
Article:
Gómez Santiago, C, Paci, E orcid.org/0000-0002-4891-2768 and Donnelly, D orcid.org/0000-0001-6048-6995 (2018) A mechanism for agonist activation of the glucagon-like peptide-1 (GLP-1) receptor through modelling & molecular dynamics. Biochemical and Biophysical Research Communications, 498 (2). pp. 359-365. ISSN 0006-291X

https://doi.org/10.1016/j.bbrc.2018.01.110

A MECHANISM FOR AGONIST ACTIVATION OF THE GLUCAGON-LIKE PEPTIDE-1 (GLP-1) RECEPTOR THROUGH MODELLING & MOLECULAR DYNAMICS

Carla Gómez Santiago[a], Emanuele Paci[b] and Dan Donnelly[a£]

[a]School of Biomedical Sciences, [b]School of Molecular and Cell Biology,
Faculty of Biological Sciences, University of Leeds, Leeds LS2 9JT, UK.

Carla Gómez Santiago	bscgs@leeds.ac.uk
Emanuele Paci	e.paci@leeds.ac.uk
Dan Donnelly	d.donnelly@leeds.ac.uk

[£]Corresponding Author

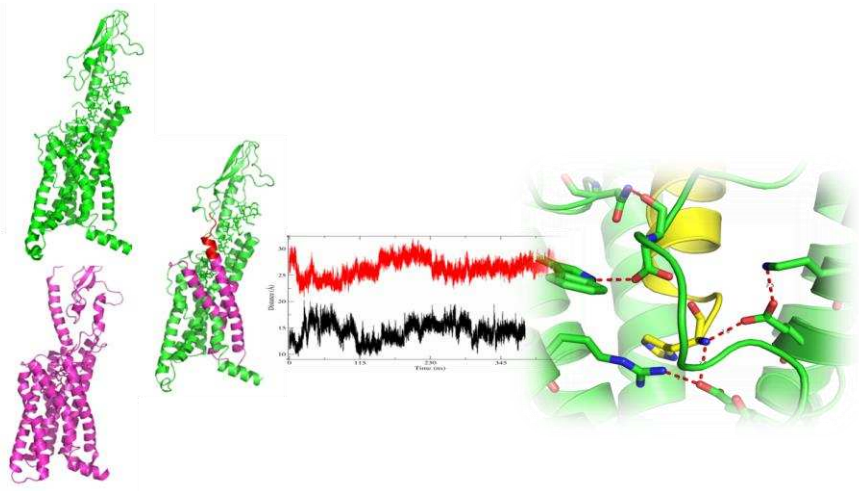
Abbreviations:

GPCR	G protein-coupled receptor
GLP-1	Glucagon-like peptide 1
GLP-1R	GLP-1 receptor
NTD	N-terminal extracellular domain
TMD	transmembrane domain
TM1-TM7	transmembrane helices 1-7
EM	electron microscopy
MD	molecular dynamics
RMSD	root mean square deviation

ABSTRACT

The receptor for glucagon-like peptide 1 (GLP-1R) is a validated drug target for the treatment of type 2 diabetes and obesity. Recently the first three structures of GLP-1R were published – an X-ray structure of the apo transmembrane domain in the inactive conformation; an X-ray structure of the full-length receptor bound to a truncated peptide agonist; and a cryo-EM structure of the full-length receptor bound with GLP-1 and coupled to the G protein G_s. Since the inactive structure was incomplete, and the two active-state structures shared significant differences, we utilised all available knowledge to build hybrid models of the full length active and inactive state receptors. The two models were simulated using molecular dynamics and the output trajectories analysed and compared to reveal insights into the mechanism for agonist-mediated receptor activation. His-7, Glu-9 and Asp-15 of GLP-1 act together to destabilise transmembrane helix 6 and extracellular loop 3 in order to generate an active conformation of GLP-1R.

GRAPHICAL ABSTRACT



KEY WORDS

GLP-1, GPCR, incretin, diabetes, molecular dynamics, model

INTRODUCTION

Glucagon-like peptide-1 (GLP-1) is a hormone secreted by enteroendocrine L cells in response to feeding. It has multiple physiological effects, most notable of which are the potentiation of glucose-induced insulin secretion and the induction of satiety [reviewed in 1]. As such, the GLP-1 receptor (GLP-1R) has been successfully targeted for anti-diabetic and anti-obesity therapy through the discovery of a number of peptidic GLP-1 analogues that have superior pharmacokinetic properties compared to those of the parent hormone [e.g. 2, 3]. Nevertheless, although there is emerging evidence for routes towards orally available peptide mimetics [4], they will require expensive dosing regimens, and there remains an unmet need for the development of commercially viable orally available non-peptidic drugs. In addition, the ability of GLP-1R to signal through both G-protein and G protein-independent pathways [5-7], opens up the possibility for new biased therapeutics that focus upon one particular pathway. However, in order to maximise the opportunities for rational discovery of such non-peptidic and/or biased agents, there is a requirement for a detailed understanding of how ligands bind and activate GLP-R

GLP-1R is a typical class B/secretin-family G protein-coupled receptor (GPCR) and as such is composed of two domains – the N-terminal extracellular domain (NTD) comprising the first ~115 residues, and the transmembrane domain (TMD) comprising the remaining 7 transmembrane helices and connecting regions. GLP-1 interacts with GLP-1R through a two-domain mechanism in which the C-terminal half of the α-helical ligand binds to the NTD, while the N-terminal half binds to the TMD [8, 9]. The latter interaction is critical for agonist-induced receptor activation and hence is the region of most interest when considering the discovery of novel GLP-1 agonists. Until recently, knowledge of the details of peptide-GLP-1R interactions

relied upon a large body of site-directed mutagenesis data (reviewed in [1]) coupled with molecular models that were derived from the isolated TMD structure of the glucagon receptor [10-12]. Although useful, the inactive conformation of the glucagon receptor template used, alongside the absence of a TMD-bound peptide ligand, resulted in GLP-1/GLP-1R models that required a degree of interpretation and speculation. However, a step change in our knowledge of GLP-1R structure came with the simultaneous publication of three papers in mid-2017 which described GLP-1R in its inactive and active states [13-15].

The inactive structure of the TMD of GLP-1R was obtained through two structures (pdb codes 5VEW and 5VEX) of GLP-1R, each coupled with a different negative allosteric modulator which helped to stabilise the inactive conformation [13]. The crystallised receptor construct lacked the NTD and had a number of alterations from the wild type sequence. The active state of GLP-1R could be observed through two further molecular structures – one being the cryo-EM structure of rabbit GLP-1R in complex with both G_s and GLP-1 (5VAI; 3.9Å resolution in region of interest; [14]), and the other being the X-ray crystal structure of a thermo-stabilised GLP-1R in complex with a non-natural C-truncated peptidic agonist (5NX2; [15]). The latter structure (5NX2) contained 11 stabilising mutations and displayed wild type-like agonist affinity but with reduced potency, suggesting a partially active conformation. The agonist was an 11-residue peptide based on the N-terminus of GLP-1, sharing five identical residues alongside several non-natural variants. Despite these differences, the resultant binding site correlated closely with the published site-directed mutagenesis data. The cryo-EM structure (5VAI) described a 150 kDa complex that included GLP-1, the complete rabbit GLP-1R, and the G_s heterotrimeric G-protein stabilised by a nanobody (Nb35). This provided the first view of GLP-1R coupled to both its natural ligand and its primary G protein.

However, despite many similarities between the 5VAI and 5NX2 active GLP-1R structures, there remain a number of significant differences between these active state structures which impede a full understanding of agonist binding and activation (discussed below). Therefore, in order to fully understand how agonists bind and activate GLP-1R, we set out to build a high quality, complete, active-state, human GLP-1R model, docked with GLP-1, using a combination of both the 5VAI and 5NX2 structures as templates. This hybrid model was subjected to molecular dynamics simulations to refine the model and ascertain the dynamic details of the peptide-receptor interaction. A second objective was to generate a complete inactive wild type human GLP-1R model, by combining the inactive TMD of 5VEW with the empty NTD of 5NX2, and then to examine the stability of this model using molecular dynamics in order to compare the inactive conformation to that of the agonist-bound GLP-1R.

MATERIALS AND METHODS

All molecular model building manipulations were carried out using the tools embedded within PyMol (The PyMOL Molecular Graphics System, Version 1.7.2.3 Schrödinger, LLC.) unless otherwise stated.

Building active GLP-1R (**Supplementary Figure S1A**)**.** (i) 5VAI and 5NX2 were structurally aligned by superimposing the residues of their TMDs using the align function in PyMol. (ii) All residues in 5NX2 were deleted, apart from E139-K197 and E373-F393. (iii) The G-protein, nanobody, ligand and regions E139-K197 and E373-F393 were deleted from the 5VAI structure. (iv) The remaining atoms from both molecules were merged and saved to a single pdb file which was then used as a template to build a model of human GLP-1R using the homology modelling

server SWISS-MODEL [16; http://swissmodel.expasy.org/]. Residues Ser-129 to Glu-138, connecting the NTD and TMD, were built by SWISS-MODEL, as were any missing or altered side chains in the template. All other residues were built from the template and matched the starting conformations. The output pdb file from SWISS-MODEL represented the apo hybrid model of active human GLP-1R from Thr-29 to Arg-421. (v) GLP-1 and α5 of G_s were added back into the model by structurally aligning the SWISS-MODEL output with 5VAI, extracting the co-ordinates for the GLP-1 and α5 (374-394) segments from 5VAI, and merging them with the active GLP-1R hybrid model.

