, 0.00075% LMNG, 0.00025% glycol-diosgenin (GDN), 0.0002% CHS and 10 μM CCK-8. The monomeric complex peak was collected and concentrated to ~5 mg ml^{-1} for cryo-EM studies.

Cryo-electron microscopy grid preparation and image collection. For preparation of cryo-EM grids, 2.5 μl of each purified CCK$_A$R–G protein complex was applied individually onto glow-discharged holey carbon grids (Quantifoil, Au300 R1.2/1.3) in a Vitrobot chamber (FEI Vitrobot Mark IV). The chamber was set to 100% humidity at 4 °C. Extra samples were blotted for 2 s and vitrified by plunging into liquid ethane. Grids were stored in liquid nitrogen for condition screening and data collection usage.

Automatic data collection of CCK-8–CCK$_A$R–G protein complexes was performed on an FEI Titan Krios system at 300 kV. The microscope was operated with a nominal magnification of ×81,000 in counting mode, corresponding to a pixel size of 1.045 Å for the micrographs. A total of 5,415 videos for the dataset of the CCK-8–CCK$_A$R–G$_q$–scFv16 complex, 5,008 for the dataset of the CCK-8–CCK$_A$R–G$_s$ complex and 4,811 for the dataset of the CCK-8–CCK$_A$R–G$_i$–scFv16 complex were collected, respectively, by a Gatan K3 Summit direct electron detector with a Gatan energy filter (operated with a slit width of 20 eV; GIF) using SerialEM software. The images were recorded at a dose rate of ~26.7 e Å^{-2} s^{-1} with a defocus ranging from −0.5 to −3.0 μm. The total exposure time was 3 s and intermediate frames were recorded in intervals of 0.083 s, resulting in a total of 36 frames per micrograph.

Image processing and map reconstruction. Image stacks were subjected to beam-induced motion correction and aligned using MotionCor 2.1. Contrast transfer function (CTF) parameters were estimated by Ctffind4. Data processing was performed using RELION-3.0[52]. Micrographs with measured resolution worse than 4.0 Å and micrographs imaged within the carbon area were discarded, generating 3,806 micrographs for the CCK-8–CCK$_A$R–G$_q$–scFv16 dataset, 4,963 for the CCK-8–CCK$_A$R–G$_s$ dataset and 4,543 for the CCK-8–CCK$_A$R–G$_i$–scFv16 dataset for further data processing. For particle selection, two-dimensional (2D) and 3D classifications were performed on a binned dataset with a pixel size of 2.09 Å. About 2,000 particles were manually selected and subjected to 2D classification. Representative averages were chosen as a template for particle autopicking. The autopicking process produced 3,405,355 particles for the CCK-8–CCK$_A$R–G$_q$–scFv16 complex, 4,680,972 for the CCK-8–CCK$_A$R–G$_s$ complex and 4,270,010 for the CCK-8–CCK$_A$R–G$_i$–scFv16 complex, which were subjected to reference-free 2D classifications to discard bad particles. Initial reference map models for 3D classification were generated by Relion using representative 2D averages. For the CCK-8–CCK$_A$R–G$_q$–scFv16 complex, the particles selected from 2D classification were subjected to six rounds of 3D classification, resulting in a single well-defined subset with 555,628 particles. For the CCK-8–CCK$_A$R–G$_s$

complex, particles resulting from 2D classification were subjected to five rounds of 3D classification, resulting in two well-defined subsets with 499,924 particles. For the CCK-8–CCK$_A$R–G$_i$–scFv16 complex, particles selected from 2D classification were subjected to seven rounds of 3D classifications, resulting in two well-defined subsets with 140,602 particles. Further 3D refinement, CTF refinement, Bayesian polishing and DeepEnhancer processing generated density maps with an indicated global resolution of 2.9 Å for the CCK-8–CCK$_A$R–G$_q$–scFv16 complex, 3.1 Å for the CCK-8–CCK$_A$R–G$_s$ complex and 3.2 Å for the CCK-8–CCK$_A$R–G$_i$–scFv16 complex, respectively, at a Fourier shell correlation of 0.143.

