doi:10.1038/nature22378

Human GLP-1 receptor transmembrane domain structure in complex with allosteric modulators

Gaojie Song[1], Dehua Yang[2]*, Yuxia Wang[1]*, Chris de Graaf[3]*, Qingtong Zhou[1], Shanshan Jiang[4], Kaiwen Liu[1,5,6], Xiaoqing Cai[2], Antao Dai[2], Guangyao Lin[5], Dongsheng Liu[1], Fan Wu[1,5,6], Yiran Wu[1], Suwen Zhao[1,5], Li Ye[4], Gye Won Han[7], Jesper Lau[8], Beili Wu[5,6,9], Michael A. Hanson[10], Zhi-Jie Liu[1,5,11], Ming-Wei Wang[2,4,5] & Raymond C. Stevens[1,5]

The glucagon-like peptide-1 receptor (GLP-1R) and the glucagon receptor (GCGR) are members of the secretin-like class B family of G-protein-coupled receptors (GPCRs) and have opposing physiological roles in insulin release and glucose homeostasis[1]. The treatment of type 2 diabetes requires positive modulation of GLP-1R to inhibit glucagon secretion and stimulate insulin secretion in a glucose-dependent manner[2]. Here we report crystal structures of the human GLP-1R transmembrane domain in complex with two different negative allosteric modulators, PF-06372222 and NNC0640, at 2.7 and 3.0 Å resolution, respectively. The structures reveal a common binding pocket for negative allosteric modulators, present in both GLP-1R and GCGR[3] and located outside helices V–VII near the intracellular half of the receptor. The receptor is in an inactive conformation with compounds that restrict movement of the intracellular tip of helix VI, a movement that is generally associated with activation mechanisms in class A GPCRs[4–6]. Molecular modelling and mutagenesis studies indicate that agonist positive allosteric modulators target the same general region, but in a distinct sub-pocket at the interface between helices V and VI, which may facilitate the formation of an intracellular binding site that enhances G-protein coupling.

Glucagon-like peptide-1 (GLP-1) is one of the key incretin hormones secreted in response to food intake and gastric motility, and is responsible for glucose homeostasis via the stimulation of insulin secretion through activation of GLP-1R[1]. Peptide analogues of GLP-1 have been successfully developed for the treatment of type 2 diabetes[2], but the development of therapeutically viable non-peptidic GLP-1R agonists has been unsuccessful. Agonist positive allosteric modulators (PAMs) have been identified and used to investigate ligand-directed biased cellular signalling of GLP-1R[7–9]. Previous studies have provided evidence for a two-domain binding mechanism of GLP-1 with its cognate receptor[1,10–13]. Structures of the transmembrane domain (TMD) of the class B GPCRs corticotropin-releasing factor receptor 1 (CRF_1R)[14] and GCGR[3,15] have been reported. Despite these recent advances in structural characterization, little is known about the molecular mechanisms of positive and negative allosteric modulation of GLP-1R and GCGR. To provide a foundation for the discovery of therapeutic agents that allosterically target the GLP-1R and GCGR signalling pathways, we have solved structures of the human GLP-1R TMD in complex with two negative allosteric modulators (NAMs), and complemented these structures with mutagenesis and modelling studies to map the binding site and further our understanding of the activation mechanism for agonist PAMs of GLP-1R.

To facilitate crystallization of the GLP-1R TMD, we generated a thermostabilized construct with 10 mutations, including a disulfide bond (I317$^{5.47b}$C–G361$^{6.50b}$C) (numbers in superscript refer to the Wootten numbering system for class B GPCRs[16]) that links the middle regions of helices V and VI and a GCGR mimicking mutation C347$^{6.36b}$F in the allosteric modulator binding pocket that stabilizes the interaction interface for NAMs (Methods and Extended Data Fig. 1). These NAMs were previously optimized for GCGR antagonism, but certain analogues were found to antagonize GLP-1R as well. The final modified construct yielded crystals that diffracted to 2.7 Å for PF-06372222 and 3.0 Å for NNC0640 (Fig. 1a, c and Extended Data Table 1).

The TMD architecture of GLP-1R is similar to that of GCGR, consistent with the similarity in their primary sequences (45% identical in the TMDs; Fig. 1b). GLP-1R preserves the conserved and functionally important[17] disulfide bond C226$^{3.29b}$–C296^{ECL2}, and contains most of the interhelical hydrogen bonds present in other class B structures[3,14,15] (Extended Data Fig. 2). Helix I of GLP-1R is 2.5 helical-turns shorter than the long stalk region found in the initial structure of the GCGR TMD (PDB code 4L6R)[15] (Fig. 1b). In an accompanying paper[18], the full-length structure of GCGR reveals a rearrangement of the stalk region with the N-terminal helix unwinding to form an extended β-sheet with two strands from the first extracellular loop (ECL1), suggesting a degree of flexibility in this region that may be associated with the functional mechanisms of class B GPCRs.

The NAMs PF-06372222 and NNC0640 bind in a similar pocket outside of helices V–VII as MK-0893 in the crystal structure of the thermostabilized GCGR TMD (PDB code 5EE7)[3] (Fig. 2a–c). The anionic carboxylic acid (PF-06372222) and tetrazole (NNC0640) moieties of the NAMs target a polar cleft between helices VI and VII, and form hydrogen-bond interactions with S352$^{6.41b}$ and N406$^{7.61b}$ in GLP-1R. MK-0893 forms similar interactions with homologous residues in GCGR, and makes an additional hydrogen bond to the side chain of R346$^{6.37b}$ in GCGR. PF-06372222 and NNC0640 each form a hydrogen bond with T355$^{6.44b}$ in GLP-1R, directing their hydrophobic dimethyl cyclobutane (PF-06372222) and cyclohexyl (NNC0640) moieties parallel to helix VI and the lipid bilayer. By contrast, MK-0893 does not hydrogen bond with the homologous T353$^{6.44b}$ in GCGR, and its dichlorophenyl moiety is directed perpendicular to helix VI and the lipid bilayer. The trifluoromethyl-pyrazole group of PF-06372222 binds a hydrophobic surface area of 43 Å^2 consisting of I328$^{5.58b}$, V331$^{5.61b}$, V332$^{5.62b}$ and L335$^{5.65b}$ of helix V (Fig. 2a), which is also targeted by the methylsulfone-phenyl group of NNC0640 in GLP-1R (33 Å^2 buried surface area; Fig. 2b), but only partially interacts with MK-0893 (9 Å^2; Fig. 2c)[3] and NNC0640 (4 Å^2)[18] in GCGR (Fig. 2e).

The structures demonstrate that these NAMs can accommodate different binding modes, including the variation in hydrophobic

---

[1]iHuman Institute, ShanghaiTech University, 393 Middle Huaxia Road, Shanghai 201210, China. [2]The National Center for Drug Screening and the CAS Key Laboratory of Receptor Research, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, 189 Guo Shou Jing Road, Shanghai 201203, China. [3]Division of Medicinal Chemistry, Faculty of Sciences, Amsterdam Institute for Molecules, Medicines and Systems (AIMMS), Vrije Universiteit Amsterdam, De Boelelaan 1108, 1081 HZ Amsterdam, The Netherlands. [4]School of Pharmacy, Fudan University, 826 Zhang Heng Road, Shanghai 201203, China. [5]School of Life Science and Technology, ShanghaiTech University, 393 Middle Huaxia Road, Shanghai 201210, China. [6]University of Chinese Academy of Sciences, No. 19A Yuquan Road, Beijing 100049, China. [7]Department of Chemistry, Bridge Institute, University of Southern California, 3430 S. Vermont Avenue, Los Angeles, California 90089, USA. [8]Novo Nordisk, Måløv DK-2760, Denmark. [9]The CAS Key Laboratory of Receptor Research, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, 555 Zuchongzhi Road, Shanghai 201203, China. [10]GPCR Consortium, San Marcos, California 92078, USA. [11]Institute of Molecular and Clinical Medicine, Kunming Medical University, Kunming 650500, China.
*These authors contributed equally to this work.

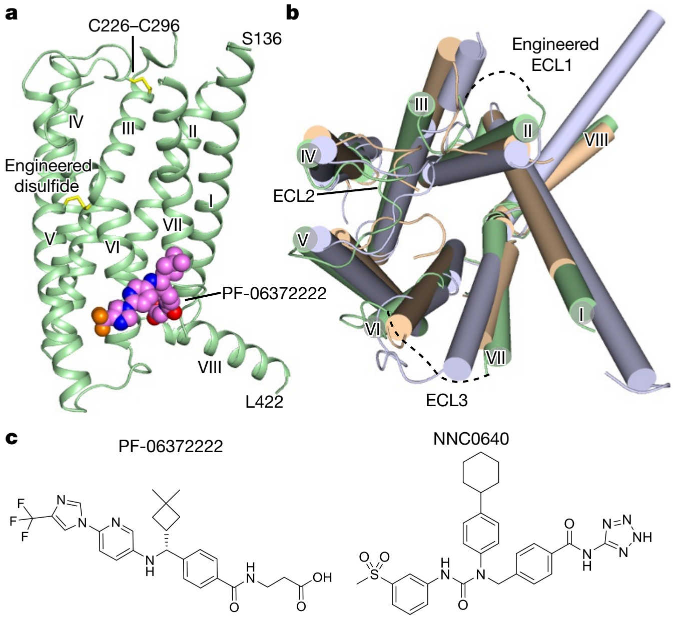
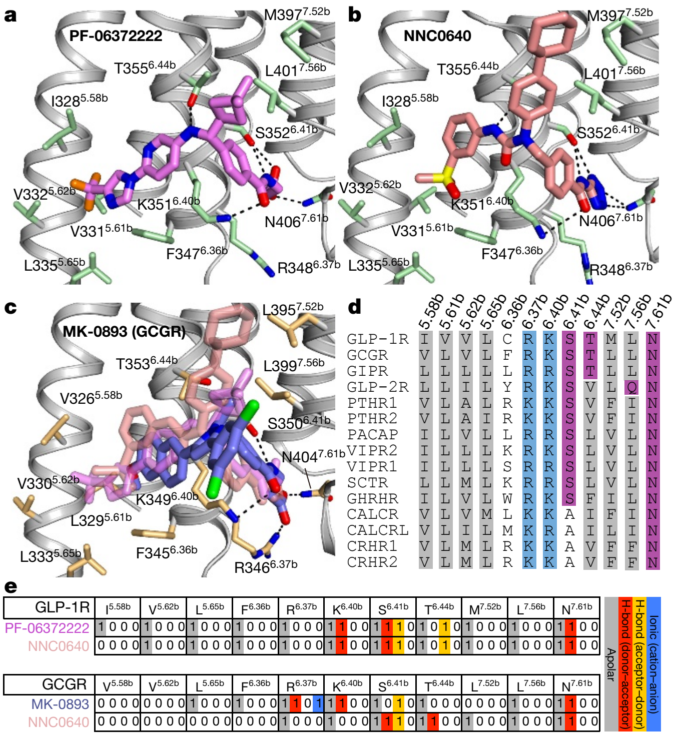


Figure 1 | Overall structure of the GLP-1R TMD. a, GLP-1R is represented as a green cartoon. Representative ligand PF-06372222 is shown as spheres with purple carbon atoms. Disulfide bonds are shown as yellow sticks and labelled. Helices are labelled I–VII. **b**, Top view of GLP-1R (green cylinders) compared with the initial GCGR structure (PDB code 4L6R, blue) and the thermostabilized GCGR structure (PDB code 5EE7, yellow). **c**, Chemical structures of co-crystallized ligands PF-06372222 and NNC0640.