Building inactive GLP-1R (**Supplementary Figure S1B**)**.** The NTD from 5NX2 represents the only apo structure of this domain and was hence selected as the NTD to fuse with the TMD from 5VEW. (i) One monomer from the crystallographic dimer of 5VEW was deleted, as was the fusion partner. (ii) The remaining residues from 5VEW were structurally aligned with 5NX2 by superimposing the residues of their TMDs using the align function in PyMol. (iii) All residues in the TMD of 5NX2 were deleted, leaving only the NTD residues from Thr-29 to Glu-133 and also Met204 to Asp215 (the latter is the part of ECL1 which contacts the NTD on 5NX2 but is missing in 5NEW). (iv) The remaining atoms were merged and saved to a single pdb file. (v) This was used as a template to build a model of inactive human GLP-1R using SWISS-MODEL. All the missing regions in the template were built by SWISS-MODEL, as were all missing or mutated side chains. All other residues built from the template matched the starting conformations. The output represented the apo hybrid model of inactive human GLP-1R from Thr-29 to Arg-421.

Molecular Dynamics

The starting models were first orientated with respect to the hydrocarbon core of the lipid bilayer utilising the OMP database and server [17]. The output pdb files were then used in the bilayer builder input generator of CHARMM-GUI [18]. The four disulphide bonds of GLP-1R were first defined and the models were then embedded into a 90 x 90 $Å^2$ POPC bilayer membrane before a water/solvent box was added (TIP3P and 0.15 M NaCl). Simulations were all performed at 303.15K temperature and 1 atm pressure, using Langevin dynamics with a damping coefficient of 1/ps to control the temperature and a Nose-Hoover Langevin piston to control the pressure. A timestep of 2 fs was used, and Particle Mesh Ewald to account for long-range electrostatics. A simple POPC bilayer has been used previously to simulate a model of GLP-1R [12] and indeed we found that this bilayer closely matched the boundaries predicted by OMP (see supplementary Figure S2). The models were subjected to equilibration (70 ns) and production (385-460 ns) using NAMD, a program developed by the Theoretical and Computational Biophysics Group in the Beckman Institute for Advanced Science and Technology at the University of Illinois at Urbana-Champaign [19]. Output trajectories were analysed using WORDOM and VMD [20, 21]. The simulation trajectory for the active hybrid model was analysed in detail between 280-370 ns using VMD in order to estimate the fraction of time in which all residue-residue hydrogen bonds were formed (defined as heavy atom distance of 3.5 Å or less, with an angle of 25° or less). A similar analysis was carried out for the inactive hybrid model trajectory between 200-380 ns. Example models in Figure 3 & 4 and in the Supplementary information are taken from the central frame of each of these two analysis windows (at 285 ns and 325 ns respectively for the active and inactive models).

RESULTS AND DISCUSSION

Rationale for a hybrid GLP-1R model

Despite many similarities between the TMD in the 5VAI and 5NX2 active GLP-1R structures, there are a number of differences that need to be resolved for a full understanding of agonist binding and activation mechanisms. While the relative positions of the NTD and TMD differ in the active 5VAI and 5NX2 structures, this is perhaps not surprising given the that flexibility between these domains in related receptors has been analysed and discussed previously [22]. Furthermore, since GLP-1 straddles the two domains in 5VAI, which limits their relative movement, the shorter 11-residue ligand in 5NX2 is bound to the TMD and is largely independent of NTD binding, hence enabling it more conformational flexibility between the domains.

The most surprising difference between 5NX2 and 5VAI involves the conformation of TM6. While both structures display the expected outward movement of TM6 relative to that observed in the inactive 5VEW structure, to create the required binding site for α5 of G_s, the observed conformation and position of TM6 differs significantly in each active structure. TM6 in 5VAI is highly distorted, non-linear and closely packed against the rest of the TM bundle in the central region (Figure 1A), in close agreement to that observed in the structure of the calcitonin receptor TMD in complex with G_s (Figure 1B; pdb code 5UZ7; [23]). However, TM6 in 5NX2 is much more regular and linear but is displaced away from the rest of the TM bundle such that there appears to be an unusual channel joining the extracellular peptide binding cavity to the intracellular space (Figure 1C). In our molecular dynamics simulations using 5NX2 without its ligand, the TM bundle fractured between TM5 and TM6 and allowed bilayer lipids to penetrate the receptor core (data not shown). Hence, given this 5NX2 instability, coupled with the good correlation between 5VAI and 5UZ7, and also that 5VAI represents a wild type GLP-1R sequence

coupled to GLP-1 and G_s, we elected to utile the TM6 conformation from 5VAI in our hybrid model.

Nevertheless, there are some aspects of the 5VAI structure that appear to be less secure than 5NX2. In particular, the Cα positions of Arg-380 and Lys-288 at the extracellular end of TM7 are displaced by one helical turn between the two structures (Supplementary Figure S3). While in 2NX2 they interact with the ligand as expected from site-directed mutagenesis data and previous modelling studies (reviewed in [1]), in the 5VAI structure they point away from the TM bundle and would hence interact with the lipid head-group region of the bilayer. Given the known important role of Arg-380 and Lys-383 in peptide binding, we elected to take this region of TM7 from 5NX2 when constructing the hybrid human GLP-1R model. As a consequence, we also took the contacting regions on TM1 and part of TM2 from 5NX2 in the hybrid model (Figures 2A and S1). While we believe such hybrid modelling represents the best approach (in the absence of reliable high-resolution structures) for creating a starting model for molecular dynamics simulations, there are likely to be a number approximations and associated errors in the model. For example, we must rely on SWISS-MODEL to accurately replace side chain conformations and to fill in missing gaps in the template primary structure, including the "stalk" region linking the two domains. Furthermore, the newly created protein-protein interactions and the melding points between the different fragments used to create the hybrid models may result in less than optimal interactions and conformations, although we have used long relaxation times during the equilibration stages which should have fixed local inaccuracies. Nevertheless, it is important to interpret the output carefully in light of the mutagenesis and molecular pharmacological data in order to identify meaningful outcomes.

Simulations of hybrid active and inactive human GLP-1R models

Due to the higher resolution of the single template structure used to build the TMD of the inactive model, relative to the two lower resolution structures used to build the active hybrid TMD, we expected the former to reach stability earlier in the MD simulations. Indeed, this was the case with a positional root mean square deviation (RMSD) relative to the starting model TMD plateauing at 60 ns and remaining stable throughout the remainder of the 385 ns simulation (Figure 1C). The active hybrid model took longer to reach stability but nevertheless the RMSD plateaued after 230 ns (Figure 2C).

The principle difference between the inactive and active conformations of GPCRs is believed to be the outwards movement of the cytoplasmic half of TM6 away from the TM bundle in order to create the binding site for α5 of the G protein [24]. Such a movement can be observed when comparing the inactive 5VEX and 5VEW structures with those of the active conformations 5VAI and 5NX2 [13-15]. For example, the distance between the Cα atoms of residues Arg-176 (TM2) and Arg-348 (TM6) is 23.7 Å in 5NX2 and 25.3 Å in 5VAI, whereas it is only 12.1 Å in 5VEW. These distances were maintained throughout the simulations of both our inactive and active models (Figure 2D). The outward movement of the cytoplasmic end of TM6 upon activation is accompanied by a large perturbation of the helix at Gly-361 and a movement of the segment of TM6 incorporating His-363 and Glu-364 towards the cytoplasmic side (Cα movements of 6.2 Å and 4.4 Å respectively based on the starting models – see Supplementary Figure S4). In addition, TM7 tilts towards TM6 at the extracellular end, with the hinge being at Gly-395, with an additional clockwise twist of the helix to move Arg-380 and Arg-383 from the exterior of the protein into the core of the receptor where they interact directly with the ligand (Figure 3). ECL3, which links the top of TM6 and TM7, undergoes a significant conformational rearrangement as a

result of these helical movements. The space created by the movement of TM7 towards TM6 is filled by the tilting of the extracellular side of TM1 around Gly-151. The final consequence is the outward movement of TM6 at the cytoplasmic end where the G protein docks. In our simulations, residues 382-394 region of the α5 helix of G_s remained stable and helical through the 460 ns active state simulation (Supplementary Figure S6G).