Model building and refinement. For the CCK$_A$R–G$_q$ complex, the initial G$_q$ protein and scFv16 model were adopted from the cryo-EM structure of the M$_1$R–G$_{11}$ protein complex (PDB 6OIJ)[19]. The initial CCK$_A$R model was generated by an online homology model building tool[53]. All models were docked into the EM density map using Chimera[54], followed by iterative manual adjustment and rebuilding in COOT[55] and ISOLDE[56], and real-space refinement using Phenix programs[57]. The model statistics were validated using Phenix comprehensive validation. A model of the refined CCK$_A$R from the CCK$_A$R–G$_q$ complex was used for the other two complexes. Models from PTH1R–G$_s$ (PDB 6NBF) and FPR2–G$_i$ (PDB 6OMM) were used as templates for the model building of G$_s$ in the CCK$_A$R–G$_s$ complex and G$_{i1}$–scFv16 in the CCK$_A$R–G$_i$ complex, respectively. The fitted models were then built in the same way as the CCK$_A$R–G$_q$ complex. The final refinement statistics are provided in Supplementary Table 1. All figures were prepared using PyMol and Chimera software.

Radiolabeled ligand-binding assay. The WT or mutant CCK$_A$Rs were transiently transfected into HEK 293T/17 cells (purchased from the Cell Bank at the Chinese Academy of Sciences), which were cultured in a poly-D-lysine-coated 96-well plate. After 24 h, the cells were washed twice and incubated with blocking buffer (Dulbecco's modified Eagle medium (DMEM) supplemented with 33 mM HEPES, and 0.1% (wt/vol) bovine serum albumin (BSA), pH 7.4) for 2 h at 37 °C. After three washes with ice-cold phosphate-buffered saline (PBS), the cells were treated by a constant concentration of ^{125}I-CCK-8 (40 pM, PerkinElmer) plus eight different doses of CCK-8 (1 pM to 10 μM) for 3 h at r.t. Cells were washed three times with ice-cold PBS and lysed by 50 μl of lysis buffer (PBS supplemented with 20 mM Tris-HCl and 1% (vol/vol) Triton X-100, pH 7.4). Subsequently, the plates were counted for radioactivity (counts per minute) in a scintillation counter (MicroBeta^2 plate counter, PerkinElmer) using 150 μl of scintillation cocktail (OptiPhase SuperMix, PerkinElmer).

G-protein dissociation assay. G-protein dissociation was monitored by BRET experiments performed as previously reported[58]. Briefly, a C-terminal fragment of the G-protein-coupled receptor kinase 3 (GRK3ct) fused to a luciferase serves as a BRET donor. Gβγ dimer is labeled with the fluorescent protein Venus, a BRET acceptor. Upon G-protein heterotrimer activation, free Gβγ-Venus is released and binds to membrane-associated GRK3ct-luciferase, leading to an increased signal detectable by BRET.

HEK 293T/17 cells were seeded onto a 10 μg ml^{-1} Matrigel-coated six-well plate (1 × 10^6 cells per well). After 4 h of culture, WT or mutant CCK$_A$R (0.84 μg), Gα (Gα$_q$, Gα$_s$ and Gα$_i$, 2.1 μg each), Gβ (0.42 μg), Gγ (0.42 μg) and GRK (0.42 μg) were transiently transfected with Lipofectamine LTX reagent (Invitrogen). At 24 h post-transfection, cells were washed once with DMEM medium (no phenol red) and detached by EDTA. Cells were then collected by centrifugation at 1,000 r.p.m. for 5 min and resuspended in DMEM medium. Approximately 75,000 cells per well were distributed in 96-well flat-bottomed white microplates (PerkinElmer). The NanoBRET substrate (furimazine, 25 μl per well, Promega) was added, and the BRET signal (535 nm/475 nm ratio) was determined using an EnVision multilabel plate reader (PerkinElmer). The average baseline value recorded before CCK-8 stimulation was subtracted from BRET signal values.

NanoBiT G-protein recruitment assay. The recruitment of CCK$_A$R to G$_i$ protein was detected in *sf*9 cells using the NanoBiT method as previously reported[59]. Briefly, the LgBiT fragment of NanoBiT luciferase was fused to the C terminus of CCK$_A$R. SmBiT was fused to the C terminus of the Gβ subunit with a 15-amino-acid flexible linker. CCK$_A$R-LgBiT, Gα$_{i1}$, SmBiT-fused human Gβ1 and human Gγ2 were co-expressed in *sf*9 insect cells. Cell pellets were collected by centrifugation after infection for 48 h. The cell suspension was dispensed in a 96-well plate (64,000 cells per well) at a volume of 80 μl diluted in the assay buffer (Hanks' balanced salt solution buffer supplemented with 10 mM HEPES, pH 7.4) and incubated for 30 min at 37 °C. The cells were then reacted with 10 μl of 50 mM coelenterazine H (Yeasen) for 2 h at r.t. The luminescence signal was measured using an EnVision plate reader (PerkinElmer) at 30-s intervals (25 °C). The baseline was measured before CCK-8 addition for eight intervals, and the measurements continued for 20 intervals following ligand addition. Data were corrected to baseline measurements and the results were analyzed using GraphPad Prism 8.0 (Graphpad Software).