	5.58b	5.61b	5.62b	5.65b	6.36b	6.37b	6.40b	6.41b	6.44b	7.52b	7.56b	7.61b
GLP-1R	I	V	V	L	C	R	K	S	T	M	L	N
GCGR	V	L	V	L	F	R	K	S	T	L	L	N
GIPR	L	L	L	L	L	R	R	S	T	L	L	N
GLP-2R	L	L	I	L	Y	R	K	S	V	L	Q	N
PTHR1	V	L	A	L	R	K	K	S	V	F	I	N
PTHR2	V	L	A	I	R	K	K	S	V	F	I	N
PACAP	I	L	V	L	L	R	R	S	L	V	L	N
VIPR2	I	L	L	L	K	R	K	S	L	V	L	N
VIPR1	I	L	L	L	S	R	R	S	L	V	L	N
SCTR	L	L	M	L	K	R	R	S	L	V	L	N
GHRHR	I	L	V	L	W	R	K	S	F	I	L	N
CALCR	V	L	V	M	L	K	K	A	I	F	I	N
CALCRL	V	L	I	L	M	K	R	A	I	L	I	N
CRHR1	V	L	M	L	R	K	K	A	V	F	F	N
CRHR2	V	L	M	L	R	K	K	A	V	F	F	N

GLP-1R	I$^{5.58b}$	V$^{5.62b}$	L$^{5.65b}$	F$^{6.36b}$	R$^{6.37b}$	K$^{6.40b}$	S$^{6.41b}$	T$^{6.44b}$	M$^{7.52b}$	L$^{7.56b}$	N$^{7.61b}$
PF-06372222	1 0 0 0	1 0 0 0	1 0 0 0	1 0 0 0	1 0 0 0	1 1 0 0	1 1 1 0	1 0 1 0	1 0 0 0	1 0 0 0	1 1 0 0
NNC0640	0 0 0 0	1 0 0 0	1 0 0 0	1 0 0 0	1 0 0 0	1 1 0 0	1 1 1 0	1 0 1 0	1 0 0 0	1 0 0 0	1 1 0 0

GCGR	V$^{5.58b}$	V$^{5.62b}$	L$^{5.65b}$	F$^{6.36b}$	R$^{6.37b}$	K$^{6.40b}$	S$^{6.41b}$	T$^{6.44b}$	L$^{7.52b}$	L$^{7.56b}$	N$^{7.61b}$
MK-0893	0 0 0 0	0 0 0 0	1 0 0 0	1 0 0 0	1 1 0 1	1 1 0 0	1 0 1 0	1 0 0 0	0 0 0 0	1 0 0 0	1 1 0 0
NNC0640	0 0 0 0	0 0 0 0	0 0 0 0	0 0 0 0	1 0 0 0	1 0 0 0	1 1 1 0	1 1 0 0	1 0 0 0	1 0 0 0	1 1 0 0

Figure 2 | NAM-binding pockets. a–c, Ligand-binding interfaces of GLP-1R–PF-06372222 (**a**), GLP-1R–NNC0640 (**b**) and GCGR–MK-0893 (**c**)[3]. Interfaces are shown in identical orientations after superposition of the entire domain. Receptors are shown in grey cartoon representation. Key side chains are shown as sticks and coloured green and yellow for GLP-1R and GCGR, respectively. Hydrogen bonds are shown as dashes. Carbon atoms of PF-06372222, NNC0640 and MK-0893 are coloured purple, salmon and blue, respectively. Others elements are coloured as follows: oxygen, red; nitrogen, dark blue; sulfur, yellow; chlorine, green; fluorine, orange. In **c**, the structures of superimposed PF-06372222 (purple) and NNC0640 (salmon) are shown for comparison (20% transparency). **d**, Alignment of key residues in the ligand-binding pocket among human class B receptors. The conserved hydrophobic, neutral hydrophilic and basic residues are coloured grey, blue and purple, respectively. **e**, The receptor–ligand interaction patterns are described by interaction fingerprint bit strings encoding different interaction types. ‘0’ denotes no interaction; ‘1’ indicates that an interaction is observed. Colour codes are listed on the right. The interaction data of GCGR–NNC0640 are from ref. 18.

contact surface with the membrane. A comparison of the surface areas that are either buried by the receptor or solvent/membrane accessible of all 86 crystallized GPCR ligands demonstrates that four NAMs (PF-06372222, NNC0640, MK-0893 and BPTU) target the extra-helical binding sites of GLP-1R, GCGR and the purinergic receptor $P2Y_1$ (ref. 19), while the allosteric agonist (TAK875) interacts with the orthosteric and extra-helical binding sites of the free fatty acid receptor 1 (FFAR1)[20]. Thus, their exposed/buried surface ratio is larger than that of the rest (Extended Data Fig. 3), and the affinity of those ligands with corresponding receptors is determined not only by interactions with buried receptor-binding pockets, but also by favourable hydrophobic interactions with the membrane[19].

Residues in the NAM binding pocket are mostly conserved within class B receptors (Fig. 2d). Three of the five residues (R$^{6.37b}$, K$^{6.40b}$ and N$^{7.61b}$) in the hydrophilic subpocket between helices VI and VII are highly conserved. S$^{6.41b}$ is conserved in most secretin-like receptors, whereas T$^{6.44b}$ is only present in GLP-1R, GCGR and GIPR (gastric inhibitory polypeptide receptor) and may be an important determinant for selectivity of small molecule allosteric modulators. The hydrophobic surface of helix V targeted by NAMs is composed of variable hydrophobic aliphatic residues among class B GPCRs. On the other side of the pocket, the residues M$^{7.52b}$ and L$^{7.56b}$ of helix VII are also conserved.

Mutagenesis studies align with the binding modes revealed in the GLP-1R crystal structures (Figs 2, 3 and Extended Data Fig. 4). S352$^{6.41b}$A abolishes the ability of all three NAMs to antagonize GLP-1-mediated cAMP accumulation. T355$^{6.44b}$A also abolishes the antagonistic potency of PF-06372222 or NNC0640, but its effect on the potency of MK-0893 is greatly reduced, consistent with the absence of a hydrogen bond in the GCGR–MK-0893 structure (Fig. 3a–c). I328$^{5.58b}$N and V332$^{5.62b}$N mutants diminish the potency of PF-06372222, but have no notable effect on the potency of NNC0640 and MK-0893, consistent with the larger hydrophobic interaction surface of these residues with the trifluoromethyl group of PF-06372222. The bulky V332$^{5.62b}$W and L335$^{5.65b}$W mutants decrease NAM potencies (PF-06372222 in particular), probably as a consequence of steric hindrance, whereas the I328$^{5.58b}$W mutant increases the potencies of all three NAMs (Fig. 3a–c), probably by increasing the hydrophobic interactions between tryptophan and these ligands. The enhanced potency of NAMs on the C347$^{6.36b}$F mutant of GLP-1R (Fig. 3a–c), and attenuated potency on the reverse F345$^{6.36b}$C mutant of GCGR (Extended Data Fig. 4), confirm the similar binding mode of these NAMs by GLP-1R and GCGR. Most NAMs reported for GCGR, such as MK-0893 (ref. 21), NNC0640 (ref. 22) and PF-06372222 (ref. 23), are selective for GCGR over GLP-1R and our results show that C/F$^{6.36b}$ is an important determinant of GLP-1R/GCGR selectivity, thus providing a means for rationally designing highly potent GLP-1R NAM tool compounds through enhancement of GLP-1R-specific interactions. These tool compounds would provide an invaluable resource for *in vitro* and *in vivo* pharmacological studies allowing receptor modulation without affecting the peptide ligand-binding pocket.

Comparative molecular dynamics simulations indicate that the C347$^{6.36b}$F mutant maximizes van der Waals interactions with all three NAMs by stabilizing the aliphatic side chain of K351$^{6.40b}$ in an optimal conformation for hydrophobic interactions with NAMs (Extended Data Fig. 5). The C347$^{6.36b}$F mutation alone does not affect cAMP signalling potency of GLP-1 in a wild-type background. By contrast, the engineered disulfide bond I317$^{5.47b}$C–G361$^{6.50b}$C resulted in a complete loss of signalling in response to GLP-1 both in the crystallization construct and in the wild-type receptor (Fig. 3d), confirming our hypothesis that

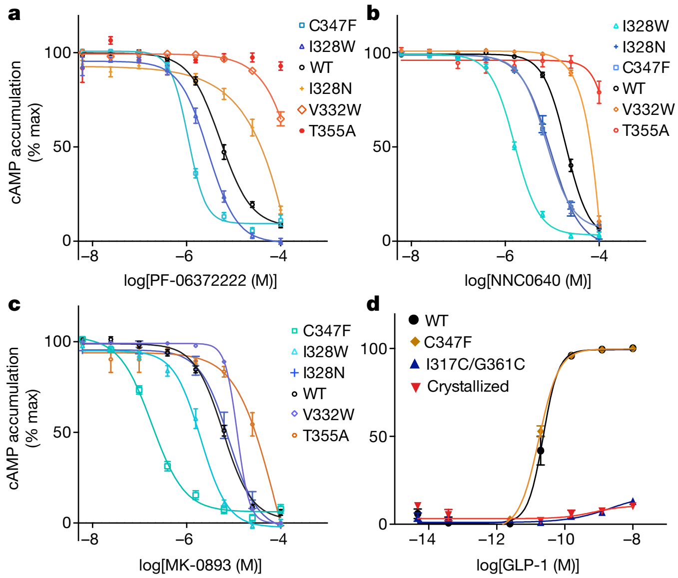


Figure 3 | Structural determinants of NAM potency. a–**c**, Representative effects of binding-pocket mutations on the antagonistic potency of PF-06372222 (**a**), NNC0640 (**b**) and MK-0893 (**c**) on wild-type (WT) GLP-1R by the cAMP accumulation assay. Wild-type GLP-1R (black) and GLP-1R mutations that show >4-fold increase (cyan), <4-fold effect (blue), 4–10-fold decrease (orange), or >10-fold decrease (red) on the potency of each NAM are indicated in the curves, and the order of codes from top to bottom is based on potency (higher to lower). **d**, Comparison of representative constructs to the wild-type GLP-1R. All constructs are with full-length sequence for signalling of GLP-1. The single (C347F) and double (I317C/G361C) mutations are based on the wild-type sequence, and the crystallized construct includes 10 themostabilizing mutations compared to the wild-type sequence (see Extended Data Fig. 1). In **a**–**c**, expressed cells were pre-activated by a constant amount of GLP-1 and then competed with serial dilutions of NAMs to examine cAMP accumulation; in **d**, expressed cells were activated by a series of concentrations of GLP-1 and then cAMP accumulation was measured. Experiments were repeated at least three times and error bars represent s.e.m. of quadruplicate measurements.

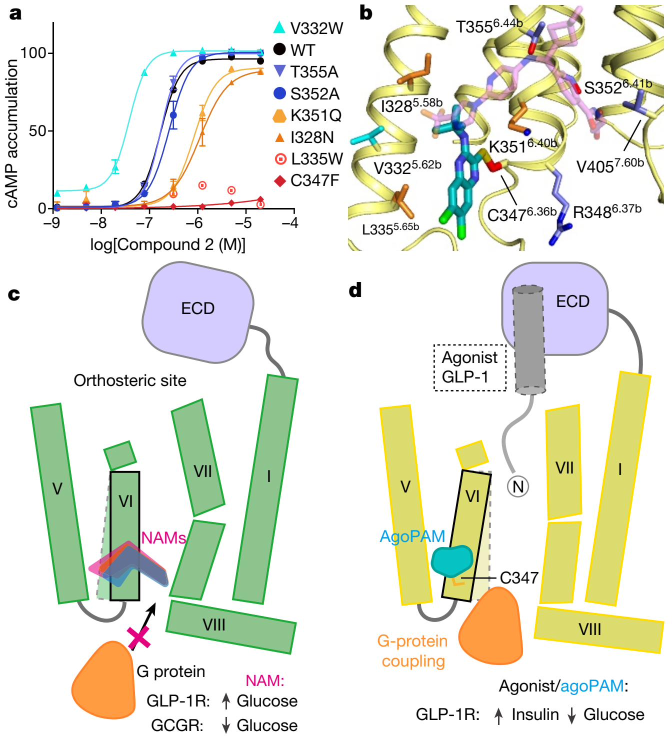


Figure 4 | Modelling of the agonist PAM compound 2 binding mode and proposed mechanism of activation. a, Representative mutation effects on the agonism of compound 2. **b**, Proposed binding mode of compound 2 based on modelling, simulation and mutation studies. Dose–response curves of GLP-1R mutants that show >4-fold increase (cyan), <4-fold effect (blue), 4–10-fold decrease (orange), or >10-fold decrease (red) in compound 2 potency compared with wild-type GLP-1R (black) are indicated in curves (**a**, and Extended Data Fig. 8) and mapped on a representative molecular dynamics simulation snapshot (at 500 ns) of GLP-1R bound to compound 2 (**b**). Experiments were repeated at least 3 times and error bars represent s.e.m. of quadruplicate measurements. In **b**, PF-06372222 crystal structure is also shown for comparison (pink, transparent). **c**, **d**, Schematic diagram of the inhibition of GLP-1R or GCGR by NAMs (**c**) and activation of GLP-1R by agonist PAM (agoPAM; **d**). NAMs (PF-06372222, NNC0640 and MK-0893 are coloured purple, salmon and blue, respectively) bind and insert into the cleft between helices VI and VII and push the equilibrium of helix VI towards inactive conformation (solid line); while agonist PAM (compound 2) binds between helices V and VI, pulling the covalently linked C347$^{6.36b}$ towards helix V (solid line), thus providing an intercellular binding site for G protein (active conformation). Green (**c**) and yellow (**d**) schemes indicate inactive and active conformations, respectively. In **d**, the endogenous agonist GLP-1 binding model is also shown for comparison and both agonist and agonist PAM can independently trigger downstream signalling. Signalling consequences are indicated as directional arrows in **c** and **d**. ECD, extracellular domain.