In the inactive hybrid model, Arg-190 was able to remain hydrogen bonded with Gln-394 for 40% of the trajectory (Figure 3A, blue). However, as can be seen, this interaction was abolished by ligand binding (Figure 3A, green) since Arg-190 (yellow) instead interacted with Glu-9* of the ligand (84%; GLP-1 residues will be identified with subscript * throughout). In addition, Glu-9* also interacts with the side chain hydroxyl groups of Tyr-148 and Tyr-152 (78% and 76% respectively), both residues that have been identified as being involved in GLP-1 recognition (reviewed in [1]). The movement of Arg-190 to interact with Glu-9* results in Gln-394 being free to move and interact with the exposed main chain carboxyl oxygen of Pro-358 of TM6, providing a helix cap for the cytoplasmic half of this helix following its disruption at Gly-361 (Figure 3B). This represents a way in which agonist binding could stabilise one type of active state of GLP-1R. GLP-1 potency is sensitive to Arg-190 substitution, as is a synthetic variant of oxyntomodulin in which the native Gln at the third position was replaced by Glu [25]. Given that native oxyntomodulin (and synthetic Q9-GLP-1) can activate the cAMP pathway but are not sensitive to Arg-190 substitution [25], it is likely that they stabilise a different active state which does not rely upon direct Arg-190 binding or the breakage of the Arg-190/Gln-394 interaction.

A second clear change in hydrogen bond formation resulting from ligand binding can be observed with Arg-380. In the inactive model, this residue forms a very stable hydrogen bond with Glu-373 throughout the trajectory (96%). Upon ligand binding, Arg-380 undergoes a significant

movement to interact with Asp-15* of the ligand (77%) (in agreement with mutagenesis data [1]), while the interaction with Glu-373 is also maintained (82%), albeit requiring a substantial conformational rearrangement of ECL3 (Figure 3D). Hence the movement of Arg-380 towards Asp-15* of the ligand encourages the rearrangement of ECL3, which is ultimately an essential requirement for the movement of TM6 and the creation of a G_s binding site. The position of Arg-380 is conserved as a positively charged residue (Arg or Lys) in the receptors for glucagon, GLP-2 and GIP, all of which also have a negatively charged residue (Asp or Glu) in the ligand at the position equivalent Asp-15* in GLP-1. Indeed, the oppositely charged residues have been interchanged from ligand to receptor in an elegant study by Moon et al. [27] showing a reciprocal rescue which strongly implicates the two residues in an interaction. In this study, Arg-380 was replaced by Asp, resulting in almost a 2000-fold reduction in potency. However, Arg9*-GLP-1, which had almost 100-fold lower potency at wild-type GLP-1R, was shown to have 120-fold improved potency at the Arg-380–Asp mutant receptor.

A third arginine residue (Arg-299) also has substantially different interactions in the active and inactive model trajectories. Despite starting off pointing downwards towards the receptor core as in the 5VAI template, Arg-299 rapidly moved out of the binding pocket and then interacted with both Glu-294 (57%) and Glu-21* (40%) of the ligand (Supplementary Figure S6 A) as previously predicted from earlier mutagenesis and modelling studies [11]. Indeed, Arg-299 was placed in this external position in the 5NEX X-ray structure, adding further evidence towards this conformation in the active state. In the inactive model, Arg-299 forms a very stable interaction with both Glu-34 (97%) on the NTD, and with Asp-372 (97%) of ECL2. Indeed, a characteristic of the active state is the close interaction between ECL2 and ECL3 – for example, Glu-294, Asn-300 and Asn-304

interact with Asp-372 and Arg-376 (Supplementary Figure S6 B). This inter-loop interaction is absent in the active state.

Ligand binding in active conformation.

Glu-9* is clearly a critical residue, interacting with Arg-190 (84%), Tyr-148 (78%) and Tyr-152 (76%), and to a lesser extent with Thr-391 (16%). Likewise, Arg-380 is important since it interacts with Asp-15* (77%). Additional interactions are Thr-13* with Lys-297, and Glu-21* with Arg-299 – both residues have been shown to be involved in agonist recognition [reviewed in 1]. However, the key residue for agonist-induced receptor activation in GLP-1R is the N-terminal His-7*. In the early stages of the simulation, the positively charged N-terminal moiety interacted with Glu-9*, forming a four-way interaction with Arg-190 and Glu-364. However, this intra-ligand salt bridge eventually broke and consequently the distance between Arg-190 and Glu-264 increased. While Arg-190 continued to interact with Glu-9*, the N-terminal amine of His-7* interacted with both Glu-364* (78%) and Glu-387 (66%). It is interesting to note that, while there have been a number of mutations of these residues [reviewed in 1], it was only the double substitution by Yang et al. [28] which abolished ligand binding, suggesting that either of these Glu residues can partially compensate for the other in the less deleterious single mutations. The sharing of the interaction with the positively charged amino-terminus of GLP-1 by these two acidic side chains forms the centre of a network of salt bridges and hydrogen bonds involving several residues known to be critical for full agonist recognition (Asn-300, Trp-306, Arg-310, Asp-372, and Lys-383). It is interesting to note that Arg-310 and Lys-383 both directly stabilise Glu-364 and Glu-387, respectively, but display no direct interactions with GLP-1. This is in keeping with mutagenesis and pharmacological data which showed their mutation caused minimal disruption of GLP-1

binding affinity compared with much more significant impairment of GLP-1 mediated receptor activation [11].

We have demonstrated and rationalised the importance of interactions between (i) Asp-15* and Arg-380, rotating TM7 and re-configuring ECL3 by pulling Glu-373 towards the ligand (Supplementary Figure S6A and S6B); (ii) Glu-21* and Arg-299, freeing Asp-372 and contributing further to the ECL3 reconfiguration (Supplementary Figure S6A and S6B); (iii) His-7* with Glu-364 (stabilised by Arg-310) and with Glu-387 (stabilised by Lys-383) – the movement of Glu-364 is possibly a critical characteristic of the movement and distortion of TM6 during the formation of the active state (Supplementary Figure S6C and S6D); and (iv) Glu-9* and Arg-190, freeing Gln-394 to stabilise the distortion in TM6 (Supplementary Figure S6E-G). By acting together, these various ligand-receptor interactions result in the stabilisation of an active receptor state in which TM6 moves to enable G_s to bind (Supplementary Figure S6G). Since N-terminally truncated GLP-1 can activate GLP-1R at high ligand concentrations in recombinant systems [26], it is likely that interactions (i) and (ii) are sufficient to generate the active state through the Asp-15*/Glu-21*-mediated disruption of ECL3. However, clearly the additional interactions via His-7* make this transition substantially more efficient.

In summary, we have demonstrated that fragments of several GLP-1R structures can be used to create stable and meaningful receptor models which can be simulated and analysed in order to answer important questions linking structure to function, and to propose agonist-mediated receptor activation mechanisms.

ACKNOWLEDGEMENTS

CGS is funded by CONACYT (Mexico).

REFERENCES

[1] de Graaf C., Donnelly D., Wootten D., et al., Glucagon-like peptide-1 and its class B G protein-coupled receptors: A long march to therapeutic successes, Pharmacological Reviews, 68 (2016) 954-1013.

[2] Lau J., Bloch P., Schäffer L., et al., Discovery of the Once-Weekly Glucagon-Like Peptide-1 (GLP-1) Analogue Semaglutide, J. Med. Chem., 58 (2015) 7370–7380.

[3] Ladenheim, E. E., Liraglutide and Obesity: A Review of the Data so Far, Drug Design, Development and Therapy, 9 (2015) 1867–1875.

[4] Davies, M., Pieber T.R., Hartoft-Nielsen, M-L. et al., Effect of Oral Semaglutide Compared With Placebo and Subcutaneous Semaglutide on Glycemic Control in Patients With Type 2 Diabetes: A Randomized Clinical Trial, JAMA, 318 (2017) 1460-1470.

[5] Jorgensen R., Kubale V., Vrecl M., et al., Oxyntomodulin Differentially Affects Glucagon-Like Peptide-1 Receptor β-Arrestin Recruitment and Signaling through Gα, Journal of Pharmacology and Experimental Therapeutics, 322 (2007) 148-154.

[6] Wootten D., Simms J, Miller L.J., et al., Polar transmembrane interactions drive formation of ligand-specific and signal pathway-biased family B G protein-coupled receptor conformations PNAS, 110 (2013) 5211-5216.

[7] Zhang H., Sturchler E., Zhu J., et al., Autocrine selection of a GLP-1R G-protein biased agonist with potent antidiabetic effects, Nat Commun., 6 (2015) 8918.

[8] López de Maturana R., Willshaw A., Kuntzsch A., et al., The Isolated N-terminal Domain of the Glucagon-like Peptide-1 (GLP-1) Receptor Binds Exendin Peptides with Much Higher Affinity than GLP-1. J Biol Chem, 278 (2003) 10195-10200.

[9] Al-Sabah S., Donnelly D., A model for receptor-peptide binding at the glucagon-like peptide-1 (GLP-1) receptor through the analysis of truncated ligands and receptors, Br J Pharmacol, 140 (2003) 339-346.

[10] Siu F.Y., He M., de Graaf C., et al., et al. Structure of the human glucagon class B G-protein-coupled receptor, Nature, 499 (2013) 444-449.

[11] Dods R.L., Donnelly D. The peptide agonist-binding site of the glucagon-like peptide-1 (GLP-1) receptor based on site-directed mutagenesis and knowledge-based modelling, Bioscience Reports, 36 (2016) e00285.

[12] Wootten D., Reynolds C.A., Koole C., et al. A Hydrogen-Bonded Polar Network in the Core of the Glucagon-Like Peptide-1 Receptor Is a Fulcrum for Biased Agonism: Lessons from Class B Crystal Structures Molecular Pharmacology, 89 (2016) 335-347.