NanoBiT G-protein dissociation assay. G$_s$ activation was measured by a NanoBiT dissociation assay. G-protein NanoBiT split luciferase constructs were generated

by fusing the LgBiT in Gα_s and the SmBiT to Gγ (a gift from A. Inoue, Tohoku University) as previously reported[60]. In brief, HEK 293T/17 cells were plated in 10-cm plates at a density of 3×10^6 cells per plate. After 24 h, cells were transfected with 1.62 µg receptor plasmids, 0.81 µg Gα_s-LgBiT, 4.1 µg Gβ and 4.1 µg SmBiT-Gγ using Lipofectamine LTX reagent (Invitrogen). The transiently transfected cells were then seeded into poly-D-lysine-coated 96-well plates (50,000 cells per well) and grown overnight before incubation in assay buffer. Measurement of the luminescence signal was identical to the steps described above.

Surface expression assay. HEK 293T/17 cells were seeded into a six-well plate and incubated overnight. After transient transfection with WT or mutant plasmids for 24 h, the cells were collected and blocked with 5% BSA in PBS at r.t. for 15 min and incubated with primary anti-Flag antibody (1:300, Sigma-Aldrich) at r.t. for 1 h. The cells were then washed three times with PBS containing 1% BSA followed by 1 h of incubation with donkey anti-mouse Alexa Fluor 488-conjugated secondary antibody (1:1,000, Thermo Fisher) at 4 °C in the dark. After three washes, the cells were resuspended in 200 µl of PBS containing 1% BSA for detection in a NovoCyte flow cytometer (ACEA Biosciences) utilizing laser excitation and emission wavelengths of 488 nm and 519 nm, respectively. For each assay point, ~15,000 cellular events were collected, and the total fluorescence intensity of the positive expression cell population was calculated. The gating strategy and the method for the calculation of expression are shown in Supplementary Fig. 6.

Molecular docking. Before docking, hydrogens were added to CCK$_A$R and the whole system coordinates were optimized with a pH of 7.0. A grid file was then generated on the peptide pocket in our G$_q$-coupled CCK$_A$R structure. Small-molecule ligands Glaxo-11p, GW-5823 and CE-326597 were prepared in the OPLS3 force field with a pH of 7.0 to generate 3D structures. Finally, glide docking with standard precision was applied to all ligands and the structures with the best docking score were picked as outputs.

Statistics. All functional study data were analyzed using Prism 8 (GraphPad) and are shown as the mean ± s.e.m. from at least three independent experiments. Concentration–response curves were evaluated with a three-parameter logistic equation. Significance was determined with either a two-tailed Student's t-test or one-way analysis of variance Dunnett multiple comparisons test, and $P < 0.05$ was considered statistically significant.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this Article.

Data availability
Density maps and structure coordinates have been deposited in the Electron Microscopy Data Bank (EMDB) and the Protein Data Bank (PDB) with accession codes EMD-31387 and 7EZH for the CCK-8–CCK$_A$R–G$_i$–scFv16 complex, EMD-31388 and 7EZK for the CCK-8–CCK$_A$R–G$_s$ complex, and EMD-31389 and 7EZM for the CCK-8–CCK$_A$R–G$_q$–scFv16 complex. Source data are provided with this paper.

References
50. Duan, J. et al. Cryo-EM structure of an activated VIP1 receptor–G protein complex revealed by a NanoBiT tethering strategy. *Nat. Commun.* **11**, 4121 (2020).
51. Zhuang, Y. et al. Structure of formylpeptide receptor 2–G$_i$ complex reveals insights into ligand recognition and signaling. *Nat. Commun.* **11**, 885 (2020).
52. Scheres, S. H. RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* **180**, 519–530 (2012).
53. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **10**, 845–858 (2015).
54. Pettersen, E. F. et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
55. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).
56. Croll, T. I. ISOLDE: a physically realistic environment for model building into low-resolution electron-density maps. *Acta Crystallogr. D Struct. Biol.* **74**, 519–530 (2018).
57. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010).
58. Masuho, I., Martemyanov, K. A. & Lambert, N. A. Monitoring G protein activation in cells with BRET. *Methods Mol. Biol.* **1335**, 107–113 (2015).
59. Xu, P. et al. Structures of the human dopamine D3 receptor–G$_i$ complexes. *Mol. Cell* **81**, 1147–1159 (2021).
60. Inoue, A. et al. Illuminating G-protein-coupling selectivity of GPCRs. *Cell* **177**, 1933–1947 (2019).