the disulfide link between C317$^{5.47b}$ and C361$^{6.50b}$ locks the protein into an inactive conformation. While the binding of a non-TMD-binding peptide exendin-4$_{(9-39)}$ is only slightly affected, the binding of GLP-1 to the disulfide-including construct is abolished (Extended Data Fig. 6), indicating that the locked inactive conformation disrupted the binding of GLP-1 to the orthosteric pocket, and this may be reflected by an unusually tilted helix VI towards helix V near the disulfide bond that links the middle region of helices V and VI. Comparative molecular dynamics simulations support this conclusion, indicating that the engineered disulfide bond restricts the movement of helix V towards helix III, and stabilizes the hydrogen-bond interaction network associated with the ionic lock (H180$^{2.50b}$–E247$^{3.50b}$, R348$^{6.37b}$–E408)[24] that is proposed to stabilize class B GPCRs in an inactive conformation[14,15,24] (Extended Data Fig. 7).

The GLP-1R-specific C347$^{6.36b}$ residue has been identified as the site for covalent interaction with electrophilic groups of the agonist PAMs 6,7-dichloro-3-methanesulfonyl-2-*tert*-butylamino-quinoxaline (compound 2) and BTEP[25,26]. Modelling of the GLP-1R TMD structure with agonist PAM compound 2 was experimentally validated by mutagenesis studies (Fig. 4a, b). Compound 2 is proposed to be located orthogonally above helix VI forming interactions with residues at the interface of helices V and VI (Fig. 4b), supported by its unaffected potency by mutations (R348$^{6.37b}$Q, S352$^{6.41b}$A, T355$^{6.44b}$A and V405$^{7.60b}$L) located at the cleft between helices VI and VII (Extended Data Fig. 8). The C347$^{6.36b}$F mutant abolished the potency of compound 2, in line with the covalent binding model and consistent with structure–activity relationship (SAR) studies showing the importance of the electrophilic character of the dichloroquinoxaline moiety of compound 2 (refs 25, 26). Decreased potency for the I328$^{5.58b}$N and K351$^{6.40b}$Q mutants can be explained by the decreased hydrophobic interaction surface with compound 2. The enhanced potency of the V332$^{5.62b}$W mutant is consistent with our expectation that it could increase the hydrophobic interaction surface with the key tert-butyl moiety[7] of compound 2, whereas the diminished potency of the bulky L335$^{5.65b}$W mutant is likely to be the result of its steric incompatibility with the dichloroquinoxaline group of compound 2 (Fig. 4). The apparent inverse effects of mutations I328$^{5.58b}$W and V332$^{5.62b}$W by NAMs and compound 2 are consistent with the different orientations of the tert-butyl group of compound 2 and trifluoromethyl group of PF-06372222 directed towards I328$^{5.58b}$ and V332$^{5.62b}$, respectively (Figs 3, 4). Although the current binding model of compound 2 may represent only a snapshot of the pre-active state and the interface between compound 2 and GLP-1R may be reshaped upon full activation, the comparison of the binding sites and specific mutation effects on NAMs and agonist PAM can provide a template for the structure-based design of novel allosteric modulators and SAR investigations in the context of GLP-1R.

On the basis of comparative molecular dynamics simulations of NAM-bound and agonist PAM-bound GLP-1R, we propose an agonist PAM-mediated receptor activation model (Fig. 4c, d). In the NAM-bound GLP-1R, the ionic lock interactions between helices II, III, VI and VII (Extended Data Fig. 7), together with the NAM itself, restrict the movement of helix VI away from helix VII that is implicated to be required for G-protein coupling[3,24]. In our molecular dynamics simulations, the agonist PAM interacts with GLP-1R to induce a conformational change in the intracellular regions of helices V and VI that results in disruption of the intracellular ionic lock (Extended Data Fig. 7). These conformational rearrangements open a cavity at a similar location on the intracellular portion of the receptor that has been associated with G-protein coupling in class A GPCRs[4–6]. The full activation of class B GPCRs is likely to involve larger movement of helices VI and VII to facilitate efficient G-protein binding, as seen in class A GPCRs[4–6], and consistent with substituted cysteine accessibility method (SCAM) studies of a class B GPCR[27]. Despite the limitations of using an inactive NAM-bound GLP-1R template to model an agonist-bound receptor conformation, the molecular dynamics simulation of the agonist PAM-bound GLP-1R structure revealed that the cationic residue $R176^{2.46b}$ may have a similar role in G-protein coupling as $R^{3.50}$ in class A GPCRs[28]. This finding is consistent with mutation studies showing decreased GLP-1 potency of the $R176^{2.46b}$A mutant for rat GLP-1R[29], and the $R176^{2.46b}$Q mutant for human GLP-1R without a loss of GLP-1 binding affinity (Extended Data Fig. 9), as well as a decrease in gastric inhibitory polypeptide (GIP) potency for human GIPR[30].

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 9 January; accepted 20 April 2017.

1. de Graaf, C. *et al.* Glucagon-like peptide-1 and its class B G protein-coupled receptors: a long march to therapeutic successes. *Pharmacol. Rev.* **68,** 954–1013 (2016).
2. Cho, Y. M., Merchant, C. E. & Kieffer, T. J. Targeting the glucagon receptor family for diabetes and obesity therapy. *Pharmacol. Ther.* **135,** 247–278 (2012).
3. Jazayeri, A. *et al.* Extra-helical binding site of a glucagon receptor antagonist. *Nature* **533,** 274–277 (2016).
4. Rasmussen, S. G. *et al.* Crystal structure of the β2 adrenergic receptor–G_s protein complex. *Nature* **477,** 549–555 (2011).
5. Carpenter, B., Nehmé, R., Warne, T., Leslie, A. G. & Tate, C. G. Structure of the adenosine A_{2A} receptor bound to an engineered G protein. *Nature* **536,** 104–107 (2016).
6. Scheerer, P. *et al.* Crystal structure of opsin in its G-protein-interacting conformation. *Nature* **455,** 497–502 (2008).
7. Knudsen, L. B. *et al.* Small-molecule agonists for the glucagon-like peptide 1 receptor. *Proc. Natl Acad. Sci. USA* **104,** 937–942 (2007).
8. Sloop, K. W. *et al.* Novel small molecule glucagon-like peptide-1 receptor agonist stimulates insulin secretion in rodents and from human islets. *Diabetes* **59,** 3099–3107 (2010).
9. Chen, D. *et al.* A nonpeptidic agonist of glucagon-like peptide 1 receptors with efficacy in diabetic db/db mice. *Proc. Natl Acad. Sci. USA* **104,** 943–948 (2007).
10. Yang, D. *et al.* Structural determinants of binding the seven-transmembrane domain of the glucagon-like peptide-1 receptor (GLP-1R). *J. Biol. Chem.* **291,** 12991–13004 (2016).
11. Wootten, D. *et al.* The extracellular surface of the GLP-1 receptor is a molecular trigger for biased agonism. *Cell* **165,** 1632–1643 (2016).
12. Underwood, C. R. *et al.* Crystal structure of glucagon-like peptide-1 in complex with the extracellular domain of the glucagon-like peptide-1 receptor. *J. Biol. Chem.* **285,** 723–730 (2010).
13. Patterson, J. T., Li, P., Day, J. W., Gelfanov, V. M. & Dimarchi, R. D. A hydrophobic site on the GLP-1 receptor extracellular domain orients the peptide ligand for signal transduction. *Mol. Metab.* **2,** 86–91 (2013).
14. Hollenstein, K. *et al.* Structure of class B GPCR corticotropin-releasing factor receptor 1. *Nature* **499,** 438–443 (2013).
15. Siu, F. Y. *et al.* Structure of the human glucagon class B G-protein-coupled receptor. *Nature* **499,** 444–449 (2013).
16. Wootten, D., Simms, J., Miller, L. J., Christopoulos, A. & Sexton, P. M. Polar transmembrane interactions drive formation of ligand-specific and signal pathway-biased family B G protein-coupled receptor conformations. *Proc. Natl Acad. Sci. USA* **110,** 5211–5216 (2013).
17. Mann, R. J., Al-Sabah, S., de Maturana, R. L., Sinfield, J. K. & Donnelly, D. Functional coupling of Cys-226 and Cys-296 in the glucagon-like peptide-1 (GLP-1) receptor indicates a disulfide bond that is close to the activation pocket. *Peptides* **31,** 2289–2293 (2010).
18. Zhang, H. *et al.* Structure of the full-length glucagon class B G protein-coupled receptor. *Nature* http://dx.doi.org/10.1038/nature22363 (this issue).
19. Zhang, D. *et al.* Two disparate ligand-binding sites in the human P2Y1 receptor. *Nature* **520,** 317–321 (2015).
20. Srivastava, A. *et al.* High-resolution structure of the human GPR40 receptor bound to allosteric agonist TAK-875. *Nature* **513,** 124–127 (2014).
21. Xiong, Y. *et al.* Discovery of a novel glucagon receptor antagonist *N*-[(4-(1S)-1-[3-(3, 5-dichlorophenyl)-5-(6-methoxynaphthalen-2-yl)-1*H*-pyrazol-1-yl] ethylphenyl)carbonyl]-β-alanine (MK-0893) for the treatment of type II diabetes. *J. Med. Chem.* **55,** 6137–6148 (2012).
22. Lau, J. *et al.* New beta-alanine derivatives are orally available glucagon receptor antagonists. *J. Med. Chem.* **50,** 113–128 (2007).
23. Guzman-Perez, A. *et al.* The design and synthesis of a potent glucagon receptor antagonist with favorable physicochemical and pharmacokinetic properties as a candidate for the treatment of type 2 diabetes mellitus. *Bioorg. Med. Chem. Lett.* **23,** 3051–3058 (2013).
24. Hollenstein, K. *et al.* Insights into the structure of class B GPCRs. *Trends Pharmacol. Sci.* **35,** 12–22 (2014).
25. Bueno, A. B. *et al.* Positive allosteric modulation of the glucagon-like peptide-1 receptor by diverse electrophiles. *J. Biol. Chem.* **291,** 10700–10715 (2016).
26. Nolte, W. M. *et al.* A potentiator of orthosteric ligand activity at GLP-1R acts via covalent modification. *Nat. Chem. Biol.* **10,** 629–631 (2014).
27. Spyridaki, K. *et al.* Structural–functional analysis of the third transmembrane domain of the corticotropin-releasing factor type 1 receptor: role in activation and allosteric antagonism. *J. Biol. Chem.* **289,** 18966–18977 (2014).
28. Venkatakrishnan, A. J. *et al.* Molecular signatures of G-protein-coupled receptors. *Nature* **494,** 185–194 (2013).
29. Mathi, S. K., Chan, Y., Li, X. & Wheeler, M. B. Scanning of the glucagon-like peptide-1 receptor localizes G protein-activating determinants primarily to the N terminus of the third intracellular loop. *Mol. Endocrinol.* **11,** 424–432 (1997).
30. Cordomí, A. *et al.* Functional elements of the gastric inhibitory polypeptide receptor: Comparison between secretin- and rhodopsin-like G protein-coupled receptors. *Biochem. Pharmacol.* **96,** 237–246 (2015).