[13] Song, G., Yang, D., Wang, Y., et al., Human GLP-1 receptor transmembrane domain structure in complex with allosteric modulators, Nature, 546 (2017) 312-315.

[14] Zhang, Y., Sun, B., Feng, D. et al. Cryo-EM structure of the activated GLP-1 receptor in complex with a G protein, Nature, 546 (2017) 248-253.

[15] Jazayeri, A., Rappas, M., Brown, A.J.H. et al., Crystal structure of the GLP-1 receptor bound to a peptide agonist, Nature, 546 (2017) 254-258.

[16] Biasini M., Bienert S., Waterhouse A., et al., SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information, Nucleic Acids Res, 42 (2014) W252–W258.

[17] Lomize, M.A., Pogozheva, I.D., Joo, H., et al., OPM database and PPM web server: resources for positioning of proteins in membranes, Nucleic Acids Res, 40 (2012) D370-376.

[18] Brooks, B.R., Brooks, C.L., 3rd, Mackerell, A.D., Jr. et al., CHARMM: the biomolecular simulation program, J Comput Chem, 30 (2009) 1545-1614.

[19] Phillips J.C., Braun R., Wang W., et al., Scalable molecular dynamics with NAMD, Journal of Computational Chemistry, 26 (2005) 1781-1802.

[20] Humphrey W., Dalke A. and Schulten, K. VMD: visual molecular dynamics, J Mol Graph, 14 (1996) 33-38.

[21] Seeber, M., Cecchini, M., Rao, F., et al., Wordom: a program for efficient analysis of molecular dynamics simulations, Bioinformatics, 23 (2007) 2625-2627.

[22] Yang L, Yang D, de Graaf C et al., Conformational states of the full-length glucagon receptor, Nat Commun, 6 (2015) 7859.

[23] Liang, Y.L., Khoshouei, M., Radjainia, M., et al., Phase-plate cryo-EM structure of a class B GPCR-G-protein complex, Nature, 546 (2017) 118-123.

[24] Rasmussen, S.G., DeVree, B.T., Zou, Y., et al., Crystal structure of the beta2 adrenergic receptor-Gs protein complex, Nature, 477 (2011) 549-555.

[25] Wootten D., Reynolds C.A., Smith K.J., The Extracellular Surface of the GLP-1 Receptor Is a Molecular Trigger for Biased Agonism, Cell, 165 (2016) 1632-1643.

[26] Donnelly D., The structure and function of the glucagon-like peptide-1 receptor and its ligands, Brit J Pharmacol, 166 (2012) 27-41.

[27] Moon M.J., Lee Y.N., Park S., et al., Ligand binding pocket formed by evolutionarily conserved residues in the glucagon-like peptide-1 (GLP-1) receptor core domain, J. Biol. Chem. 290 (2015) 5696–5706.

[28] Yang D., de Graaf C., Yang L., et al., Structural Determinants of Binding the Seven-transmembrane Domain of the Glucagon-like Peptide-1 Receptor (GLP-1R), J. Biol. Chem. 291 (2016) 12991-3004.

FIGURE LEGENDS

FIGURE 1.

Ribbon representations of 5VAI (**A**), 5UZ7 (**B**) and 5NX2 (**C**). Ligands are shown as sticks, TM6 is shown in black, and the cavity space is shown for A and C as yellow. TM6 is highly distorted in 5VAI and 5UZ7 and bends inwards at the centre to pack closely against the rest of the helical bundle. On the other hand, TM6 in 5NX2 is more regular and leaves an unusual cavity within the helical bundle that joins the cytoplasmic side to the extracellular binding pocket.

FIGURE 2.

Ribbon representations of the hybrid inactive (**A**) and hybrid active (**B**) starting models, with GLP-1 shown in stick form in B. Regions derived from 5VEW (blue), 5NX2 (magenta) and 5VAI (green) are shown, although side chains were all converted to those of human wild type GLP-1R by SWISS-MODEL. Red regions were built by SWISS-MODEL as they were missing from the templates. **C** Root mean square deviation (RMSD) relative to the starting models for the simulations of the hybrid GLP-1R models - inactive (black, 385 ns) and active (red, 460 ns). **D** Distance between the Cα atoms of Arg-176 (TM2) and Arg-348 (TM6) – these were 25.3 Å in the starting model for the active state (red) and 12.1 Å for the inactive state (black).

FIGURE 3

Snaps shots of the inactive (blue) and active (green) models taken at 285 ns and 325 ns respectively. Stick representations have oxygens as red, nitrogens as blue, and hydrogen bonds as red dashes. **A** Overlay of both models showing how Arg-190 and Gln-394 interact in the inactive state, whereas in the active state Arg-190 instead interacts with Glu-9* of GLP-1 (yellow), freeing Gln-394 to move and interact elsewhere. **B** Gln-394 (yellow) in the active state provides an N-cap for the cytoplasmic half of TM6 at the carbonyl oxygen of Pro-358. **C** Arg-380 in the active receptor points towards the centre of the helical bundle and interacts with both Glu-373 and Asp-15* of GLP-1. **D** Arg-380 in the inactive receptor points away from the helical bundle and interacts with Glu-373. The ligand-induced movement of Arg-380 may be responsible for rearranging ECL3.

FIGURE 4

The binding of GLP-1 (yellow) to the active GLP-1R model (snapshot at 325 ns). Stick representations have oxygens as red, nitrogens as blue, and hydrogen bonds as red dashes. The positively charged N-terminus of GLP-1 interacts with both Glu-364 and Glu-387, each of which additionally interact with Arg-310 and Lys-383 respectively. Since the N-terminus of GLP-1 is critical for the intrinsic activity, while the side chain of Arg-310 is essential for receptor activity, it may be that their ability to interact with Glu-364 in this manner is key to the conformational rearrangement of TM6.