Author contributions
Q.L. screened the expression constructs, optimized the CCK$_A$R–G protein complexes, prepared the protein samples for final structure determination, participated in cryo-EM grid inspection and data collection, built and refined the structure models, prepared the constructs for functional assays, analyzed the structures and prepared the figures and wrote the initial manuscript. Y.Z. performed cryo-EM grid preparation, data collection and structure determination and participated in protein sample optimization and figure and manuscript preparation. T.I.C. helped build and refine the structure model. X.H. performed the molecular docking. J.D. and W.Y. designed G-protein constructs. F.Z. participated in data analysis. B.W. and Q.Z. supervised research. H.E.X. conceived and supervised the project, analyzed the structures and initiated collaborations with M.-W.W., and supervised Q.L., Y.Z., F.Z., J.D. and W.Y. M.-W.W. and D.Y. supervised X.C., A.D. and C.Y. in G-protein assay development and data analysis. M.-W.W. participated in manuscript writing. Y.J. supervised the studies, performed the structural analysis, prepared the figures and wrote the manuscript with input from all co-authors.

Extended Data Fig. 1 | Active conformation of CCK$_A$R. a-b, Structural comparison of inactive ghrelin receptor (grey), active NTSR1 (purple blue), and active CCK$_A$R (green). Side view (**a**) and intracellular view (**b**) of the overall comparison are shown. **c**, Structural rearrangements of key activation motifs (PIF, ERY, CWxP, and NPxxY) in CCK$_A$R compared to inactive ghrelin receptor and active NTSR1. NTSR1, neurotensin receptor 1.
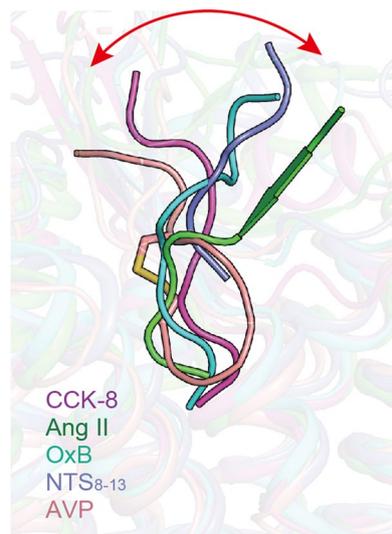
Extended Data Fig. 2 | Structure comparison of CCK-8 with other neuropeptides solved to date. The neuropeptides are shown as a cartoon. The shift of the extracellular part of neuropeptides is highlighted as a red arrow. CCK-8 in the CCK-8–CCK$_A$R–G$_q$ complex structure, magenta; Ang II, angiotensin II (PDB: 6OS0), green; OxB, orexin B (PDB: 7L1U), cyan; NTS$_{8-13}$, neurotensin 8-13 (PDB: 6OS9), purple blue; AVP, arginine vasopressin (PDB: 7DW9), salmon.
