Structure and activation of GLP-1R by a small molecule agonist

Honglei Ma[1], Wei Huang[2], Xiaoxi Wang[1], Lihua Zhao[1], Yi Jiang[1], Feng Liu[3], Wei Guo[2], Xianqiang Sun[2], Wenge Zhong[2], Daopeng Yuan[2]& H. Eric Xu[1]

[1]The CAS Key Laboratory of Receptor Research, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai 201203, China;
[2]Qilu Regor Therapeutics Inc., 1206 Zhangjiang Road, Building C, Pu Dong New District, Shanghai 201210, China;
[3]Regor Pharmaceuticals Inc., 50 Soldiers Field Pl, Boston, MA 02135, USA.

Correspondence: Daopeng Yuan (Daopeng.yuan@qlregor.com) or H. Eric Xu (Eric.xu@simm.ac.cn)

Supplementary information

Supplementary information, Data S1

Materials and methods

Constructs of GLP1-R and Gs heterotrimer

The human GLP-1R (residues 24-463) was cloned into pFastBac vector with its native signal peptide replaced by that of haemagglutinin (HA) to enhance receptor expression. A cleavable 8xHis tag was added at the C terminus of human GLP-1R, which can be removed by Tev protease. A dominant-negative Gαs (DNGαs) generated as described was cloned into pFastBac vector for the complex formation[1]. Rat Gβ1 and bovine Gγ2 were cloned into the pFastBac vector, respectively. Site-directed mutagenesis was carried out using the quick-change PCR. Mutations and all plasmid constructs were confirmed by DNA sequencing before use.

Expression and purification of GLP-1R-Gs complex

The human GLP-1R, DNGαs, Gβ1 and Gγ2 were co-expressed in Sf9 insect cells (Invitrogen) at an equal MOI using Bac-to-Bac baculovirus system. For the GLP-1R-Gs complex expression, cell cultures ware grown in protein-free insect cell culture medium (Expression Systems ESF 921) to a density of 2-3 $\times 10^6$ cells per mL and then infected with four separate viruses of human GLP1-R, DNGαs, Gβ1 and Gγ2 at an equal MOI. The infected cells were cultured at 27 ℃ for 48 h before collection by centrifugation. The pellets were stored at -80 ℃ or until use. The cell pellets were

lysed in the buffer of 20 mM HEPEs, 100 mM NaCl, 10 mM $MgCl_2$, 5 mM $CaCl_2$, pH 7.5, supplemented with Protease Inhibitor Cocktail, EDTA-Free (Bimake). The GLP1-R-Gs complex was formed in membranes by the addition of 10 μM compound RGT1383, 10 μg/mL Nb35-His and 25 mU/mL Apysase (Sigma) and incubated for 1.5 h at room temperature. The complex in the membrane was solubilized using 0.5% (w/v) n-Dodecyl β-D-maltoside (DDM, Anatrace), 0.01% (w/v) lauryl maltose neopentylglycol (LMNG, Anatrace) and 0.1% (w/v) cholesterol hemisucinate (CHS, Anatrace) for 2 h at 4 ℃. The supernatant was isolated by centrifugation at 65000 ×g for 35 min, and the solubilized complex was incubated with NiNTA resin for 1 h at 4 ℃. The resin was packed into a gravity flow column and washed with 20 column volumes of washing buffer containing 20 mM HEPEs, 100 mM NaCl, pH 7.5, 2 mM $MgCl_2$, 10 μM compound RGT1383 and 0.01% (w/v) LMNG, 0.01% (w/v) glyco-diosgenin (GDN, Anatrace), and 0.004% (w/v) CHS before bound material was eluted in buffer with 200 mM Imidazole. The protein complex was then concentrated using an Amicon Ultra Centrifugal Filter (MWCO 100 kDa) and loaded onto Superdex200 10/300 GL Increase column (GE Healthcare) with running buffer 20 mM HEPES, 100 mM NaCl, pH 7.5, 2 mM $MgCl_2$, 10 μM compound RGT1383, 0.00075% (w/v) LMNG, 0.00025% GDN and 0.0002% (w/v) CHS. The fractions for the monomeric GLP1-R-Gs complex were collected and concentrated individually for electron microscopy experiments.