Acknowledgements This work was supported by the National Natural Science Foundation of China grants 31330019 (Z.-J.L.), 31500593 (G.S.), 81373463 and 81573479 (D.Y.), the National Health and Family Planning Commission grants 2012ZX09304-011, 2013ZX09401003-005, 2013ZX09507001 and 2013ZX09507-002 (M.-W.W.), Shanghai Science and Technology Development Fund 15DZ2291600 (M.-W.W.), and Ministry of Science and Technology of China grants 2014CB910400 (Z.-J.L.) and 2015CB910104 (Z.-J.L.), the Netherlands eScience Center (NLeSC)/NWO Enabling Technologies project: 3D-e-Chem, grant 027.014.201 (C.d.G.), the European Cooperation in Science and Technology Action CM1207 GLISTEN (C.d.G.), and National Key Research and Development Program of China 2016YCF0905902 (S.Z.). We thank the Cloning, Cell Expression and Protein Purification Core Facilities of iHuman Institute for their support. We thank L. Qu, Y. Feng and C. Ji for their technical assistance, Y. Liu, H. Tao, S. Qin, W. Shui, F. Ni, C. Zhang, J. Cheng, Q. Zhao and V. Cherezov for discussions or contributions at the early stages of this project, and A. Walker and S. Reedtz-Runge for critical review of the manuscript. We thank the Shanghai Municipal Government, ShanghaiTech University and GPCR Consortium for financial support. The synchrotron radiation experiments were performed at the BL41XU of Spring-8 with approval of the Japan Synchrotron Radiation Research Institute (JASRI) (proposal no. 2016B2708 and 2016B2724), and beamline BL17U1 (Shanghai Synchrotron Radiation Facility [SSRF], China). Computational resources were supported by the Shanghai Supercomputer Center, and the Supercomputing Center of ShanghaiTech University.

Author Contributions G.S., C.d.G. and M.A.H. designed constructs for crystallization. G.S., Y.W., S.J. and K.L. expressed, characterized and screened constructs and ligands for crystallization. G.S., Y.W., S.J. and F.W. purified and crystallized the receptor, optimized crystallization conditions and grew the crystals. G.S., Z.-J.L., M.A.H. and G.W.H. collected diffraction data and solved and refined the structure. G.S., D.Y. and C.d.G. designed and analysed the receptor mutagenesis studies. D.Y., X.C., A.D. and G.L. expressed the receptor, and performed the mutagenesis, functional and ligand-binding assays. Q.Z., Y.W., C.d.G. and S.Z. constructed the GLP-1R–agonist PAM model and performed and analysed MD-simulations on wild-type and mutant GLP-1R. B.W. provided advice on construct design. R.C.S. conceived of the project. R.C.S., M.-W.W., and Z.-J.L. were responsible for the overall project management and edited the manuscript. G.S., C.d.G. and M.A.H. wrote the manuscript with discussions and improvements from D.L., L.Y. and J.L.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details are available in the online version of the paper. Readers are welcome to comment on the online version of the paper. Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations. Correspondence and requests for materials should be addressed to Z.-J.L. (liuzhj@shanghaitech.edu.cn), M.-W.W. (mwwang@simm.ac.cn) or R.C.S. (stevens@shanghaitech.edu.cn).

Reviewer Information *Nature* thanks G. Lebon, T. W. Schwartz and C. Siebold for their contribution to the peer review of this work.

METHODS

No statistical methods were used to predetermine sample size. The experiments were not randomized and investigators were not blinded to allocation during experiments and outcome assessment.

Purification of GLP-1R-T4L protein and crystallization in lipidic cubic phase. The GLP-1R-T4L construct (Extended Data Fig. 1) was produced with residues 128–431, containing 10 mutations and with T4 lysozyme (T4L) replacing three residues (258–260) at intracellular loop 2 (ICL2). Nine residues from ECL1 (205–213) were further replaced by a GSG linker to facilitate growth of high-quality crystals. This construct was expressed with an N-terminal 10× histidine tag and thermostabilized *Escherichia coli* apocytochrome b_{562}RIL (BRIL) to increase the expression yield that was removed during purification with a TEV site inserted between the BRIL and GPCR. The 10 mutations (see Extended Data Fig. 1) in the crystallization construct are screened from a total of 40 single point mutants and 29 pairs of double-cysteine mutants (designed for disulfide bond formation), based mainly on analytical size exclusion chromatography and thermostability assay[31] results. Our initial pool of single point and double-cysteine mutants are based on structural modelling combined with our experience from structure determination of related receptors or other GPCRs. For thermostability assay, CPM (*N*-([4-(7-diethylamino-4-methyl-3-coumarinyl)phenyl]maleimide) dye was dissolved in DMSO at 4 mg ml^{-1} as stock solution and diluted 1:20 in buffer (25 mM HEPES, pH 7.5, 500 mM NaCl, 5% (v/v) glycerol, 0.01% (w/v) DDM, 0.002% (w/v) CHS) before use. 1 μl of diluted CPM was added to the same buffer with approximately 0.5–2 μg receptor in a final volume of 50 μl. The thermal denaturation assay was performed in a Rotor-gene real-time PCR cycler (Qiagen). The excitation wavelength was 365 nm and the emission wavelength was 460 nm. All assays were performed over a temperature range from 25 °C to 85 °C. The stability data were processed with GraphPad Prism. Curves of representative mutations and crystallized ligands are shown in Extended Data Fig. 1.

The fusion protein was expressed as previously described[32]. In brief, the Bac-to-Bac Baculovirus System (Invitrogen) in *Spodoptera frugiperda* (*Sf*9) cells was used for expression and cells were infected at a density of 2×10^6–3×10^6 cells per ml with baculovirus at a multiplicity of infection (MOI) of 5, and cultures were grown at 27 °C and collected at 48 h after infection. Cell membrane was washed twice using low salt buffer (10 mM HEPES, pH 7.5, 20 mM KCl, 10 mM MgCl_2 and EDTA-free protease inhibitor cocktail tablets), followed by three washes with high-salt buffer (10 mM HEPES, pH 7.5, 1 M NaCl, 20 mM KCl, 10 mM MgCl_2). Before solubilization, purified membranes were incubated with 200 μM of ligand PF-06372222 ((*R*)-3-(4-((3,3-dimethylcyclobutyl)((6-(4-(trifluoromethyl)-1*H*-imidazol-1-yl)pyridin-3-yl)amino)methyl)benzamido)propanoic acid) or NNC0640 (4-[1-(4-cyclohexylphenyl)-3-(3-methanesulfonylphenyl)ureidomethyl]-*N*-(2*H*-tetrazo-5-yl)benzamide) in the presence of 2 mg ml^{-1} iodoacetamide and EDTA-free protease inhibitor cocktail (Roche) for 45 min. GLP-1R-T4L was extracted from the membrane by adding *n*-dodecyl-β-D-maltopyranoside (DDM, Affymetrix) and cholesteryl hemisuccinate (CHS, Sigma) to the membrane solution to a final concentration of 1.0% (w/v) and 0.2% (w/v), respectively, and stirring was continued at 4 °C for 2.5 h. The supernatant was isolated by centrifugation at 160,000*g* for 30 min, followed by incubation in TALON IMAC resin (Clontech) at 4 °C, overnight. The resin was washed with 20 column volumes of wash buffer 1 (50 mM HEPES, pH 7.5, 500 mM NaCl, 5% (v/v) glycerol, 0.05% (w/v) DDM, 0.01% (w/v) CHS, 30 mM imidazole and 20 μM ligand) and then followed by 10 column volumes of wash buffer 2 (25 mM HEPES, pH 7.5, 500 mM NaCl, 5% (v/v) glycerol, 0.01% (w/v) DDM, 0.002% (w/v) CHS, 30 mM imidazole and 100 μM ligand). The resin was resuspended with 2 column volumes of wash buffer 2, TEV protease was added with a molar ratio of 1:15, and the mixture was shaken at 4 °C overnight. Receptor protein was harvested the second day from the flow-through of the resin. The protein was concentrated to ~30 mg ml^{-1} with a 100 kDa molecular mass cut-off concentrator (Millipore).

Protein sample was reconstituted into lipidic cubic phase by mixing 40% of ~30 mg ml^{-1} protein with 60% lipid (10% (w/w) cholesterol, 90% (w/w) monoolein). Crystallization trials were performed using a syringe lipid mixer and the protein-lipid mixture was dispensed in 40 nl drops onto glass sandwich plates and overlaid with 800 nl precipitant solution using a NT8 (Formulatrix). For the GLP-1R–PF-06372222 complex, crystals appeared after 1 week in 0.40–0.45 M ammonium acetate, 0.1 M sodium cacodylate, pH 6.2–6.6, 35–38% PEG400, 3% (w/v) aminohexanoic acid and reached their full size ($150 \times 50 \times 10\,\mu\text{m}^3$) within 2–3 weeks (Extended Data Fig. 1). For the GLP-1R–NNC0640 complex, crystals were grown slowly from precipitant conditions containing 0.40–0.45 M ammonium acetate, 0.1 M sodium acetate, pH 5.0–5.8, 38–40% PEG400, and reached their full size ($50 \times 30 \times 5\,\mu\text{m}^3$) after 50–60 days. Crystals were harvested directly from lipidic cubic phase using 50–150 μm micromounts (M2-L19-50/150, MiTeGen), flash frozen and stored in liquid nitrogen. Initial diffractions were tested at Shanghai Synchrotron Radiation Facility (SSRF), China.

Data collection, structure solution and refinement. X-ray diffraction data were collected at the SPring-8 beam line 41XU, Hyogo, Japan, using a Rayonix MX225HE detector (X-ray wavelength 1.0000 Å). The crystals were exposed with a 10 μm minibeam for 0.2 s and 0.2° oscillation per frame, and a rastering system was applied to find the best diffracting parts of single crystals[33]. XDS[34] was used for integrating and scaling data from the 25 best-diffracting crystals for GLP-1R–PF-06372222 complexes and 27 crystals for GLP-1R–NNC0640 complexes. The GLP-1R–PF-06372222 complex was solved by molecular replacement with Phaser[35] using thermostabilized GCGR (PDB code 5EE7) as search model, and T4L (PDB code 212L) was manually docked into the density after getting the initial solution. The structure was refined iteratively with Phenix[36] and Refmac5[37] with manual examination into $|2F_o| - |F_c|$ and $|F_o| - |F_c|$ maps with Coot[38]. Final refinement was performed with Buster[39] where individual positions and TLS refinements were used along with NCS restraints. The GLP-1R–NNC0640 structure was solved using GLP-1R–PF-06372222 as starting model and refined under the same procedure. Both structures include two molecules per asymmetric unit with identical packing (Extended Data Fig. 1e). Of the three ECLs, the most conserved ECL2 was well resolved with only a few side chains missing, whereas residues 204–215 of ECL1 and 372–379 of ECL3 are disordered. On the intercellular side, residues at ICL1 and ICL3 were well resolved, and ICL2 was replaced by T4L for crystallization. Structures have been carefully refined and checked by MolProbity[40], and statistics are provided in Extended Data Table 1.

cAMP assay. The desired mutations were introduced into codon optimized amino-terminally Flag tag-labelled human GLP-1R in the pcDNA3.1/V5-His-TOPO vector (Invitrogen); this construct displayed equivalent pharmacological features to that of untagged human GLP-1R based on radioligand-binding and cAMP assays. The mutants were constructed by PCR-based site-directed mutagenesis. Sequences of receptor clones were confirmed by DNA sequencing.

HEK-293T cells (obtained from and certified by the Cell Bank at the Chinese Academy of Science and confirmed as negative for mycoplasma contamination) were cultured in DMEM supplemented with 10% (v/v) fetal bovine serum, 50 IU ml^{-1} penicillin and 50 μg ml^{-1} streptomycin. Cells were maintained at 37 °C in 5% CO_2 incubator and seeded onto 6-well cell culture plates before transfection. After overnight culture, the cells were transiently transfected with wild-type or mutant GLP-1R DNA using Lipofectamine 2000 transfection reagent (Invitrogen). After 24 h, the transfected cells were seeded onto 384-well plates (8,000 cells per well). cAMP accumulation was measured using the LANCE cAMP kit (PerkinElmer) according to the manufacturer's instructions. In brief, transfected cells were incubated for 30 min in assay buffer (DMEM, 1 mM 3-isobutyl-1-methylxanthine) with different concentrations of GLP-1 or compounds (NAMs (6.1 nM to 100 μM plus constant concentration of GLP-1 at EC_{80} potency; compound 2 (1.2 nM to 20 μM); GLP-1 (0.0048 pM to 10 nM)) at 37 °C. The reactions were stopped by addition of lysis buffer containing LANCE reagents. Plates were then incubated for 60 min at room temperature and time-resolved FRET signals were measured at 620 nm and 650 nm by an EnVision multilabel plate reader (PerkinElmer).