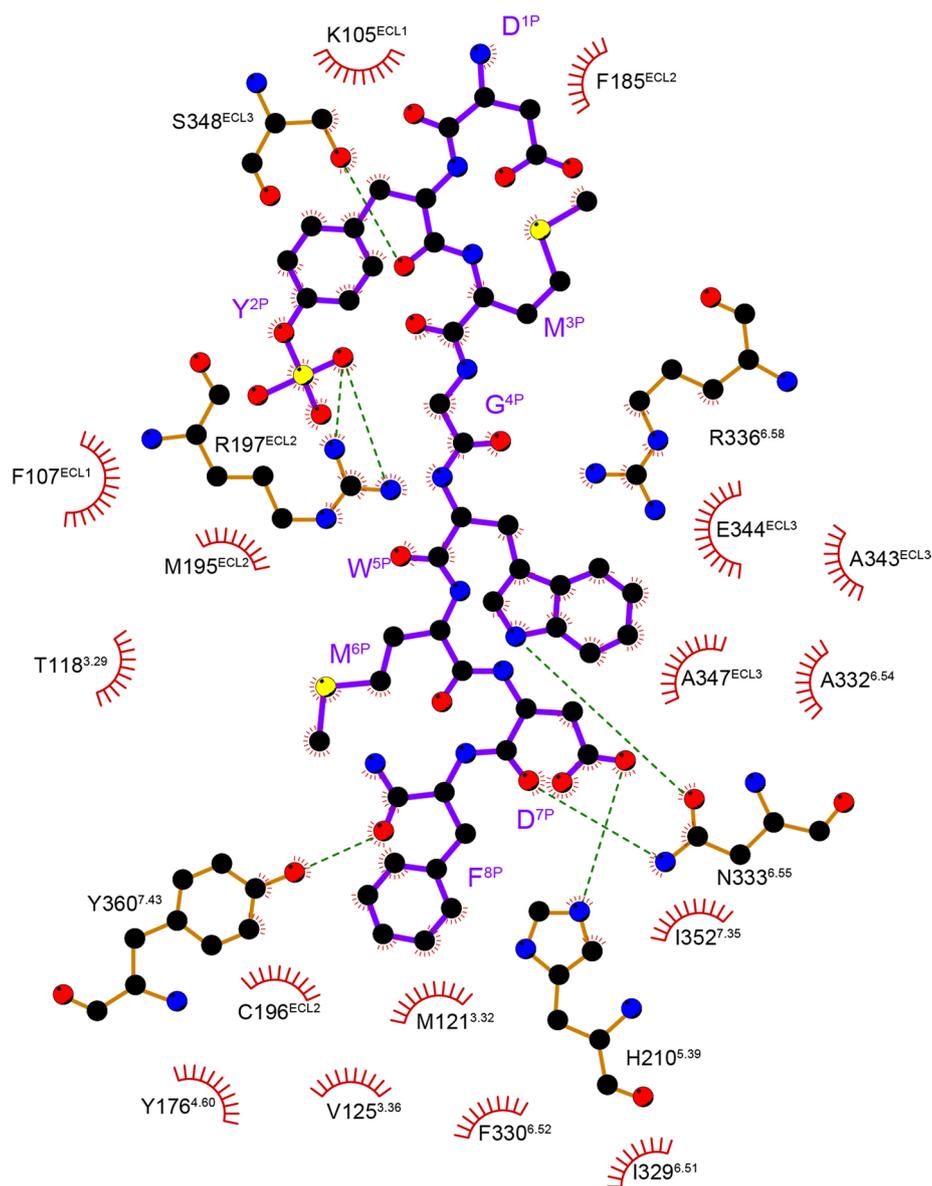
Extended Data Fig. 3 | 2D interaction plot of CCK$_A$R recognition by sulfated CCK-8. Residues in the ligand-binding pocket are colored in green. CCK-8 is displayed as magenta sticks. Polar interactions are indicated as red dashed lines.
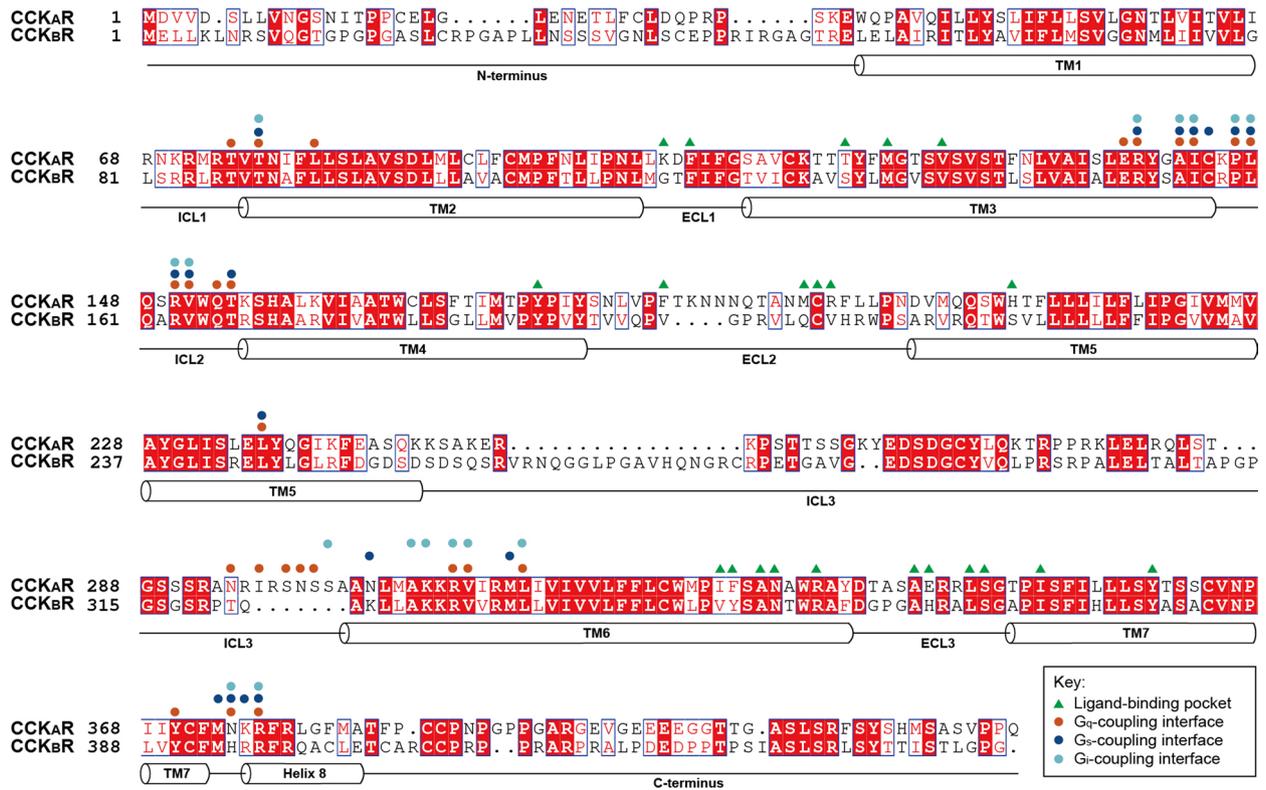
Extended Data Fig. 4 | Sequence alignment of CCK receptors. Helical secondary structures are shown based on CCK$_A$R. Residues involved in ligand-binding are labeled with green triangles. Residues involved in G protein coupling are labeled with circles (orange, G$_q$; blue, G$_s$; cyan, G$_i$).
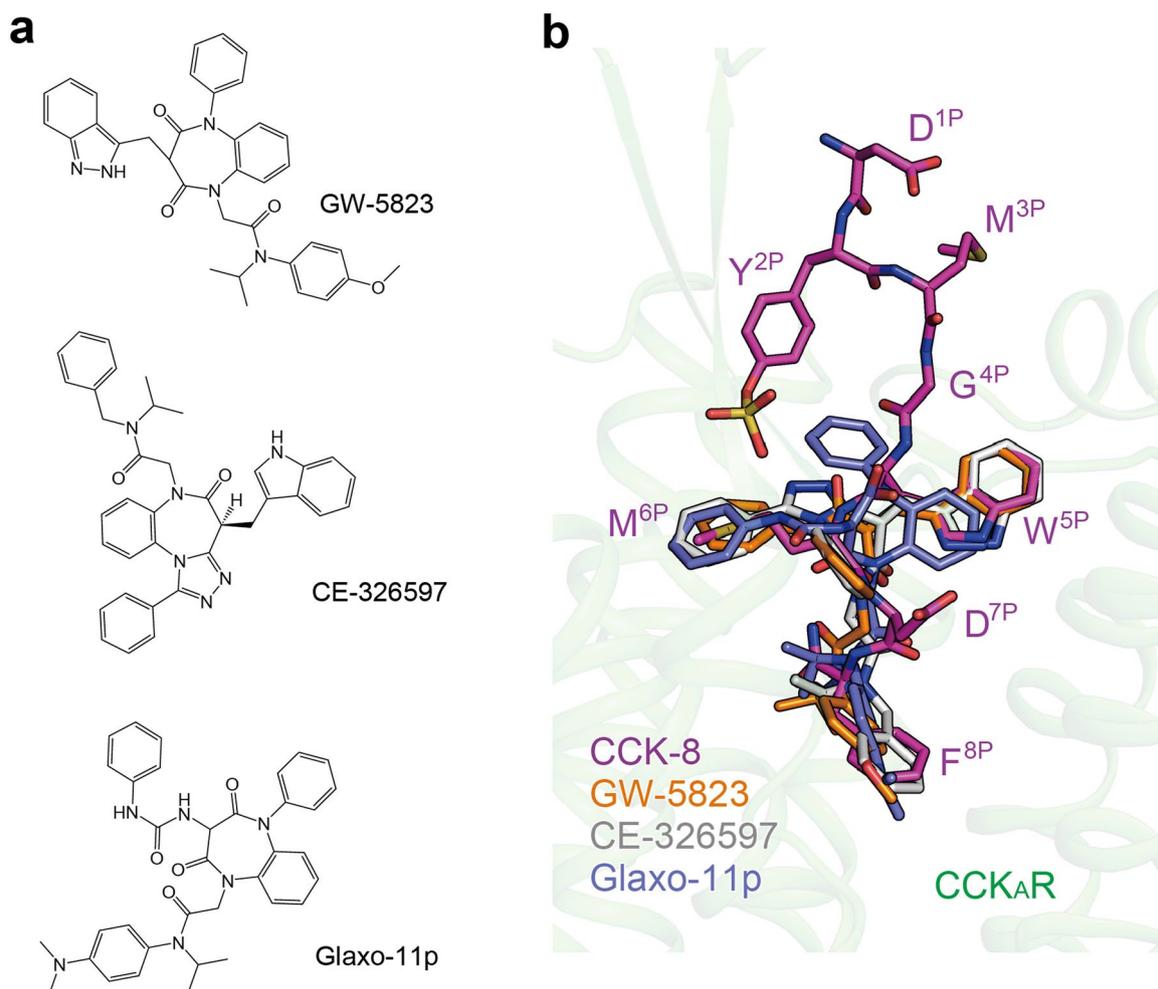
Extended Data Fig. 5 | Molecular docking of small molecule agonists to the CCK$_A$R structure. a, Chemical structures of small molecule agonists of CCK$_A$R. **b**, Comparison of the binding poses of three agonists with