Expression and purification of Nb35

Nb35 was expressed in the periplasm of E.coli strain BL21 and purified by NiNTA according to previously described methods[1] followed by size-chromatography using a Hiload 16/600 Superdex 75 column (GE healthcare) with running buffer of 20 mM HEPES, 100 mM NaCl, pH 7.5. Target fractions of Nb35 were concentrated to ~5 mg/mL and flash frozen and stored at -80 ℃.

Preparation of vitrified specimen and data acquisition

The purified human GLP-1R-Gs complex samples (9.5 mg/ml) were prepared by applying an aliquot of 3.5 μl protein sample to the glow-discharged Quantifoil holy carbon grid (Quantifoil R 1.2/1.3, 200 mesh), and blotted with filter paper for 3.0 s then plunge-frozen in liquid ethene using a FEI Vitrobot Mark IV. Cryo-EM imaging was performed on a 300 kV Titan Krios microscope (FEI) equipped with a Gatan energy filter (operated with a slit width of 20 eV) (GIF) and K3

Summit direct detection camera. The microscope was operated at a calibrated magnification of ×81000, yielding a pixel size of 1.045 Å on micrographs. In total, 2518 micrographs were collected at an electron dose rate of ~25.3 e–/Å^2 s with a defocus range of −1.2 to −2.2 μm. An accumulated dose of 77 e–/Å^2 on the sample was fractionated into a movie stack of 36 image frames.

Data processing

All movie frames were motion-corrected using MotionCor2[2]. The contrast-transfer function (CTF) parameters were estimated with CTFFIND4[3]. For the GLP-1R-Gs complex, a total of 3,047,428 particles were automatically picked from 1600 images. Particle selection, 2D classification, 3D classification and refinement were performed using RELION-3.0.8[4]. A data set of 297933 particles were subjected to 3D refinements, yielding a final map with a global nominal resolution at 4.2 Å by the 0.143 criteria of the gold-standard Fourier Shell correlation (FSC). Half-reconstructions was used to determine the local resolution of each map[4].

Model building and refinement

The initial template of the complex was built by using the cryo-EM structure of GLP-1R-Gs (Protein Data Bank accession (PDB): 6B3J). The model was docked into the EM density map, followed by iterative manual adjustment and real-space refinement using COOT[5] and PHENIX[6]. Model overfitting was evaluated against one cryo-EM half map. The final refinement statistics for the model are provided in Supplementary information, Table S1.

Molecular docking

The ligand molecule was docked to GLP-1R based on solved cryo-EM structure using the Schrödinger Drug Discovery suite (http://www.schrodinger.com). Molecular docking was performed in Maestro 12.3 using Glide. The ligand molecule was prepared using Ligprep, and the low energy conformations were sampled using MacroModel Conformational Search. The cryo-EM structure of GLP-1R was prepared for docking using Protein Preparation Wizard. The bond orders and formal charges were added for hetero groups, and hydrogens were added to all atoms in the structure. After preparation, the hydrogen atom positions were optimized with restrained

minimization using the OPLS3e force field. The minimization was terminated when the energy converged or the RMSD reached a maximum cutoff of 0.30 Å. The extra precision (XP) docking mode for a set of low energy conformations was performed on the generated grid of the protein structure.

Transfections for cAMP luciferase reporter assays

Codon-optimized cDNAs of human GLP-1R (residues 24–463) lacking the predicted signal peptide coding region were synthesized by Genewiz. The receptor constructs were fused with a human IgG leader (MGWSCIILFLVATATGVHSE) at their N-termini for cell membrane localization. The fusion proteins were expressed in 293T cells. 293T cells were plated at a density of 2×10^4 per well in 96-well plates one day before transfection. Cells were transiently co-transfected with GLP-1R expression plasmid (50 ng) and the luciferase reporter constructs with cAMP response element (CRE, 50 ng) and TK -Renilla (10 ng) using FuGENE@HD transfection reagent (Promega) at a ratio of 2:1 (reagent:DNA). Compounds were added to the cells to a final concentration of 10 μM 4 h before collection. All experiments were performed as three independent transfection experiments.