Figure 1

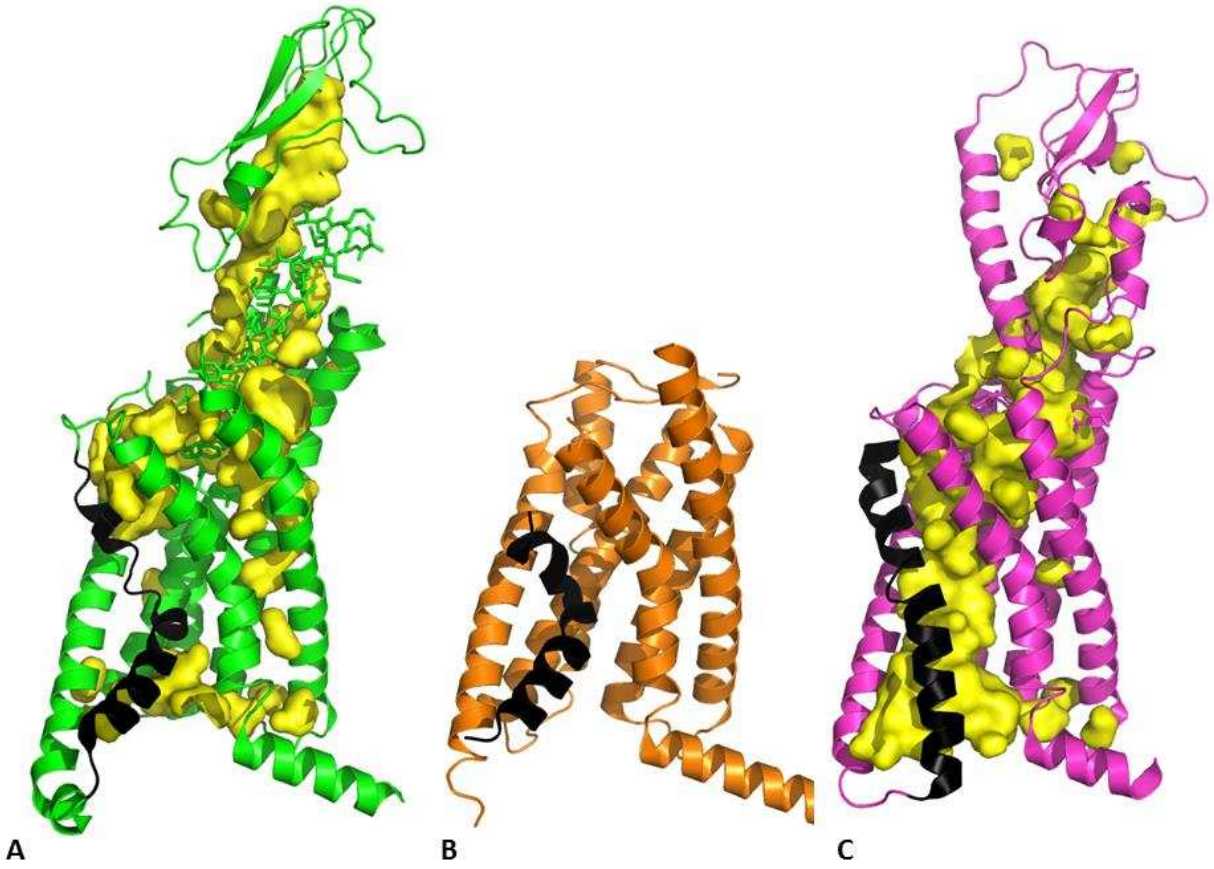


Figure 2

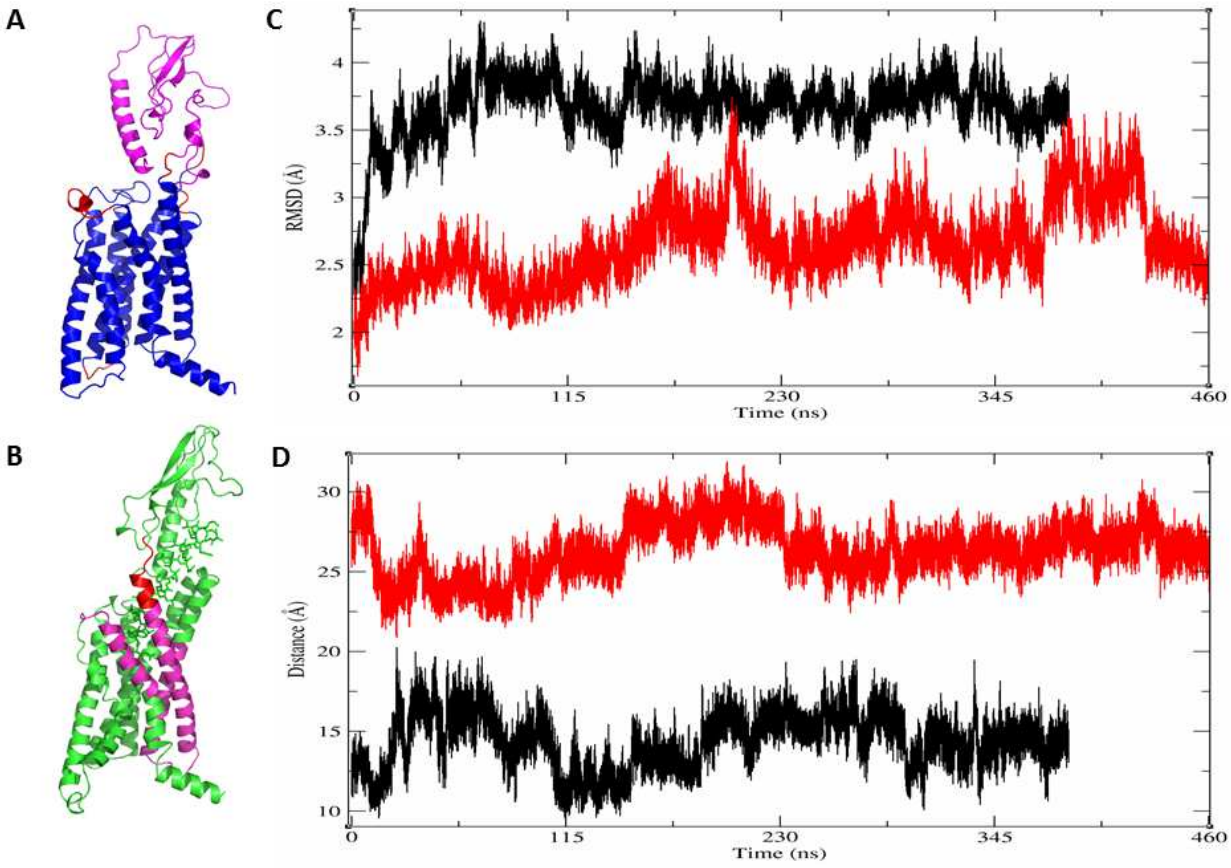


Figure 3

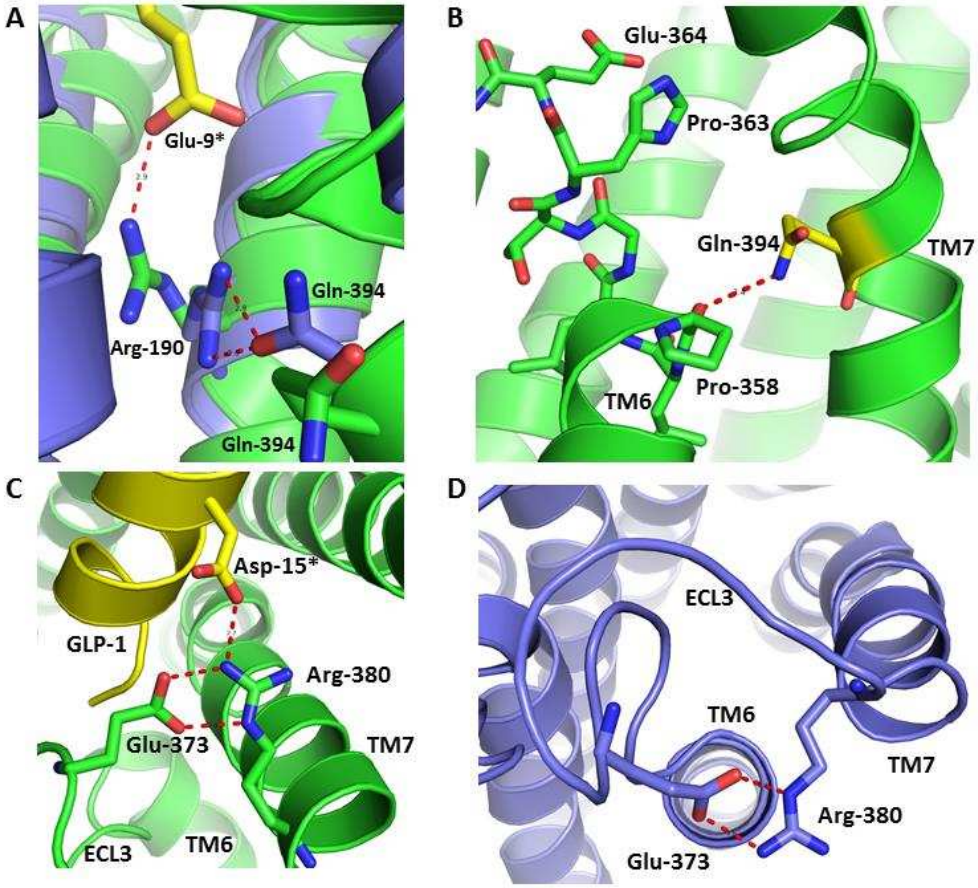
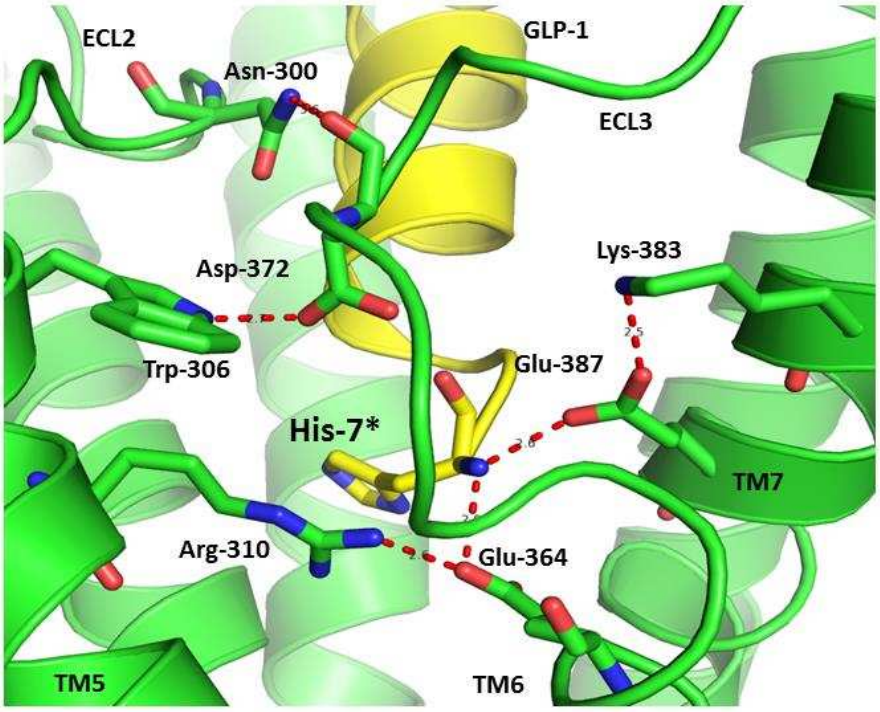

A
Glu-9*
Gln-394
Arg-190
Gln-394
B
Glu-364
Pro-363
Gln-394
TM7
TM6
Pro-358
C
Asp-15*
GLP-1
Arg-380
Glu-373
TM7
ECL3
TM6
D
ECL3
TM6
TM7
Arg-380
Glu-373