CCK-8. CCK-8, magenta; GW-5823, orange; CE-326597, grey; Glaxo-11p, purple blue. CCK-8 and small molecule agonists are shown as sticks. The amino acids of CCK-8 are labelled.
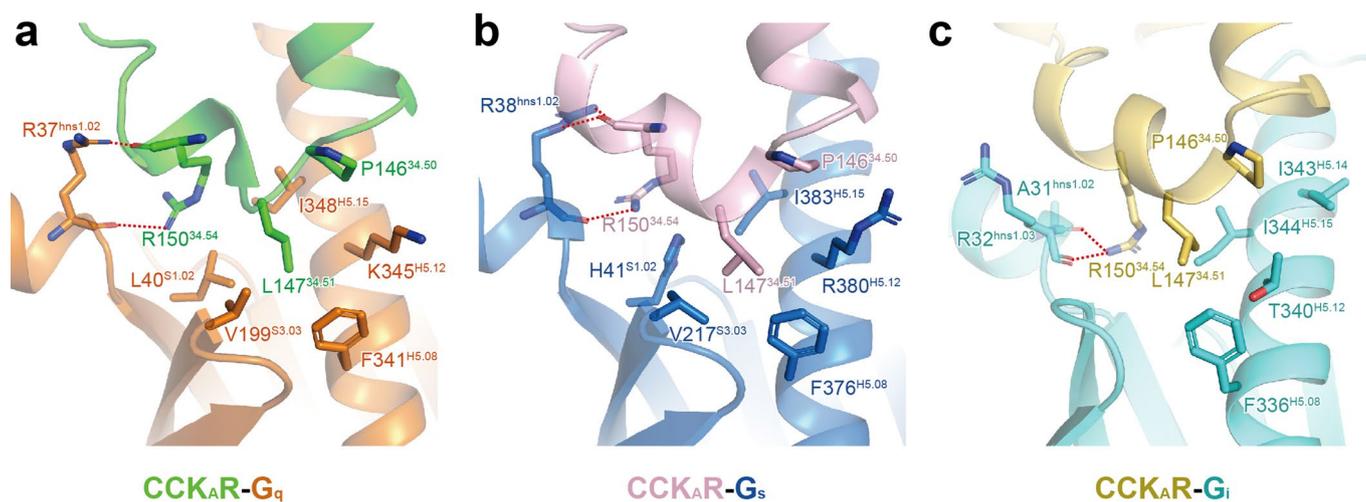
Extended Data Fig. 6 | The interface between CCK$_A$R ICL2 and different G proteins. Detailed interaction between the receptor and Gα$_q$ (**a**), Gα$_s$ (**b**), and Gα$_i$ (**c**) are shown. Side chains of related residues are shown as sticks.
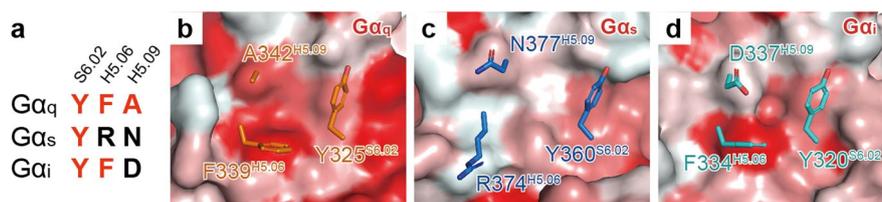
Extended Data Fig. 7 | Comparison of the hydrophobic patch in Gα$_q$ subunit to the corresponding sites in other G proteins. a, Sequence alignment of S6.02, H5.06, and H5.09 from Gα$_q$, Gα$_s$, and Gα$_i$ subunits. Residues at positions S6.02, H5.06, and H5.09 comprise the hydrophobic patch to interact with CCK$_A$R ICL3. **b-d**, Surface presentation of the patch by hydrophobicity. Side chains of residues at positions S6.02, H5.06, and H5.09 in Gα$_q$ (**b**), Gα$_s$ (**c**), and Gα$_i$ (**d**) subunits are shown.