Cell-based assay for determining the EC_{50} of compound

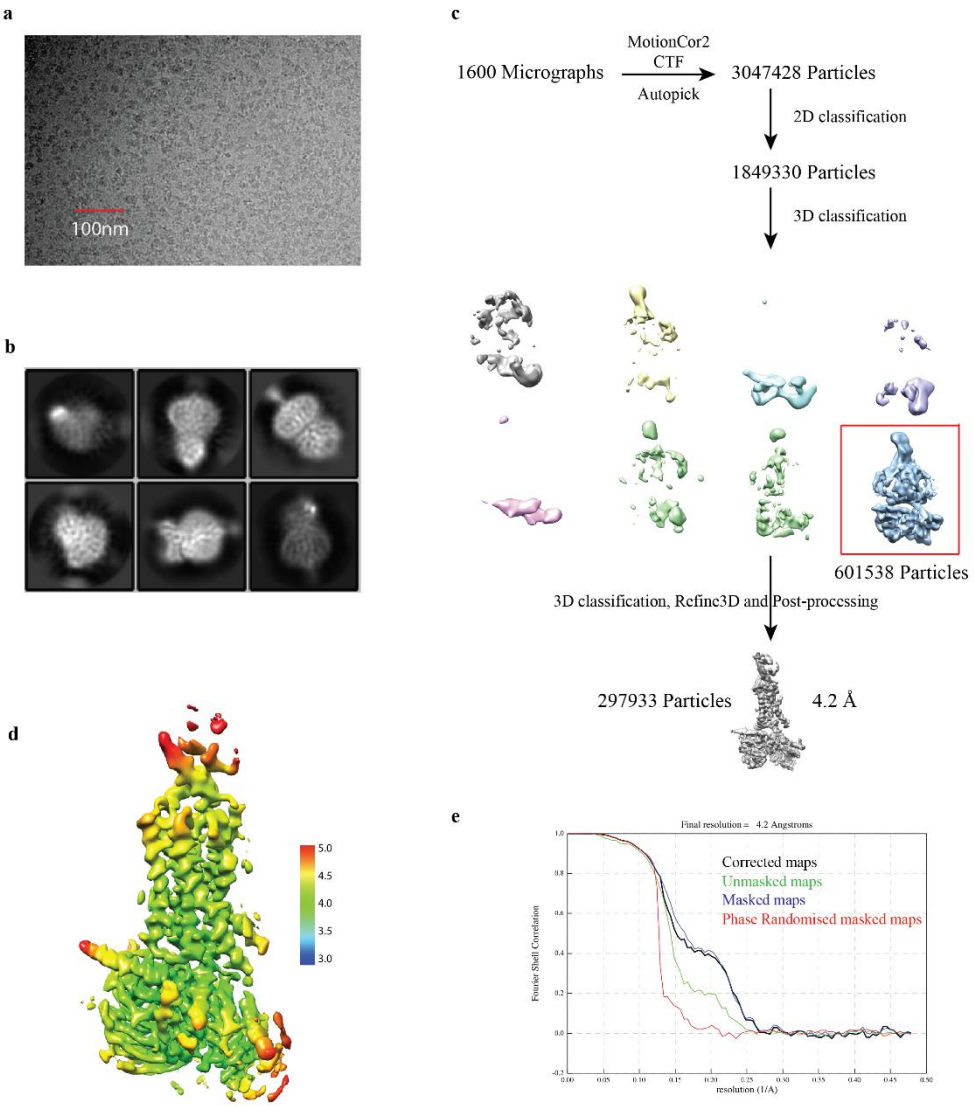
The activation of GLP-1R induced by compound RGT1383 was evaluated using 293T cells that expressed human GLP-1R and NanoLuc luciferase under control of a cAMP response element (CRE; pNL-NlucP-CREHygro, Promega). 293T cells were plated at a density of 10^4 per well in 96-well white plates (Greiner Bio-one). The second day cells were transiently co-transfected with GLP-1R expression plasmid (50 ng) and the luciferase reporter constructs with cAMP response element (CRE, 50 ng) and TK -Renilla (10 ng) using FuGENE@HD Transfection reagent(Promega) at a ratio of 2:1 (reagent:DNA). Compound with a series of concentrations was added to the cells 4 h before collection. All experiments were performed with three independent treatments of compound. The cAMP signal of GLP-1R/NlucP expressed 293T cells was determined using the Nano-Glo Luciferase Assay System (Promega) according to the manufacturer's instructions. The EC_{50} values (50% maximal effective concentration) were calculated using GraphPad Prism software version 8.0 (La Jolla, CA, USA).