Figure 4

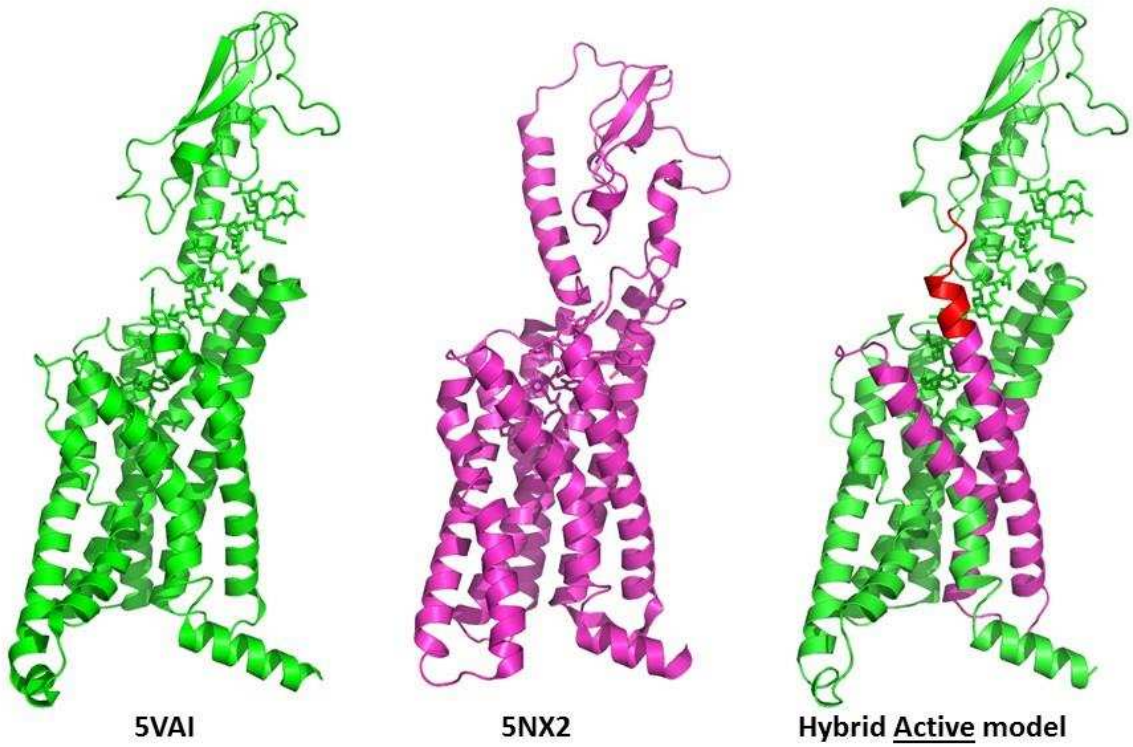

ECL2
Asn-300
GLP-1
ECL3
Asp-372
Lys-383
Trp-306
Glu-387
His-7*
TM7
Arg-310
Glu-364
TM5
TM6

Supplementary Information

A MECHANISM FOR AGONIST ACTIVATION OF THE GLUCAGON-LIKE PEPTIDE-1 (GLP-1) RECEPTOR THROUGH MODELLING & MOLECULAR DYNAMICS

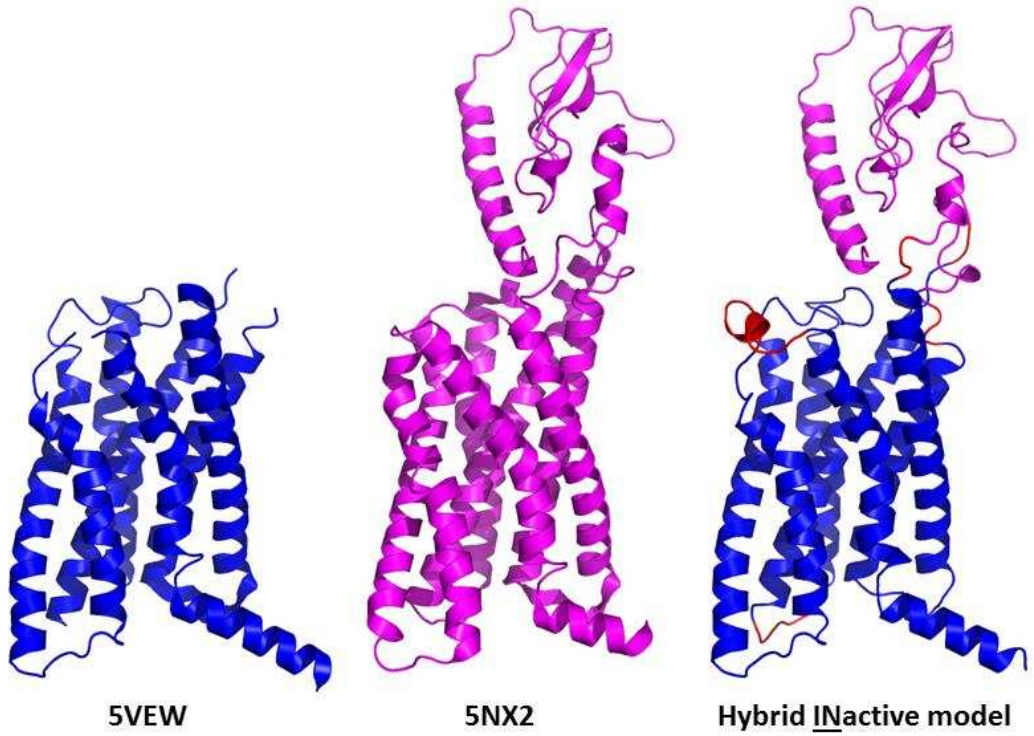
Carla Gómez Santiago [a], Emanuele Paci [b] and Dan Donnelly [a]

Figure S1



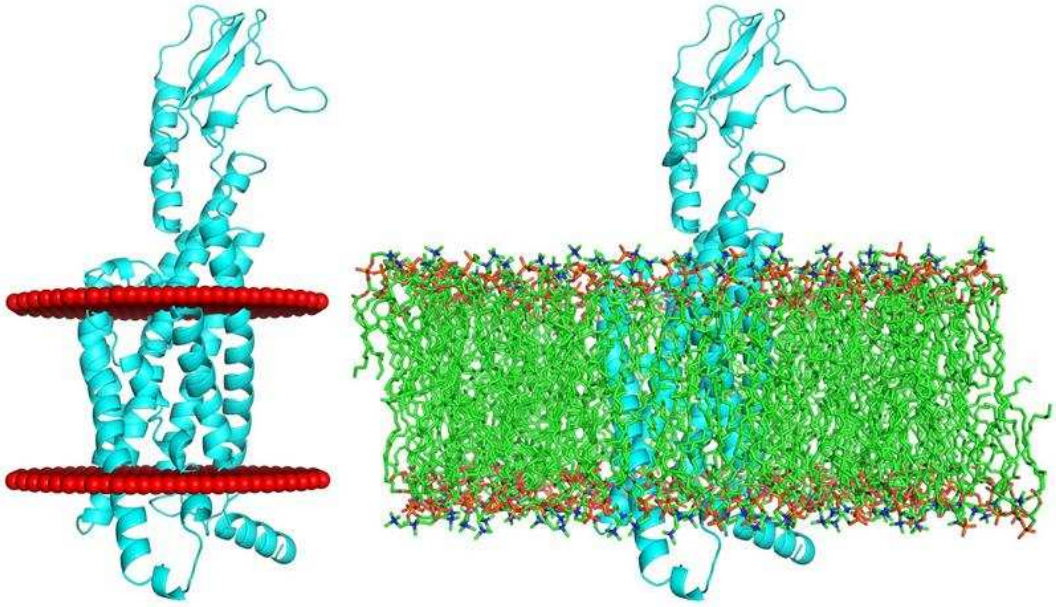
The active state model of human GLP-1R was built from the GLP-1-bound/G_s-coupled rabbit GLP-1R cryo-EM structure 5VAI (green) and the thermo-stabilised mutated human GLP-1R x-ray structure 5NX2 (magenta) as depicted above and as described in Methods. Regions missing from the template were built by SWISS-MODEL (red).