Corresponding author(s): Yi Jiang

Last updated by author(s): Jun 11, 2021

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's d, Pearson's r), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Automated data collection on the Titan Krios was performed using serialEM 3.7.11.
Data analysis	The following software was used in this study: PyMol 2.3.2, MotionCor2.1, Ctffind4, Relion 3.0, UCSF Chimera v1.13.1, UCSF ChimeraX 1.1, Phenix 1.18.2, Coot 0.9 EL, ISODLE1.1, Prism8, NovoExpress 1.2.1.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:
- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Density maps and structure coordinates have been deposited in the Electron Microscopy Data Bank (EMDB) and the Protein Data Bank (PDB) with accession codes EMD-31387 and 7EZH for the CCK-8–CCKAR–Gi–scFv16 complex; EMD-31388 and 7EZK for the CCK-8–CCKAR–Gs complex; EMD-31389 and 7EZM for the CCK-8– CCKAR–Gq–scFv16 complex. All relevant data are available from the authors and/or included in the manuscript or Supplementary Information.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences   ☐ Behavioural & social sciences   ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	For cryo-EM data, images were collected until the resolution and 3D reconstruction converges. For all the functional assay, no statistical approaches were used to predetermine the sample size. We use sample size at least of three independent experiments, commonly exploited by researchers in this field.
Data exclusions	No data were excluded.
Replication	Each experimental findings were reliably reproduced within one month.
Randomization	Randomization was not relevant to this study, as the data were collected automatically and did not involve choosing.
Blinding	Blinding was not relevant to this study, since all the data were collected automatically.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	Monoclonal Anti-Flag M2 antibody (Sigma-Aldrich, Cat#F3165) Donkey anti-Mouse IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor 488 (ThermoFisher Scientific, Cat#A21202)
Validation	Monoclonal Anti-Flag M2 antibody: https://www.sigmaaldrich.com/catalog/product/sigma/f3165?lang=zh&region=CN; Donkey anti-Mouse IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor 488: https://www.thermofisher.com/cn/zh/antibody/product/Donkey-anti-Mouse-IgG-H-L-Highly-Cross-Adsorbed-Secondary-Antibody-Polyclonal/A-21202. All antibodies were commercially purchased and have been validated by vendors. Validation data are available from the respective manufacturer's website.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	Sf9 (Expression Systems, Cat#94-001F) HEK 293T/17 (Cell Bank at the Chinese Academy of Sciences)
Authentication	Used as expression stains only, independent verification after purchase not required.
Mycoplasma contamination	Cell lines were tested and free from mycoplasma contamination.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell lines were used.

Flow Cytometry

Plots

Confirm that:

☐ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☐ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☐ All plots are contour plots with outliers or pseudocolor plots.

☐ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	Sample preparation listed in Methods.
Instrument	NovaCyte Flow Cytometer, ACEA Biosciences
Software	NovoExpress 1.2.1
Cell population abundance	Approximately 15,000 cellular events were collected and the total fluorescence intensity of positive expression cell population was calculated.
Gating strategy	Gating was determined by the Alexa-488 fluorescence intensity to differentiate positive cells and all other cells.

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.