cAMP accumulation assays

GLP-1R-mediated agonist activity was determined with a cell-based functional assay utilizing an HTRF (Homogeneous Time-Resolved Fluorescence) cAMP detection kit (cAMP Dynamic 2 Assay Kit; CisBio cat #62AM4PEC) that measures cAMP levels in the cell.

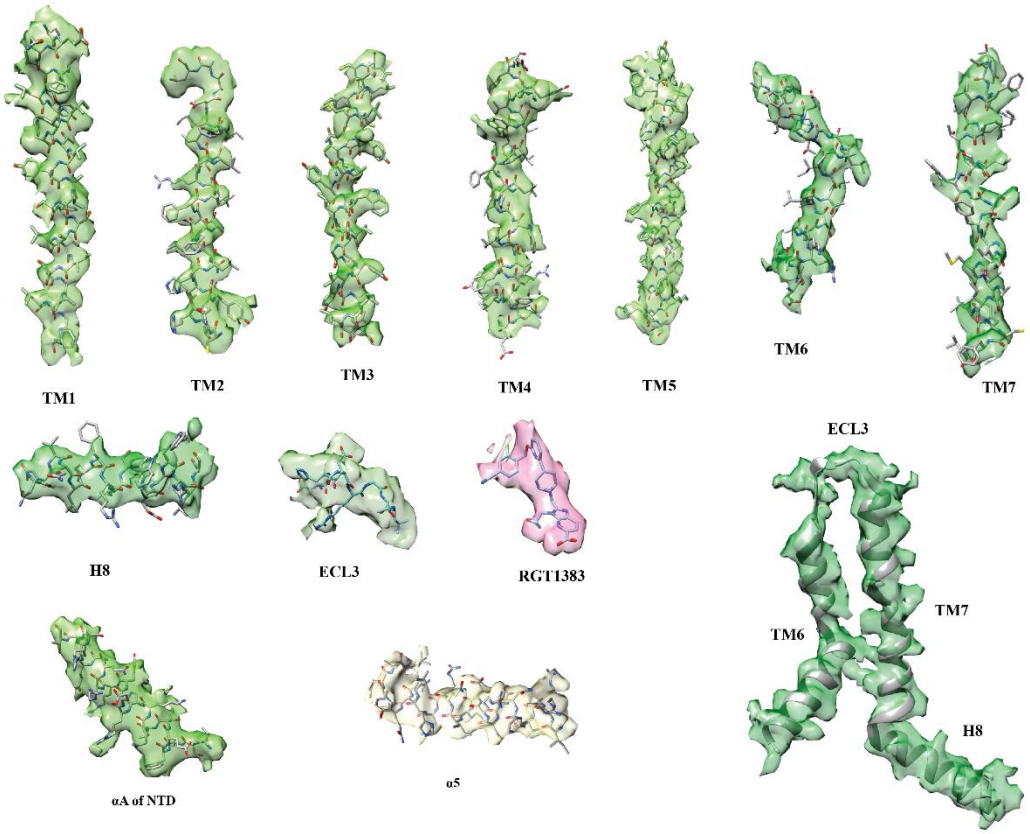
β-Arrestin recruitment assays

GLP-1R agonist activity with respect to β-arrestin recruitment was determined with a cell-based functional assay using PathHunter eXpress GLP-1R CHO-K1 β-Arrestin GPCR Assay kit (DiscoverX Cat#93-0300E2CP0M).