Whole-cell binding assay. CHO-K1 cells (obtained from ATCC and confirmed as negative for mycoplasma contamination) were cultured in F-12 medium with 10% fetal bovine serum and collected 24 h after transfection, washed twice, and incubated with blocking buffer (F-12 supplemented with 33 mM HEPES and 0.1% bovine serum albumin (BSA), pH 7.4) for 2 h at 37 °C. For homogeneous binding, the cells were incubated in binding buffer (PBS with 10% BSA, pH 7.4) with constant concentration of $[^{125}\text{I}]$GLP-1 (40 pM) or $[^{125}\text{I}]$exendin-$4_{(9\text{-}39)}$ (40 pM) and different concentrations of unlabelled GLP-1 (3.57 pM to 1 μM) or exendin-$4_{(9\text{-}39)}$ (3.57 pM to 1 μM) at room temperature for 3 h. Cells were washed three times with ice-cold PBS and lysed by 50 μl lysis buffer (PBS supplemented with 20 mM Tris-HCl, 1% Triton X-100, pH 7.4). The plates were subsequently counted for radioactivity (counts per min) in a scintillation counter (MicroBeta2 Plate Counter, PerkinElmer) using a scintillation cocktail (OptiPhase SuperMix, PerkinElmer).

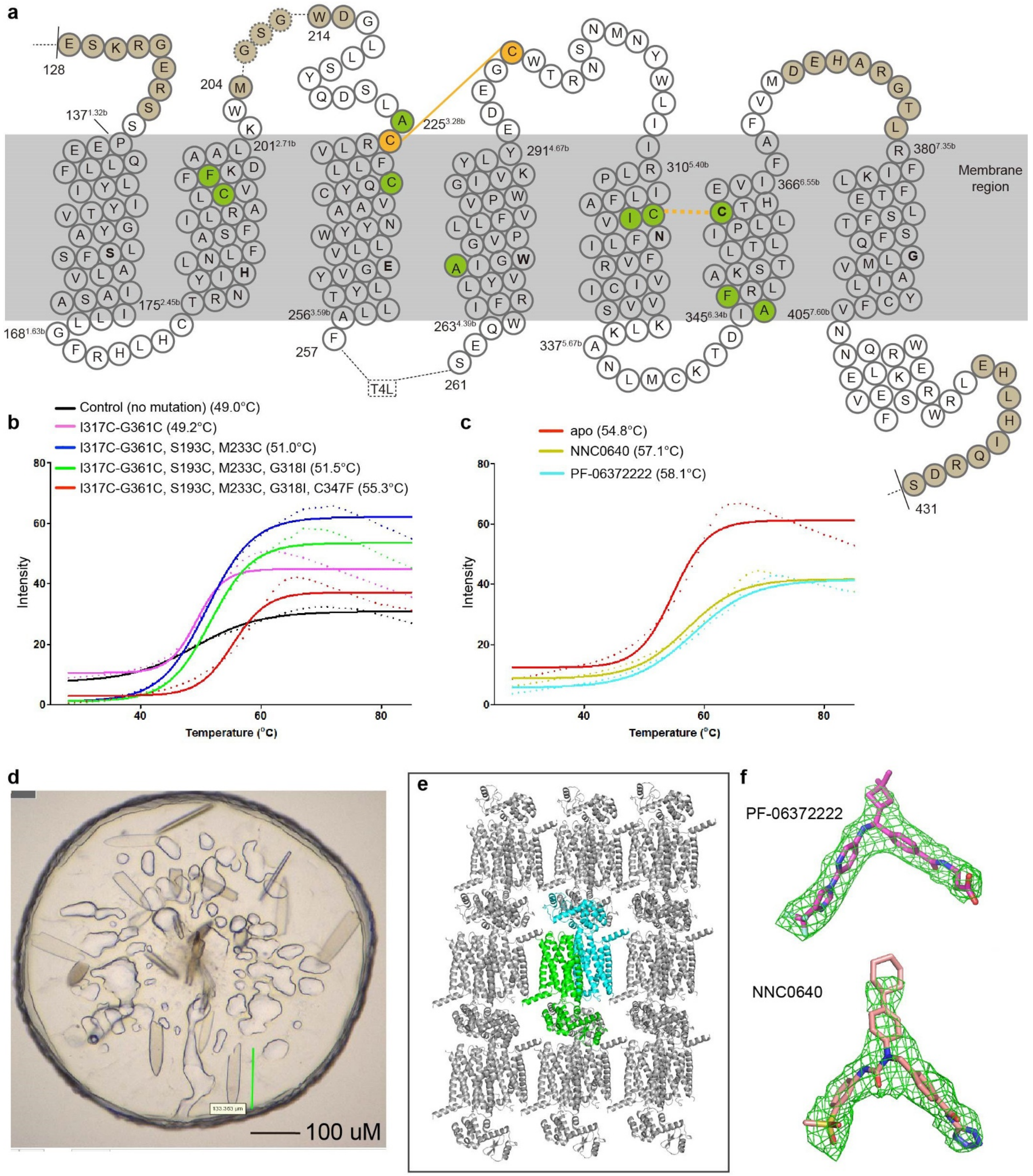
Protein-ligand docking. The crystal structure of GLP-1R was prepared using the Schrödinger Protein Preparation Wizard[41]. The CovDock application[42] implemented in the Schrödinger Suite 2016-1 was used for covalent docking of compound 2. Three-dimensional (3D) structure of compound 2 was prepared using LigPrep. The initial conformation of wild-type GLP-1R was obtained by gradually mutating back from the crystal structure through multiple rounds of single residue mutation and energy minimization. C347$^{6.36b}$ in the receptor was defined as the covalently bound reactive residue. The covalent reaction type nucleophilic substitution method was chosen and residues that were within 8 Å of the ligand or reactive residue were included in the minimization. Twenty poses were obtained and clustered to generate representative poses, which were further used as the initial structure for molecular dynamics simulations.

Molecular dynamic simulation. The simulated GLP1R–ligand complexes were embedded in a 120 Å × 120 Å POPC (1-palmytoil-2-oleoyl-*sn*-glycero-

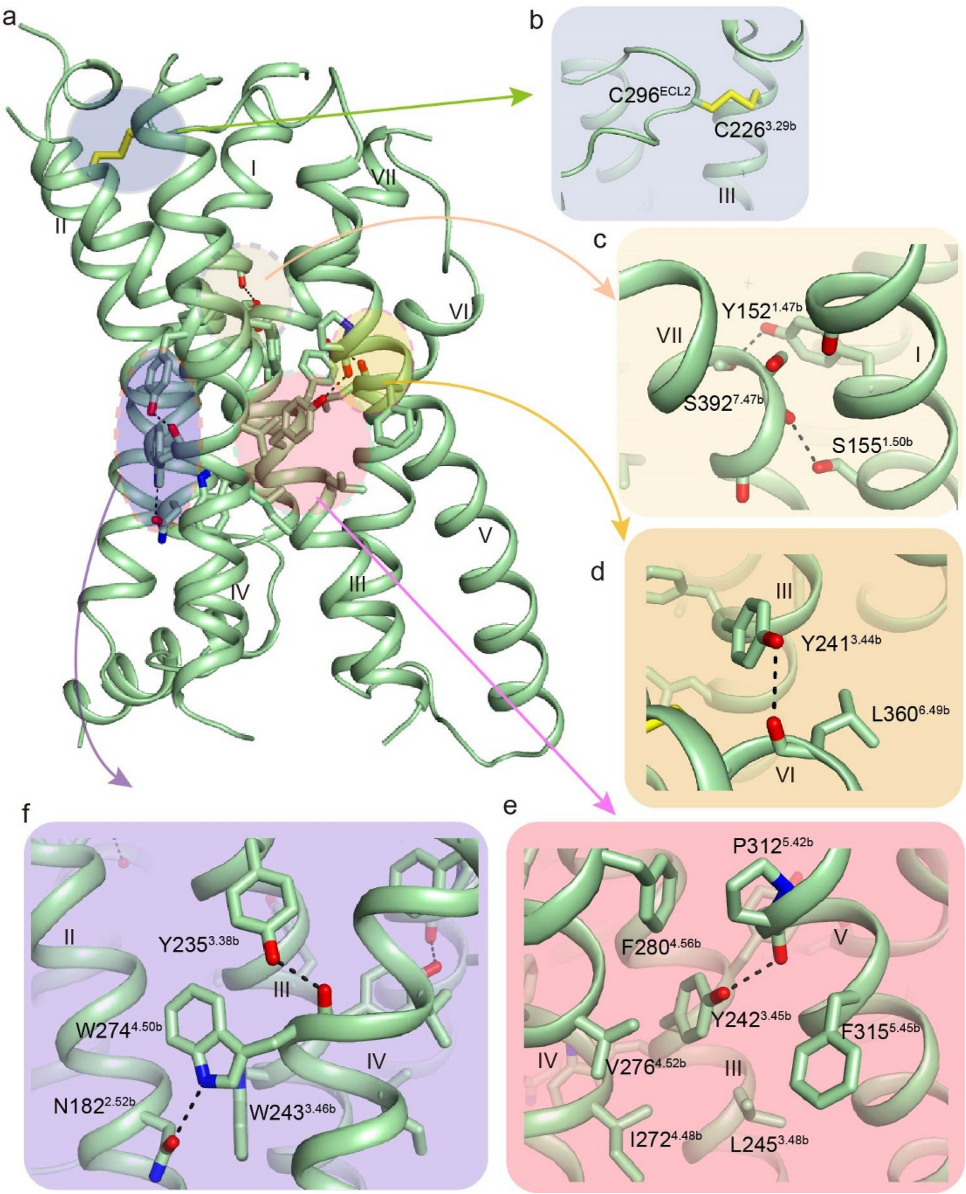
3-phosphatidylcholine) lipid bilayer and solvated by TIP3P waters with 0.15 M NaCl. The CHARMM36-CAMP force filed[43] was adopted for protein, lipid, water molecules and ions. All ligands (PF-06372222, NNC0640, MK-0893 and compound 2) were first optimized by the GAUSSIAN09 program at the B3LYP/6-31G* level, then modelled with the CHARMM CGenFF small-molecule force field, program version 1.0.0 (ref. 44). In each system, lipids located within 1 Å of the complexes were removed. Molecular dynamics simulations were performed using Gromacs 5.1.2 (ref. 45). All bonds involving hydrogen atoms were constrained using LINCS algorithm[46]. The particle mesh Ewald (PME) method was used to treat long-range electrostatic interactions with a cutoff of 14 Å. The entire system was first relaxed using the steepest descent energy minimization, followed by equilibration steps of 50 ns in total to equilibrate the lipid bilayer and the solvent while the restraints to the main chain of the protein and the ligand were reduced gradually to zero. Finally, the system was run without restraints, with a time step of 2 fs in the NPT ensemble at 300 K and 1 bar using a v-rescale thermostat[47] and Berendsen barostat[48], respectively. For each system, three 500 ns production runs were performed. The surface area was calculated by the program freeSASA[49] with default parameters.

Data availability. Atomic coordinates and structure factors have been deposited in the Protein Data Bank with accession codes 5VEW (GLP-1R–PF-06372222) and 5VEX (GLP-1R–NNC0640).