Figure S1B



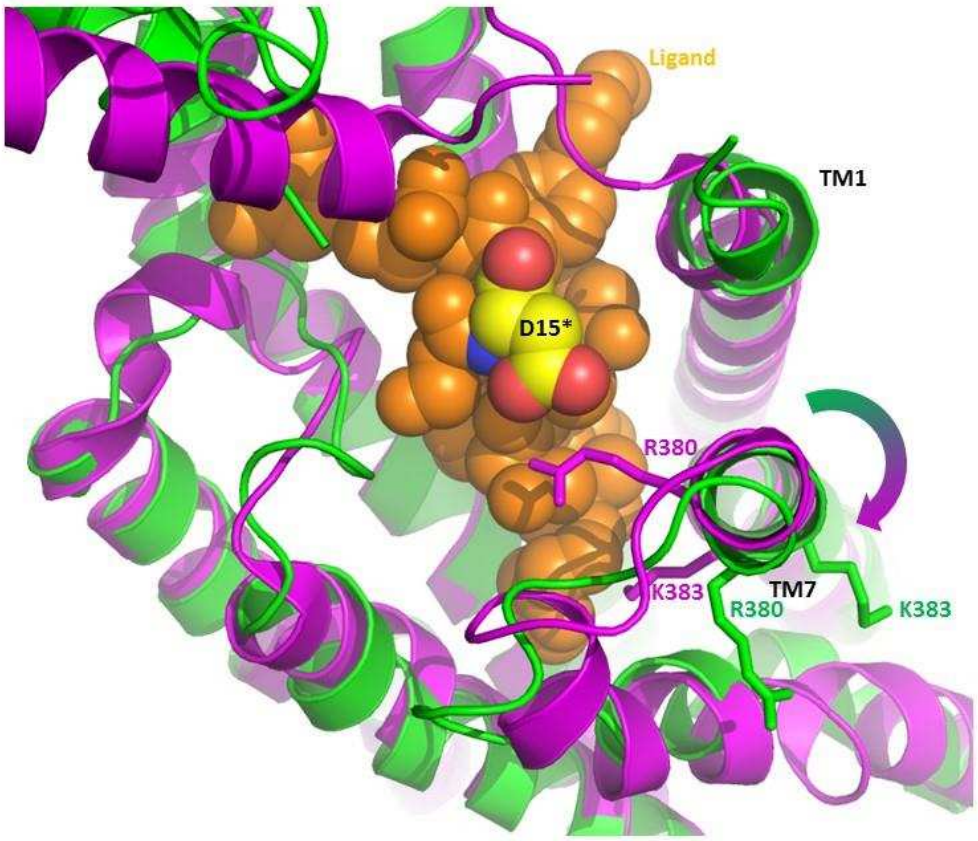
The inactive state model of human GLP-1R was built from thermo-stabilised inactive state human GLP-1R x-ray structure 5VEW (blue) and using the NTD from the thermo-stabilised mutated human GLP-1R x-ray structure 5NX2 (magenta) as depicted above and as described in Methods. Regions missing from the template were built by SWISS-MODEL (red).

Figure S2



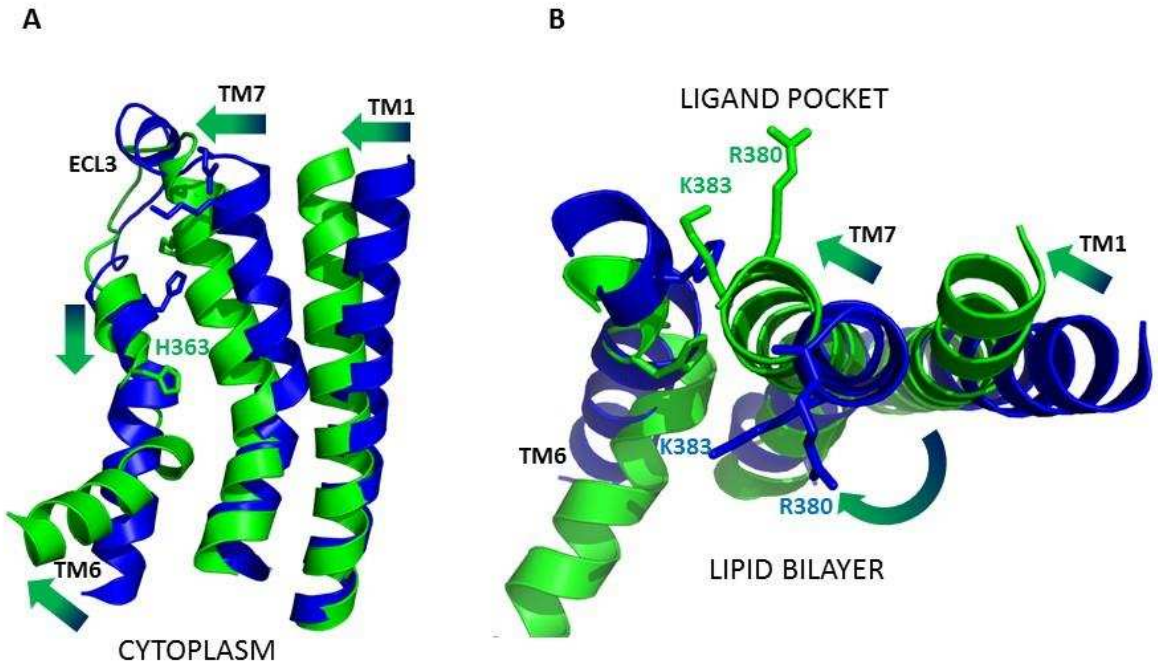
The active state model of human GLP-1R shown in ribbon form (cyan) with the hydrophobic bilayer boundaries as predicted by OMP [17] as red space filled dummy atoms (left) and the POPC bilayer as built by CHARMM-GUI [18] in stick form (right). The carbon atoms (green) depict the hydrophobic region of the bilayer, while the phosphate, oxygen and nitrogen atoms (orange, red, blue) form the phospholipid head groups just outside the OMP boundary. Hence there was very good correlation between the OMP lipid boundary prediction and the POPC bilayer built by CHARM-GUI.

Figure S3



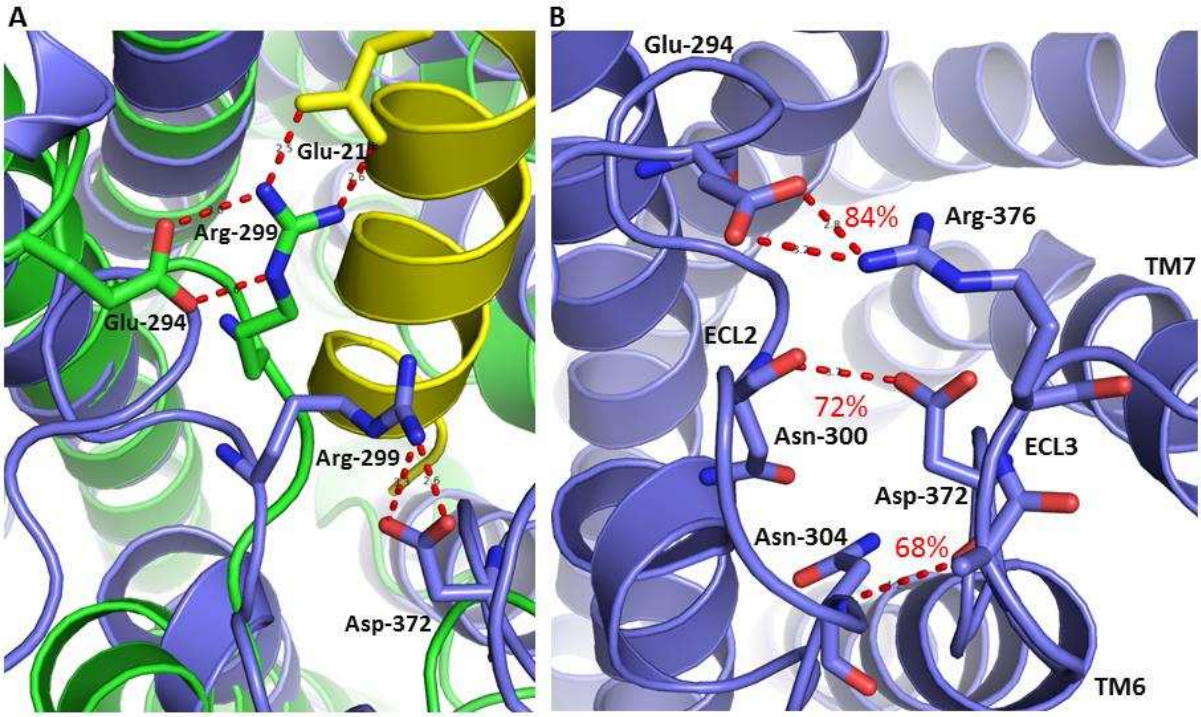
Comparison of 5NX2 (magenta with agonist as space-fill) and 5VAI (green). While Arg-380 and Lys-383 contact the ligand in 5NX2, as expected from mutagenesis data, in 5VAI they are displaced by 100° and face the lipid head-group region of the bilayer. Note that the Arg-380 side chain was built using the default conformation in PyMOL since the side chain is absent in the 5VAI co-ordinates.

Figure S4



Comparison of the starting models of the inactive (blue) and active (green) states. The expected movement at the cytoplasmic end of TM6 to enable G protein binding is coupled with the downwards movement of the top part of TM6 and a distortion below His-363 at Gly-361. TM7 bends at Gly-395 and the top part tilts towards TM6. It also twists to move Arg-380 and Lys-383 into the ligand pocket, as expected from mutagenesis data. TM1 also tilts about a hinge and moves towards TM7.

Figure S5



Snapshots of inactive (blue) and active (green) models taken at 285 ns and 325 ns respectively. Stick representations have oxygens as red, nitrogens as blue, and hydrogen bonds as red dashes. **A** Overlay of both models showing how in the inactive ECL2 residue Arg-299 interacts with ECL-3 residue Asp-372. In the active state, the ECL2-ECL3 interaction is replaced as Arg-299 interacts with Glu-292 and Glu-21* of GLP-1. **B** A number of other ECL2-ECL3 interactions are shown – all these are absent in the active state as ECL3 is rearranged due to the conformational change in TM6.

Figure S6

Snapshots of inactive (green) and active (cyan) models taken at 285 ns and 325 ns respectively. GLP-1 is shown as yellow. Stick representations have oxygens as red, nitrogens as blue, and hydrogen bonds as red dashes. Movements between sites in the inactive and active models are shown by arrows fading from green to cyan.