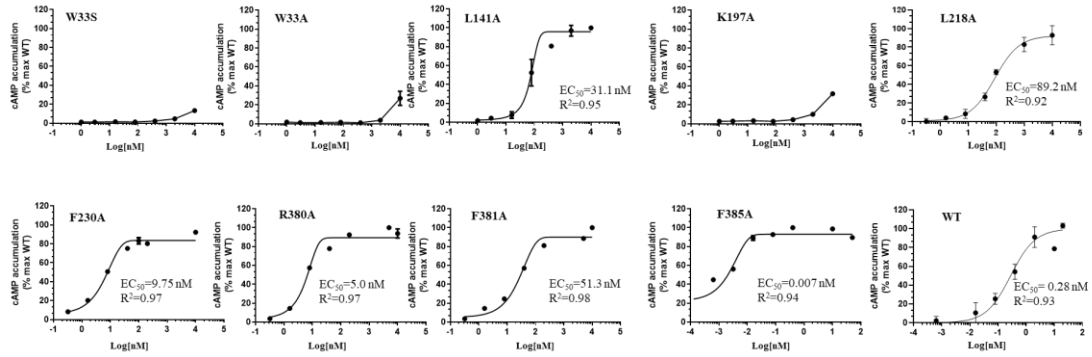
Supplementary information, Figure S1

Fig. S1 Cryo-EM data processing of the RGT1383 bound GLP-1R-Gs complex. **a,** Representative cryo-EM image of the complex. Scale bar, 100 nm. **b,** Representative 2D class averages of the complex showing different views. **c,** Workflow for image processing. **d,** Local resolution of the complex was shown in color-coded map. **e,** FSC curves of the final reconstruction showing the overall nominal resolution at 4.2 Å.