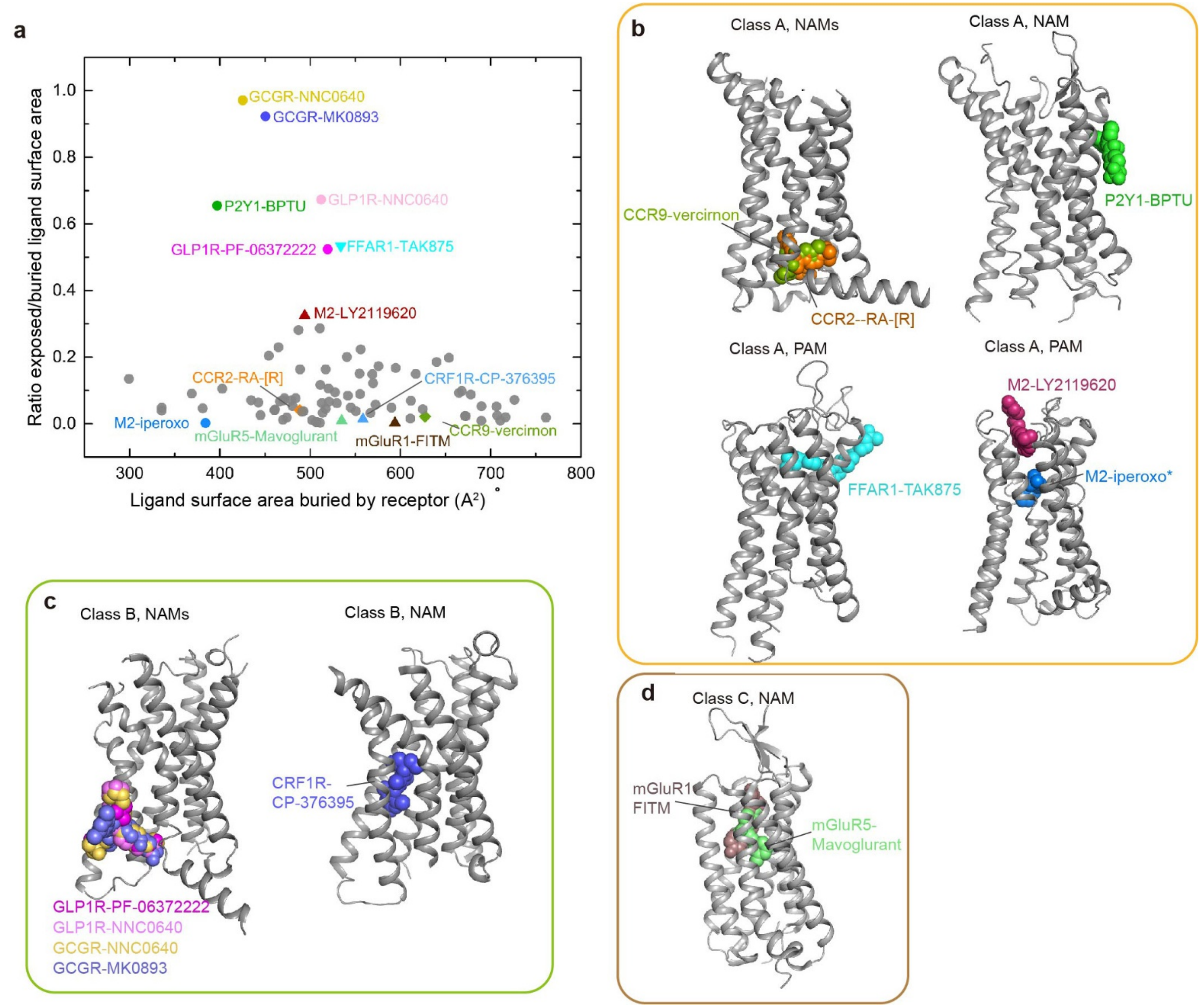
31. Chun, E. *et al.* Fusion partner toolchest for the stabilization and crystallization of G protein-coupled receptors. *Structure* **20,** 967–976 (2012).
32. Lv, X. *et al.* In vitro expression and analysis of the 826 human G protein-coupled receptors. *Protein Cell* **7,** 325–337 (2016).
33. Cherezov, V. *et al.* Rastering strategy for screening and centring of microcrystal samples of human membrane proteins with a sub-10 microm size X-ray synchrotron beam. *J. R. Soc. Interface* **6** (suppl. 5), S587–S597 (2009).
34. Kabsch, W. Xds. *Acta Crystallogr. D* **66,** 125–132 (2010).
35. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40,** 658–674 (2007).
36. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66,** 213–221 (2010).
37. Vagin, A. A. *et al.* REFMAC5 dictionary: organization of prior chemical knowledge and guidelines for its use. *Acta Crystallogr. D* **60,** 2184–2195 (2004).
38. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **66,** 486–501 (2010).
39. Smart, O. S. *et al.* Exploiting structure similarity in refinement: automated NCS and target-structure restraints in BUSTER. *Acta Crystallogr. D* **68,** 368–380 (2012).
40. Chen, V. B. *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D* **66,** 12–21 (2010).
41. Sastry, G. M., Adzhigirey, M., Day, T., Annabhimoju, R. & Sherman, W. Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *J. Comput. Aided Mol. Des.* **27,** 221–234 (2013).
42. Zhu, K. *et al.* Docking covalent inhibitors: a parameter free approach to pose prediction and scoring. *J. Chem. Inf. Model.* **54,** 1932–1940 (2014).
43. Klauda, J. B. *et al.* Update of the CHARMM all-atom additive force field for lipids: validation on six lipid types. *J. Phys. Chem. B* **114,** 7830–7843 (2010).
44. Yu, W., He, X., Vanommeslaeghe, K. & MacKerell, A. D., Jr. Extension of the CHARMM General Force Field to sulfonyl-containing compounds and its utility in biomolecular simulations. *J. Comput. Chem.* **33,** 2451–2468 (2012).
45. Pronk, S. *et al.* GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* **29,** 845–854 (2013).
46. Hess, B. P-LINCS: a parallel linear constraint solver for molecular simulation. *J. Chem. Theory Comput.* **4,** 116–122 (2008).
47. Bussi, G., Donadio, D. & Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **126,** 014101 (2007).
48. Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, A. & Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **81,** 3684 (1984).
49. Mitternacht, S. FreeSASA: An open source C library for solvent accessible surface area calculations. *F1000 Res.* **5,** 189 (2016).



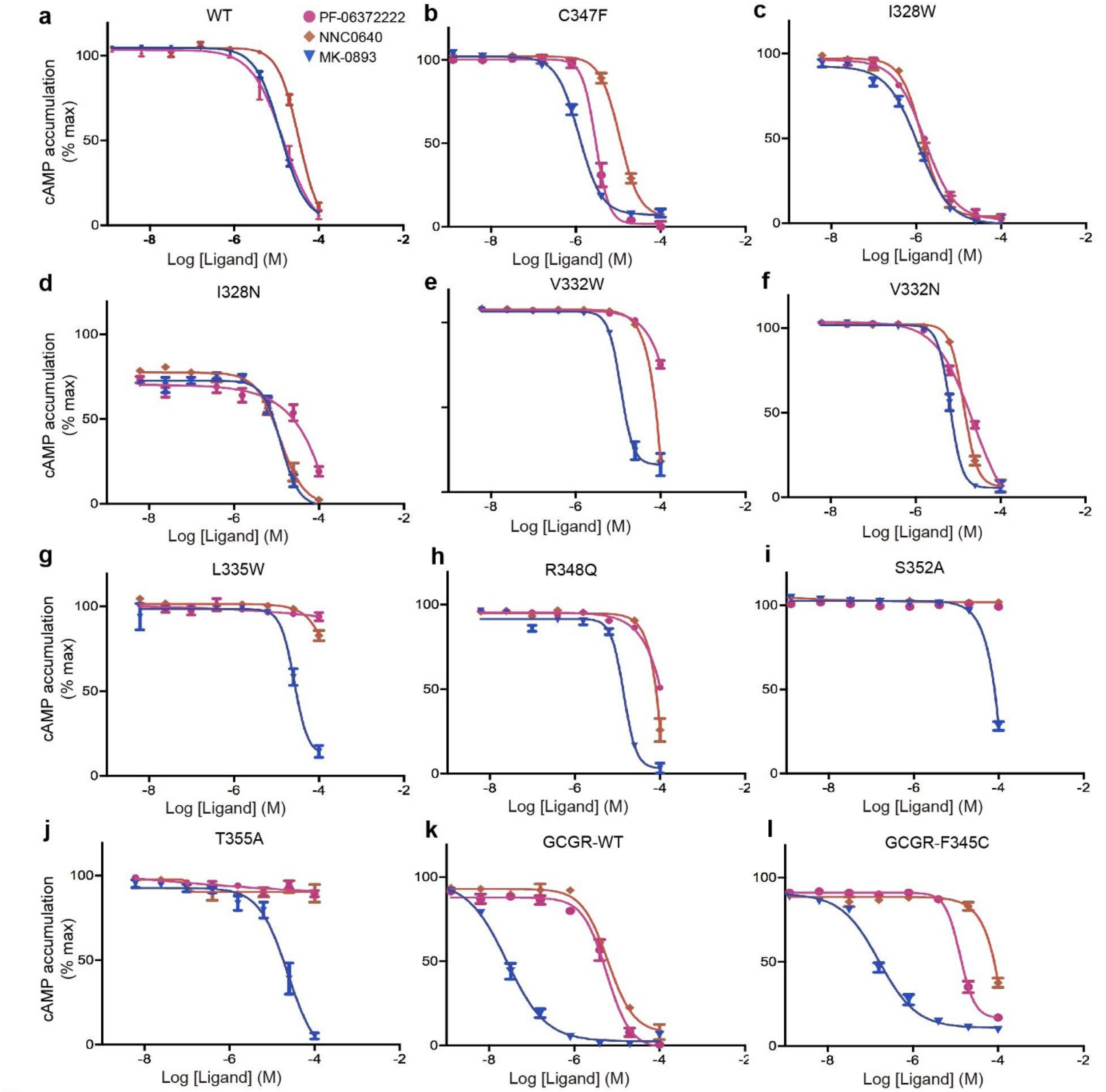
Extended Data Figure 1 | GLP-1R crystallization and structure determination. **a**, Schematic diagram of GLP-1R TMD construct (residues 128–431). Thermostabilizing mutations (green) are S193C, I196F, S225A, M233C, S271A, I317C, G318I, K346A, C347F and G361C. The single most conserved residue among class B GPCRs in each transmembrane helix (designated X.50b in the Wootten residue numbering scheme[16], in which 'X' is the transmembrane helix number) is indicated in bold. Disordered residues in the structure are shown with a brown background. Engineered ECL1 linker residues are shown as dashed circles. The endogenous and engineered disulfide bonds are shown with solid and dashed orange lines, respectively. **b**, Thermostability assay (CPM) of representative mutations. Constructs that include the disulfide bond (I317C/G361C) are much sharper in curve transition compared to the control (no mutation, black line). **c**, Thermostability assay of the apo state protein (crystallization construct) and with PF-06372222 or NNC0640. In **b** and **c**, the dotted and solid lines represent the original and fitted curves, respectively. **d**, Representative crystals of GLP-1R–PF-06372222 in lipidic cubic phase. **e**, Crystal packing of GLP-1R–PF-06372222, with the two molecules per asymmetric unit coloured green and cyan, respectively. **f**, The $|F_o| - |F_c|$ omit maps of PF-06372222 and NNC0640 contoured at 3.0σ.



Extended Data Figure 2 | GLP-1R interhelical interaction network. **a**, Overview of representative interactions. **b**, The conserved disulfide bond between C226$^{3.29b}$ and C296^{ECL2}. **c**, Hydrogen-bond interactions between helices I and VII. **d**–**f**, Interaction network between helices III and VI (**d**), helices III, IV and V (**e**), and helices II, III and IV (**f**). This figure can be compared with figure 4 in ref. 14 (CRF$_1$R) and figure 3 in ref. 15 (GCGR).



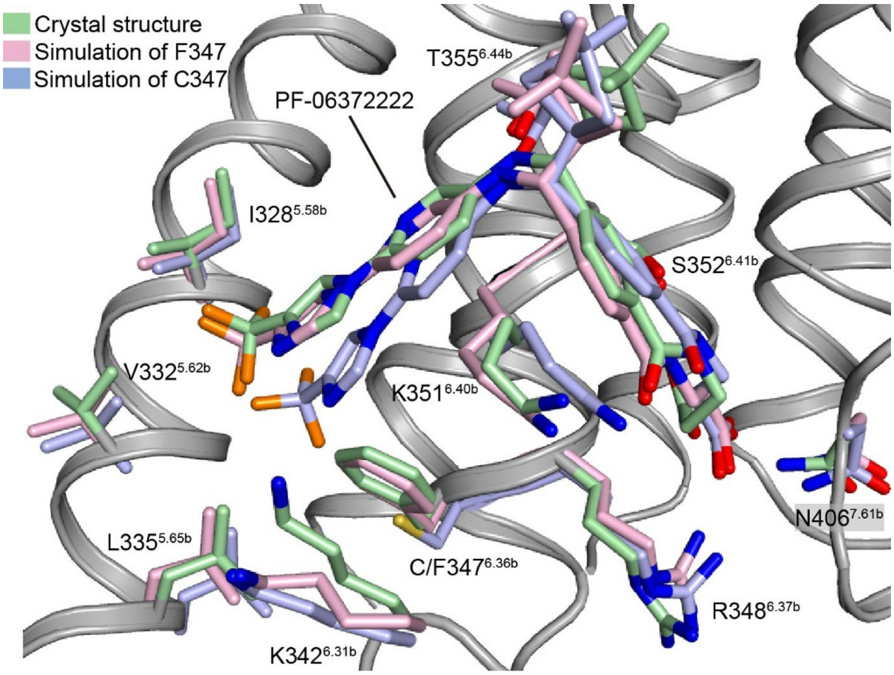
Extended Data Figure 3 | Statistics of buried and exposed surface areas of 86 crystallized GPCR ligands. **a**, Scatter plot of current crystallized ligands. *x* axis is the surface area buried by receptor; *y* axis is the ratio of exposed area to buried area. All surface areas were calculated based on the crystal structures using freeSASA[48]. PDBs with multiple chains are averaged in the plot. **b–d**, Current binding modes of GPCR allosteric modulators. For clarity, modulators that bind to similar region were arranged in the same cartoon.



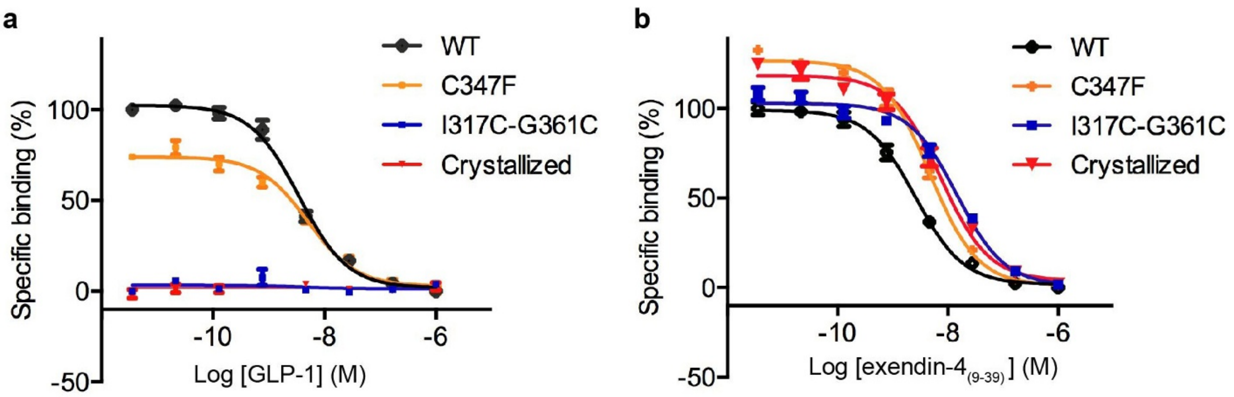
m Functional potency (pIC_{50})

	WT	C347F	I328W	I328N	V332W	V332N	L335W	R348Q	S352A	T355A	GCGR-WT	GCGR-F345C
PF-06372222	5.15±0.17	5.97±0.05	5.55±0.1	NA*	NA	4.52±0.09	NA	NA	NA	NA	5.16±0.13	4.95±0.17
NNC0640	4.63±0.09	5.12±0.09	5.79±0.04	5.13±0.09	NA	4.61±0.29	NA	NA	NA	NA	5.12±0.13	4.28±0.09
MK-0893	5.1±0.12	6.74±0.02	5.73±0.11	5.25±0.16	4.88±0.05	5.03±0.11	NA	4.77±0.04	NA	NA	7.04±0.18	6.21±0.1
Expression	100	142.6±21.1	86.9±7.2	51.5±3	205.1±48.6	75.1±18.6	55.1±12.1	173.2±33.7	143.2±24.4	95.2±14.1		

Extended Data Figure 4 | Effects of binding pocket mutations on potency of NAMs. a–l, Dose-dependent inhibition curves of NAMs on wild-type GLP-1R (**a**) or GLP-1R mutants (**b–j**), as well as wild-type GCGR (**k**) or the F345$^{6.36b}$C mutant (**l**). **m**, Summary of half-maximum inhibitory concentration (pIC_{50}) values of NAMs on the above constructs. Relative expression levels of mutated constructs were evaluated by comparing to that of wild-type GLP-1R. Experiments were repeated at least three times and error bars represent s.e.m. of quadruplicate measurements. NA, not available.



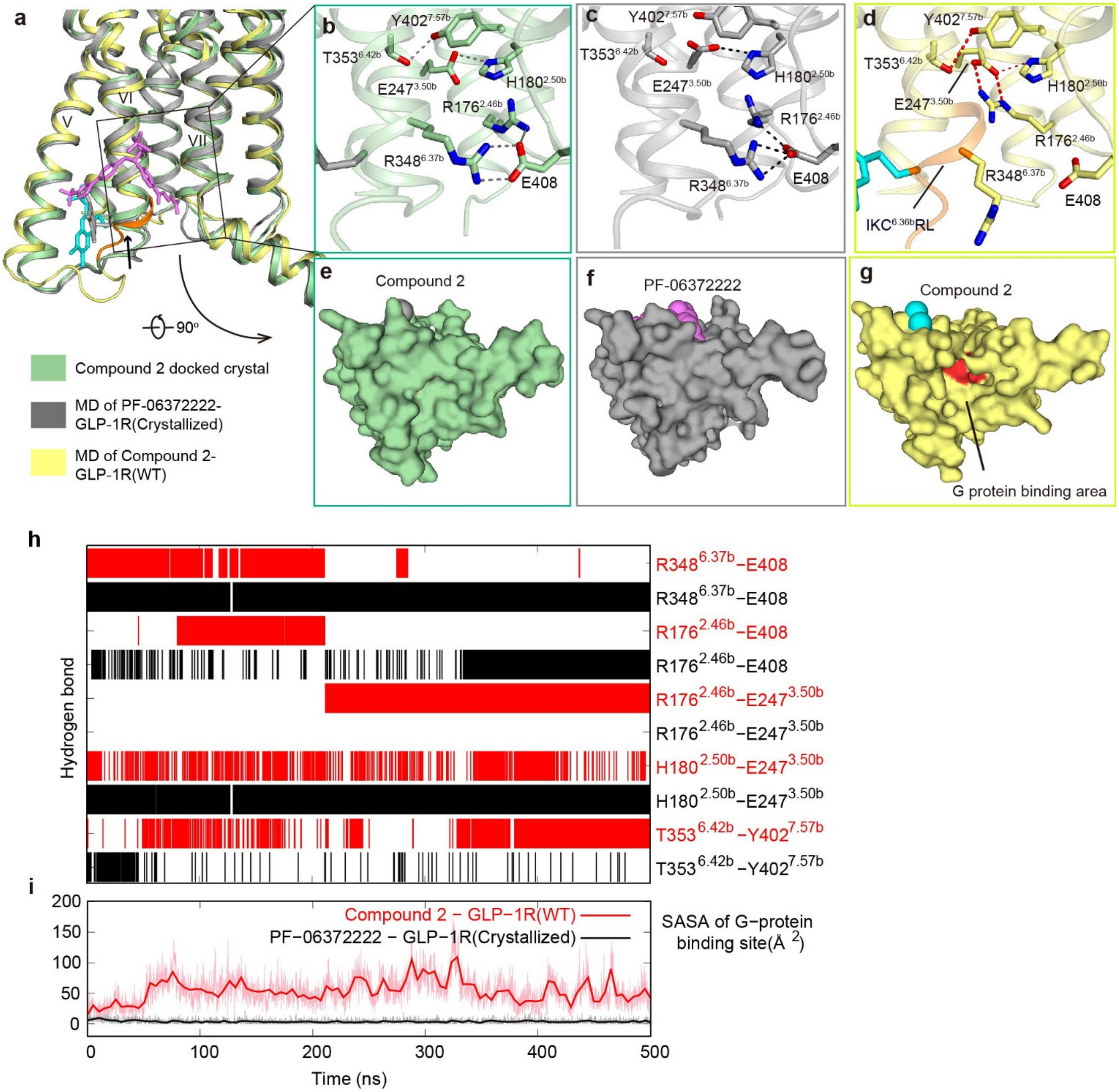
Extended Data Figure 5 | C347$^{6.36b}$F stabilizes the interaction interface between GLP-1R and NAMs. Superposition of PF-06372222-bound crystal structure (containing the C347$^{6.36b}$F mutation), with crystal structure based molecular dynamic simulations of F347 shown in blue (500 ns) and of simulations C347 (6.36b mutated back to cysteine) shown in pink (500 ns). PF-06372222 and key residues are shown as sticks. For clarity, only the backbone of the crystal structure is shown (grey). In the simulation of F347, the ligand adopts the same orientation as in the crystal structure, whereas the orientation of the ligand in the simulation of C347 is varied during the simulation process (the trifluoromethyl-pyrazole group in particular).



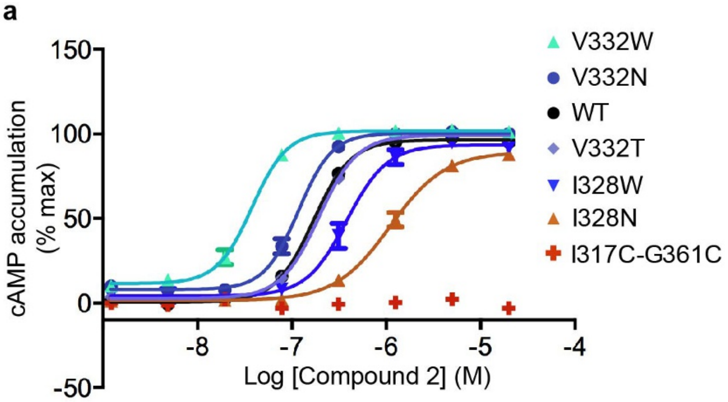
	WT	C347F	I317C-G361C	Crystallized
pIC_{50}	8.47±0.09	8.29±0.06	NB*	NB*
Span	100.7	71.32	1.896	0.1362

	WT	C347F	I317C-G361	Crystallized
pIC_{50}	8.58±0.04	8.3±0.05	7.84±0.13	8.13±0.1
Span	97.35	125.7	101.8	114.5

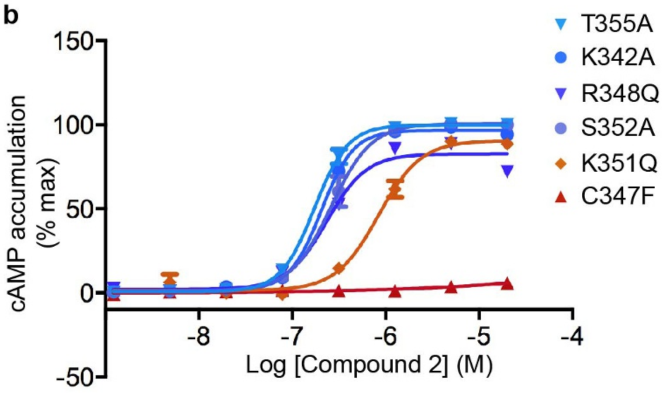
Extended Data Figure 6 | Binding of GLP-1 and exendin-4$_{(9-39)}$ to representative constructs. a, **b**, The four representative constructs used in Fig. 3d are tested for their binding properties with full agonist GLP-1 (**a**), and fragment antagonist exendin-4$_{(9-39)}$ that targets the extracellular domain (**b**). Experiments were repeated at least three times and error bars represent s.e.m. of duplicate measurements. NB, no binding.