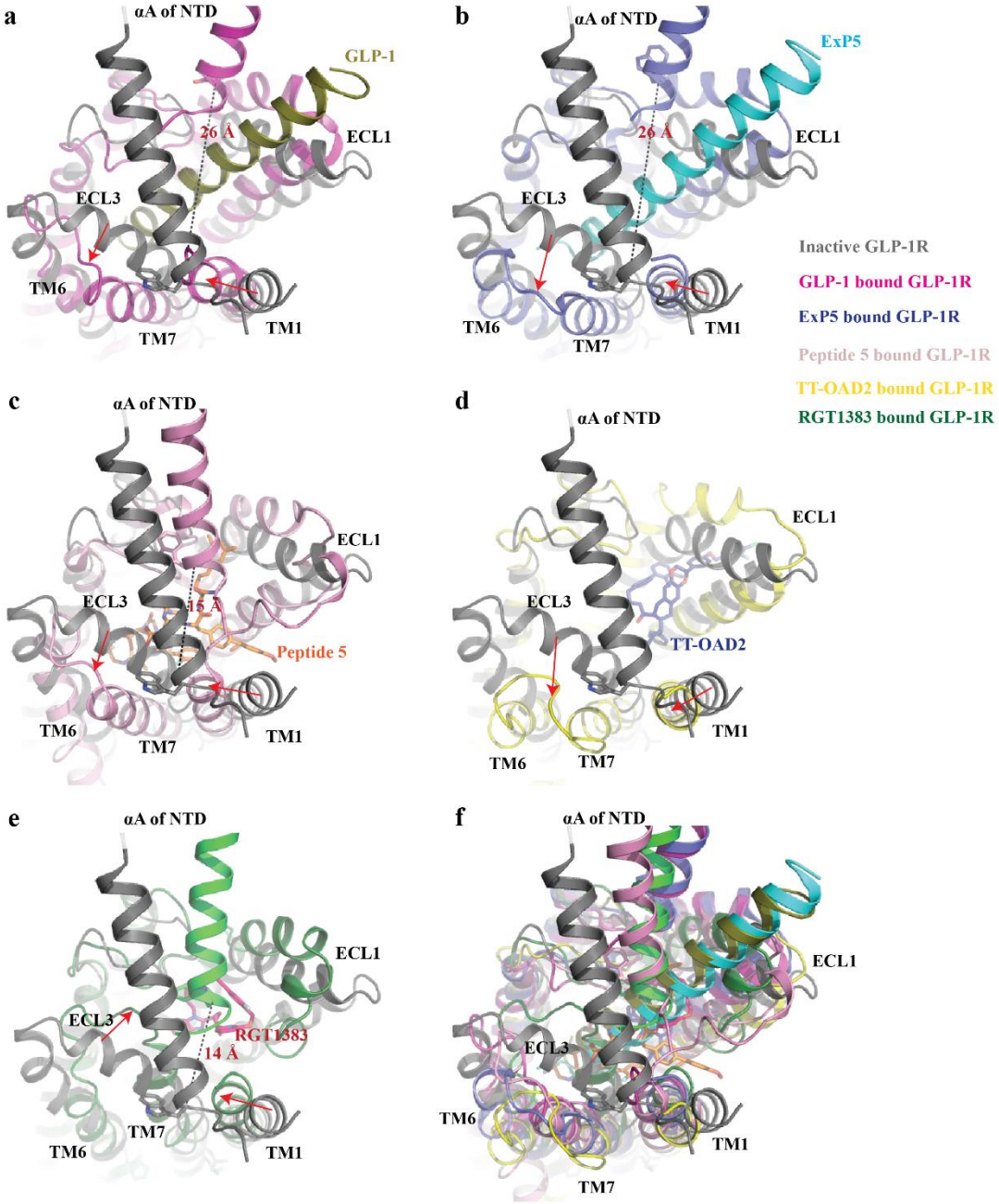
Supplementary information, Figure S2

Fig. S2 Cryo-EM map of RGT1383 bound GLP-1R-Gs complex. Cryo-EM density map and model are shown for all seven transmembrane α-helices, ECL3 and αA of NTD from the GLP-1R, $G\alpha_s$ helices α5, RGT1383.



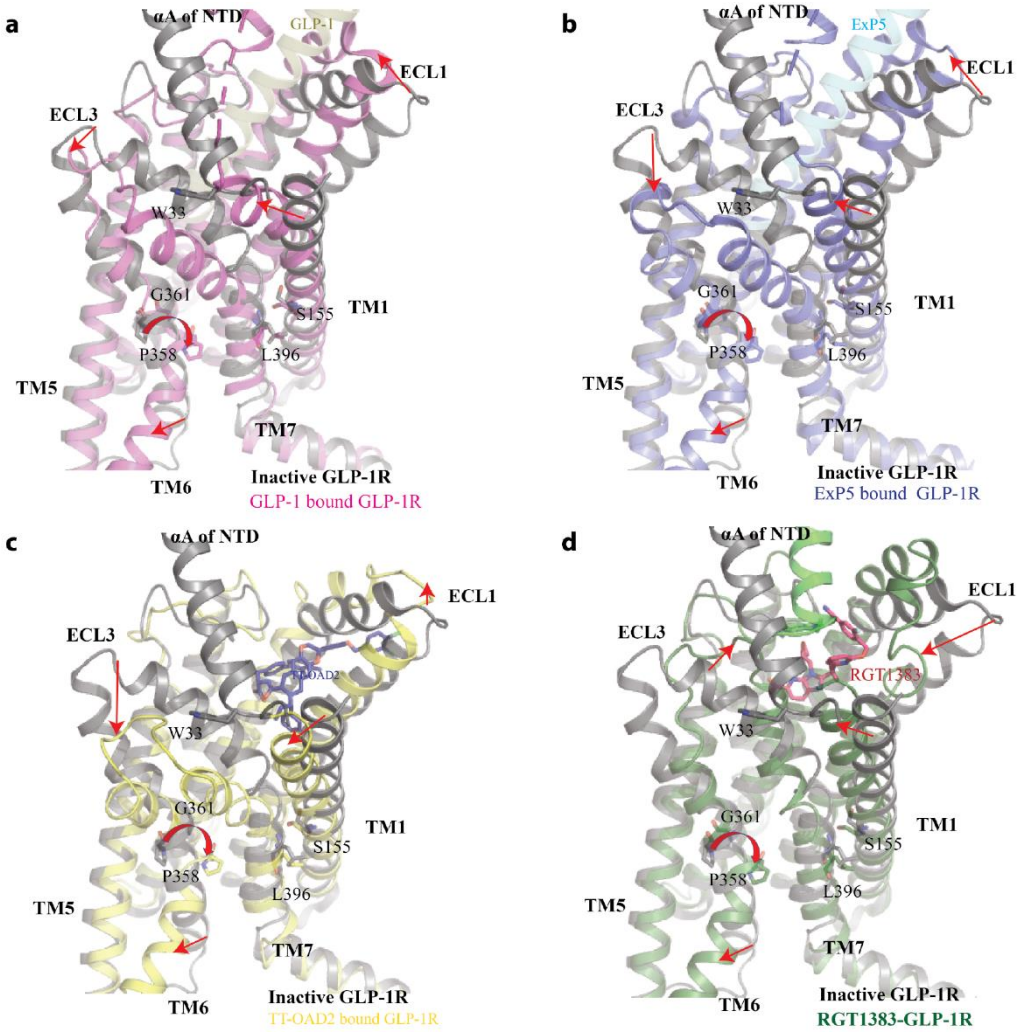
Supplementary information, Figure S3

Fig. S3 Dose-dependent cAMP signals were produced by activation of the full-length GLP-1R receptor in 293T cells that express hGLP-1R and CRE-NanoLuc luciferase reporter. The activity of the WT receptor exposed to 20 nM RGT1383 compound was set to 100%. All data are presented normalized as % max WT GLP-1R; n=3, error bars = S.D. The EC_{50} values for the compounds are noted as log nM concentrations of the compounds determined by the cAMP luciferase reporter assays in the figures.



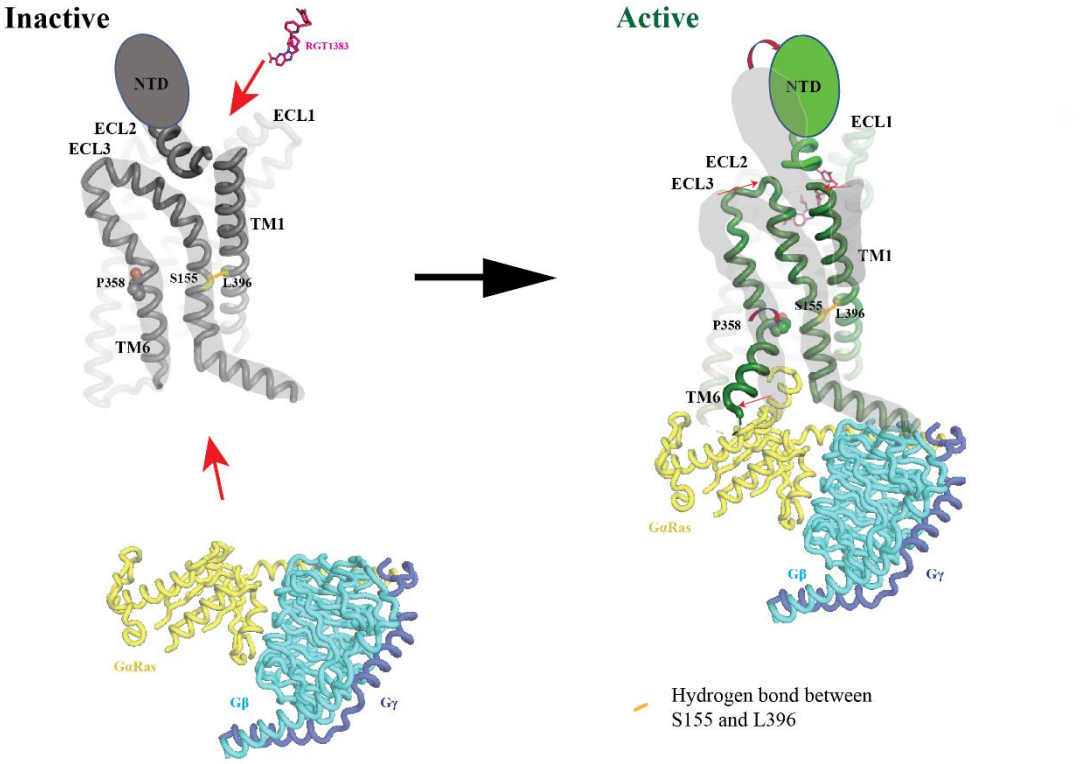
Supplementary information, Figure S4

Fig. S4 Comparison of inactive GLP-1R structure (gray; PDB code: 6ln2) with agonists bound GLP-1R structures showing the movement of αA of NTD and extracellular region in the top view. **a,** Comparison with the GLP-1 (deep olive) bound GLP-1R-Gs complex (light magenta; PDB code: 5vai). **b,** Comparison with the ExP5 (cyan) bound GLP-1R-Gs complex (light blue; PDB code: 6b3j). **c,** Comparison with the Peptide 5 (orange) bound GLP-1R (pink; PDB code: 5nx2). **d,** Comparison with the TT-OAD2 (deep blue) bound GLP-1R-Gs complex (yellow; PDB code: 6ORV). **e,** Comparison with the RGT1383 (hot pink) bound GLP-1R-Gs complex (deep green; PDB code: 7C2E). **f,** Overlap of all structures. Arrows show the extent of motion.