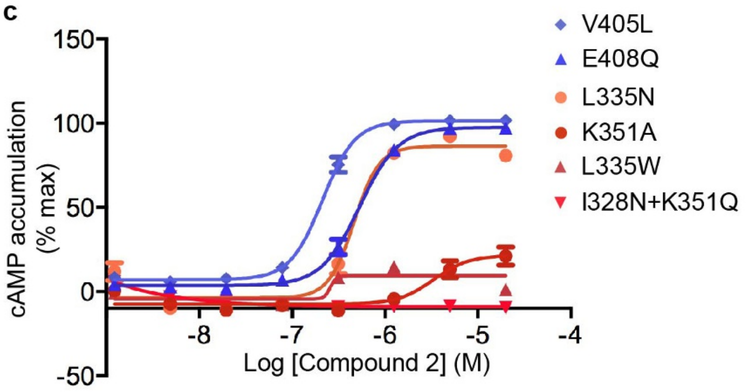
Extended Data Figure 7 | Conformational changes revealed by molecular dynamics simulation. **a**, Comparison of compound 2 (grey) docked to the GLP-1R crystal structure, after 500 ns molecular dynamics simulation (compound 2 in cyan), and molecular dynamics simulation of PF-06372222-bound GLP-1R (PF-06372222 in pink). **b**, The hydrogen-bond interaction network between residues associated with the ionic lock observed in the GLP-1R crystal structure[14,15,24]. **c**, In molecular dynamics simulation of the PF-06372222-bound GLP-1R crystal structure, the ionic lock hydrogen bond network is preserved. **d**, Molecular dynamics simulation of compound 2 covalently bound to wild-type GLP-1R reveals unwinding of the N-terminal end of helix VI (IKC$^{6.36b}$RL, coloured orange) and re-organization of the ionic lock interaction network. Unwinding of helix VI disrupts the ionic lock interactions between R348$^{6.37b}$ and E408$^{7.63b}$ and destabilizes the hydrogen-bond interaction between H180$^{2.50b}$ and E247$^{3.50b}$. These conformational changes allow R176$^{2.46b}$ to hydrogen bond with E247$^{3.50b}$, and reinforce the hydrogen bond between T353$^{6.44b}$ and Y402$^{7.57b}$ compared to the PF-06372222-bound GLP-1R molecular dynamics simulation. **e**–**g**, Intracellular views (surface representation) of **b**–**d**. **h**, Hydrogen-bond interactions between key residues during simulations in **c** and **d**. Hydrogen bonds were determined with the g_hbond program in the Gromacs[45], using a hydrogen bond distance cut-off of 3.5 Å and angle cut-off of 120°–240°. **i**, Six residues around the intracellular ionic lock of GLP-1R (H180$^{2.50b}$, L251$^{3.54b}$, L349$^{6.40b}$, S350$^{6.41b}$, T353$^{6.44b}$ and Y402$^{7.57b}$) were selected to calculate the solvent-accessible surface areas (SASA) using the program freeSASA[49]. Compared to the crystal structure or simulation of PF-06372222, these residues (marked red in **e**–**g**) were exposed to solvent by 40–100 Å^2 in the simulation of compound 2 (**g**, **i**). In **e** and **f**, these residues are buried, and thus are not visible, while in **g**, exposure of these residues provides space for the G protein to bind.



	WT	I328N	I328W	V332T	V332N	V332W	I317C-G361C
pEC_{50}	6.77	5.99±0.06	6.54±0.12	6.74±0.03	6.93±0	7.39±0.03	NA*

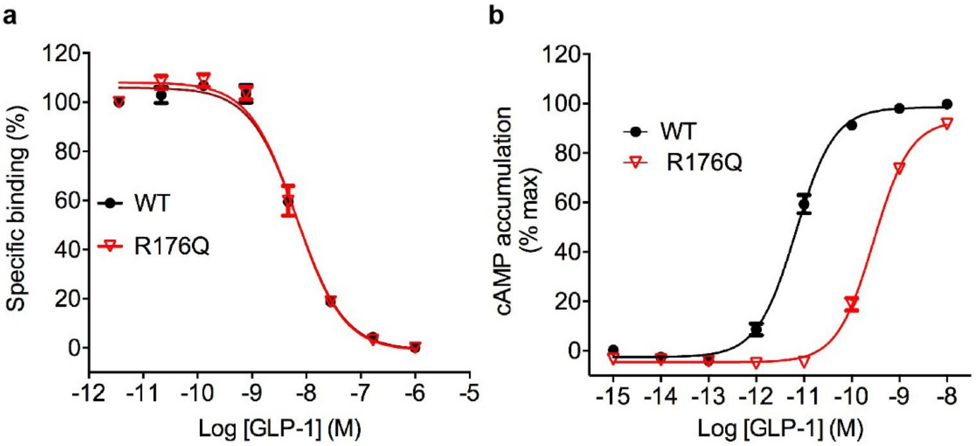


	K342A	C347F	R348Q	K351Q	S352A	T355A
pEC_{50}	6.72±0.03	NA*	6.65±0.06	6.11±0.03	6.66±0.08	6.84±0.08



	V405L	K351A	E408Q	L335N	L335W	I328N+K351Q
pEC_{50}	6.83±0.15	5.69±0.21	6.22±0.06	6.36±0.03	NA*	NA*

Extended Data Figure 8 | Effects of the PAM binding pocket mutations on potency of compound 2. a–c, Wild-type GLP-1R and 18 mutants (including 2 double mutants) were compared for their effects on potency of compound 2. Curves were coloured based on the same criteria as in Fig. 4 and codes were ranked based on their pEC_{50} (negative logarithm of the half-maximum effective concentration (EC_{50})) values (listed in the tables). Experiments were repeated at least three times and error bars represent s.e.m. of quadruplicate measurements.

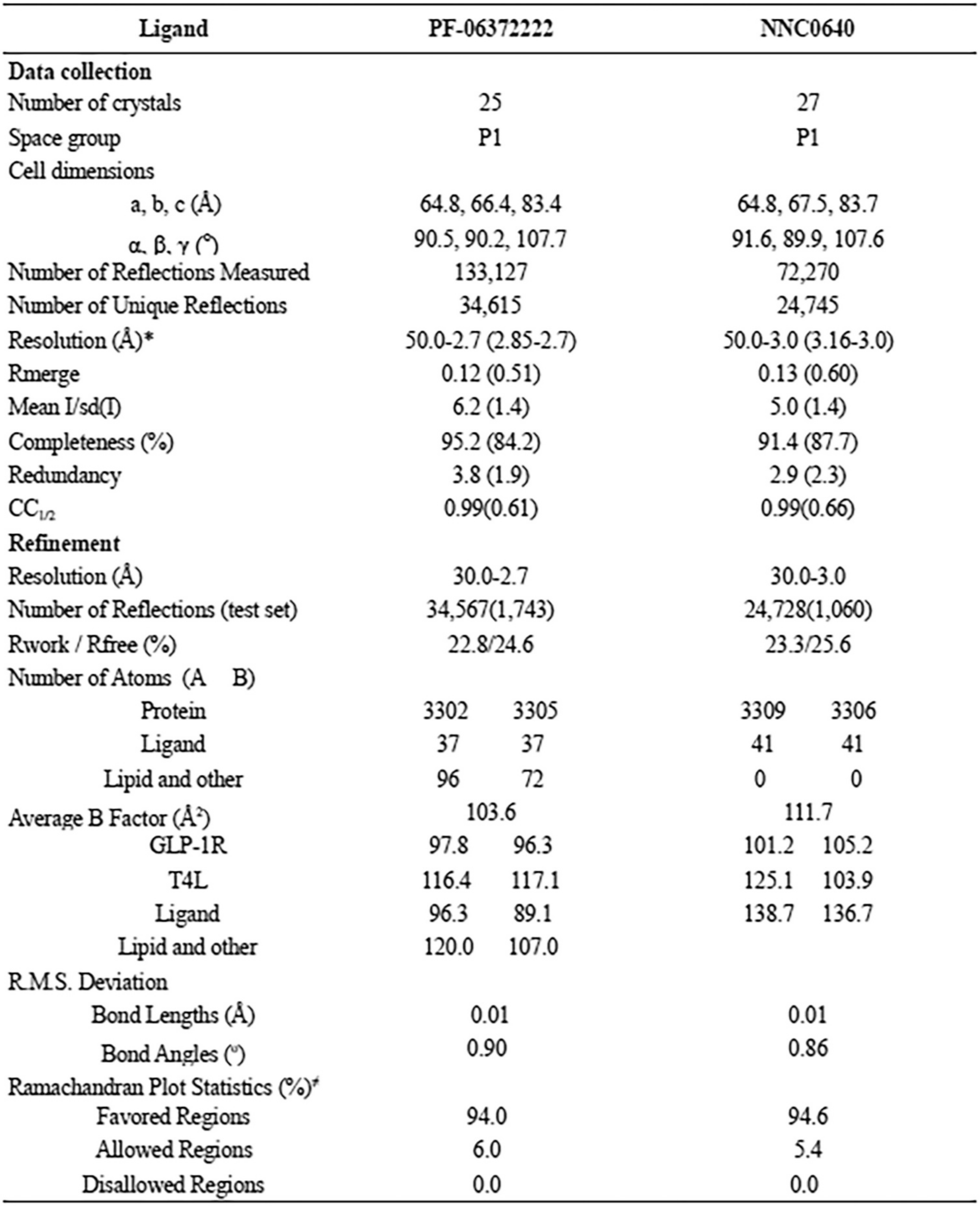


	WT	R176Q
pIC_{50}	8.17±0.03	8.19±0.13

	WT	R176Q
pEC_{50}	11.18±0.1	9.54±0.08

Extended Data Figure 9 | The R176Q mutation decreases the potency of GLP-1R, but does not affect its binding capacity with GLP-1. **a**, **b**, Comparison of wild-type GLP-1R and the R176$^{2.46b}$Q mutant by GLP-1 binding assay (**a**) and functional cAMP accumulation assay (**b**). cAMP accumulation assay is conducted as described in the text, and the binding assay was carried out using radiolabelled GLP-1 as a tracer and competing it with a serial dilution of unlabelled GLP-1. Experiments were repeated three times and error bars represent s.e.m. of duplicate (binding assay) or quadruplicate (cAMP assay) measurements.

Extended Data Table 1 | Data collection and structure refinement statistics



Ligand	PF-06372222		NNC0640	
Data collection				
Number of crystals	25		27	
Space group	P1		P1	
Cell dimensions				
a, b, c (Å)	64.8, 66.4, 83.4		64.8, 67.5, 83.7	
α, β, γ (°)	90.5, 90.2, 107.7		91.6, 89.9, 107.6	
Number of Reflections Measured	133,127		72,270	
Number of Unique Reflections	34,615		24,745	
Resolution (Å)*	50.0-2.7 (2.85-2.7)		50.0-3.0 (3.16-3.0)	
Rmerge	0.12 (0.51)		0.13 (0.60)	
Mean I/sd(I)	6.2 (1.4)		5.0 (1.4)	
Completeness (%)	95.2 (84.2)		91.4 (87.7)	
Redundancy	3.8 (1.9)		2.9 (2.3)	
$CC_{1/2}$	0.99(0.61)		0.99(0.66)	
Refinement				
Resolution (Å)	30.0-2.7		30.0-3.0	
Number of Reflections (test set)	34,567(1,743)		24,728(1,060)	
Rwork / Rfree (%)	22.8/24.6		23.3/25.6	
Number of Atoms (A B)				
Protein	3302	3305	3309	3306
Ligand	37	37	41	41
Lipid and other	96	72	0	0
Average B Factor (Å^2)	103.6		111.7	
GLP-1R	97.8	96.3	101.2	105.2
T4L	116.4	117.1	125.1	103.9
Ligand	96.3	89.1	138.7	136.7
Lipid and other	120.0	107.0		
R.M.S. Deviation				
Bond Lengths (Å)	0.01		0.01	
Bond Angles (°)	0.90		0.86	
Ramachandran Plot Statistics (%)#				
Favored Regions	94.0		94.6	
Allowed Regions	6.0		5.4	
Disallowed Regions	0.0		0.0	

*Highest resolution shell is shown in parentheses.

#As defined in MolProbity[40].