Supplementary information, Figure S5

Fig. S5 Comparison of inactive GLP-1R structure (gray; PDB code: 6ln2) with agonists bound GLP-1R-Gs structures showing the movement of GLP-1R in the side view. **a,** Comparison with the GLP-1 (deep olive) bound GLP-1R-Gs complex (light magenta; PDB code: 5vai). **b,** Comparison with the ExP5 (cyan) bound GLP-1R-Gs complex (light blue; PDB code: 6b3j). **c,** Comparison with the TT-OAD2 (deep blue) bound GLP-1R-Gs complex (yellow; PDB code: 6ORV). **d,** Comparison with the RGT1383 (hot pink) bound GLP-1R-Gs complex (deep green; PDB code: 7C2E). Arrows show the extent of motion.



Supplementary information, Figure S6

Fig. S6 Cartoon representation of GLP-1R activation by RGT1383. RGT1383 bound GLP-1R facilitates the movement of NTD toward ECL2, followed by inward movement of extracellular TM1 and rearrangement of TM7 which consequently lead to rearrangement of ECL3, downward rotation of extracellular half of TM6 (measured from P358) and subsequent outward movement of intracellular part of TM6 for the accommodation of Gs protein.

Supplementary information, Video S1

Structural comparison of RGT1383 (hot pink) bound GLP-1R (deep green) with inactive GLP-1R (gray).

Supplementary information, Video S2

Structural comparison of RGT1383 (hot pink) bound GLP-1R (deep green) with GLP-1 (deep olive) bound GLP-1R (light magenta).

Supplementary information, Table S1

Cryo-EM data collection, refinement and validation statistics

	GLP-1R-Gs complex (EMDB-30274) (PDB 7C2E)
Data collection and processing	
Magnification	81000
Voltage (kV)	300
Electron exposure (e–/Å^2)	25.3
Defocus range (μm)	-1.2 to -2.2
Pixel size (Å)	1.045
Symmetry imposed	C1
Initial particle images (no.)	3047428
Final particle images (no.)	297933
Map resolution (Å)	4.2
FSC threshold	0.143
Map resolution range (Å)	3.5-5.0
Refinement	
Initial model used (PDB code)	6B3J
Model resolution (Å)	4.2
FSC threshold	0.143
Model resolution range (Å)	3.5-5.0
Map sharpening *B* factor (Å^2)	-240
Model composition	
Non-hydrogen atoms	8828
Protein residues	8787
Ligands	41
B factors (Å^2)	
Protein	123.80
Ligand	79.74
R.m.s. deviations	
Bond lengths (Å)	0.003
Bond angles (°)	0.745
Validation	
MolProbity score	1.52
Clashscore	4.02
Poor rotamers (%)	0
Ramachandran plot	
Favored (%)	95.15
Allowed (%)	4.85
Disallowed (%)	0

References

1 Liang, Y. L. *et al.* Phase-plate cryo-EM structure of a biased agonist-bound human GLP-1 receptor-Gs complex. *Nature* **555**, 121-125, doi:10.1038/nature25773 (2018).

2 Zheng, S. Q. *et al.* MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat Methods* **14**, 331-332, doi:10.1038/nmeth.4193 (2017).

3 Mindell, J. A. & Grigorieff, N. Accurate determination of local defocus and specimen tilt in electron microscopy. *J Struct Biol* **142**, 334-347, doi:10.1016/s1047-8477(03)00069-8 (2003).

4 Scheres, S. H. RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J Struct Biol* **180**, 519-530, doi:10.1016/j.jsb.2012.09.006 (2012).

5 Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126-2132, doi:10.1107/S0907444904019158 (2004).

6 Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-221, doi:10.1107/S0907444909052925 (2010).