A & B Arg-380 and Glu-373 interact and move towards the ligand, with Arg-380 interacting with Asp-15*. Arg-299 also moves towards the ligand and interacts with Glu-21* (and Glu-294). The Arg-299 to Asp-372 interaction seen in the inactive conformation is broken, as well as numerous other ECL2-ECL3 interactions.

C & D Lys-383 and Glu-387 are distant in the inactive conformation but move together in the active state, with Glu-387 interacting directly with the N-terminal nitrogen of His-7*. The Arg-310 to Glu-364 interaction observed in the inactive state is reformed in the active state but Glu-364 moves towards the ligand and also interacts directly with the N-terminal nitrogen of His-7*.

E & F Arg-190 to Gln-394 interaction seen in the inactive state is broken in the active state since Arg-190 moves to interact with Glu-9* of GLP-1. Gln-394 is then able to hydrogen bond to the carbonyl oxygen of Pro-358, possibly helping to stabilise the active conformation in which the lower part of TM6 moves outwards to enable G_s to bind (**G**).

Figure S6

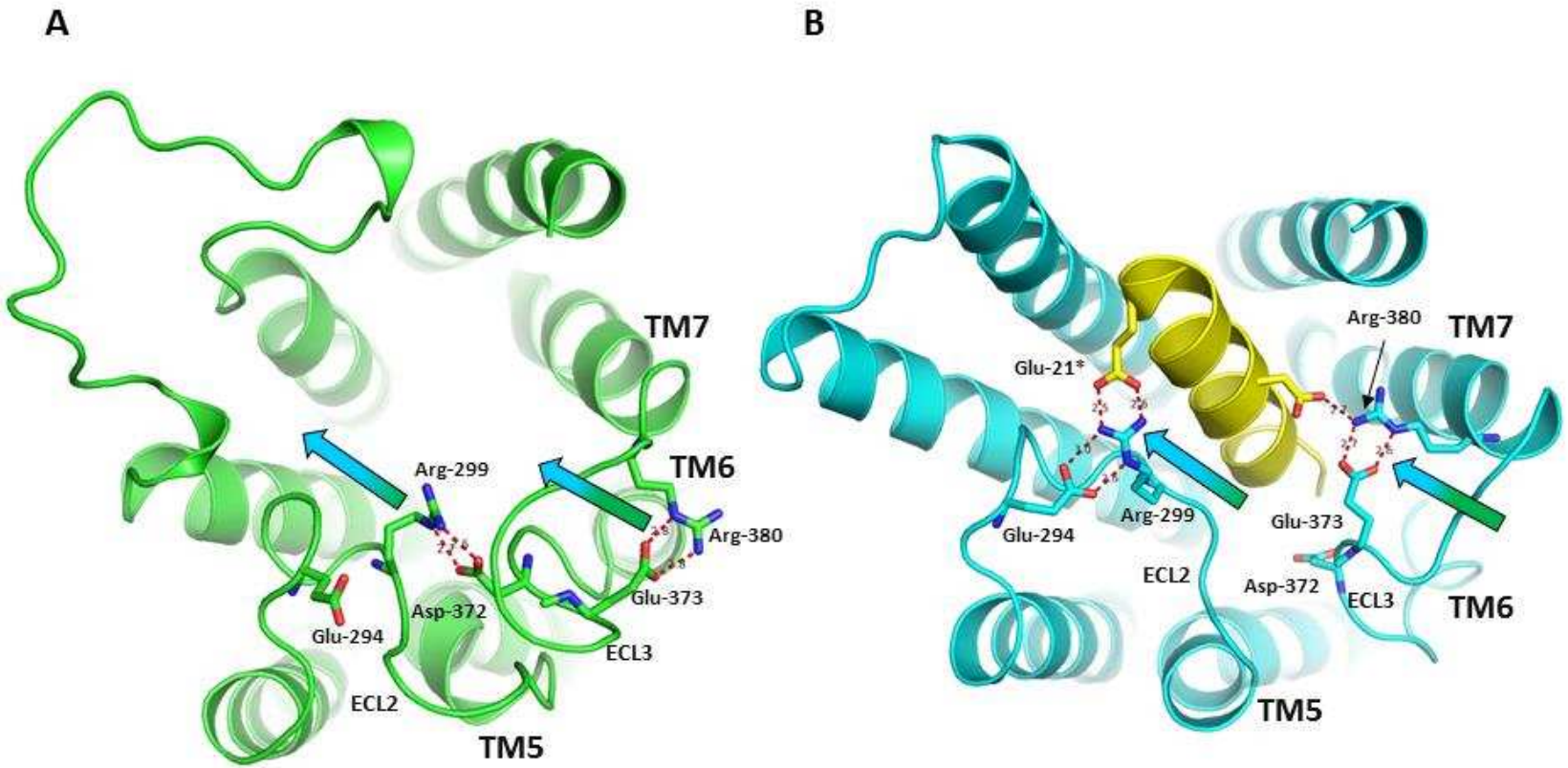

A
TM7
TM6
Arg-299
Arg-380
Glu-373
Asp-372
Glu-294
ECL3
ECL2
TM5
B
Arg-380
TM7
Glu-21*
Glu-294
Arg-299
Glu-373
ECL2
Asp-372
ECL3
TM6
TM5

Figure S6

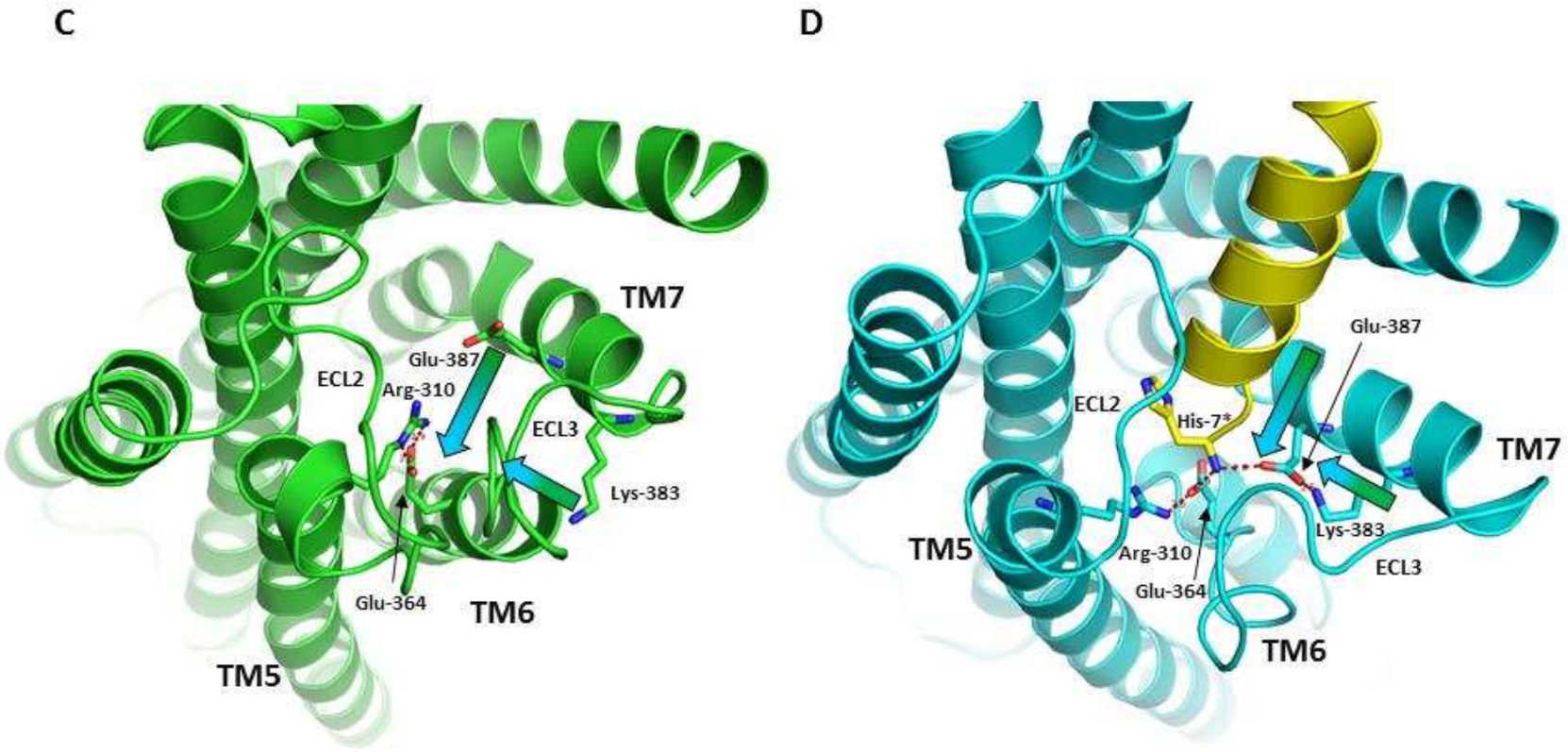

C
TM7
Glu-387
ECL2
Arg-310
ECL3
Lys-383
Glu-364
TM6
TM5
D
Glu-387
ECL2
His-7*
TM7
TM5
Lys-383
Arg-310
ECL3
Glu-364
TM6

Figure S6

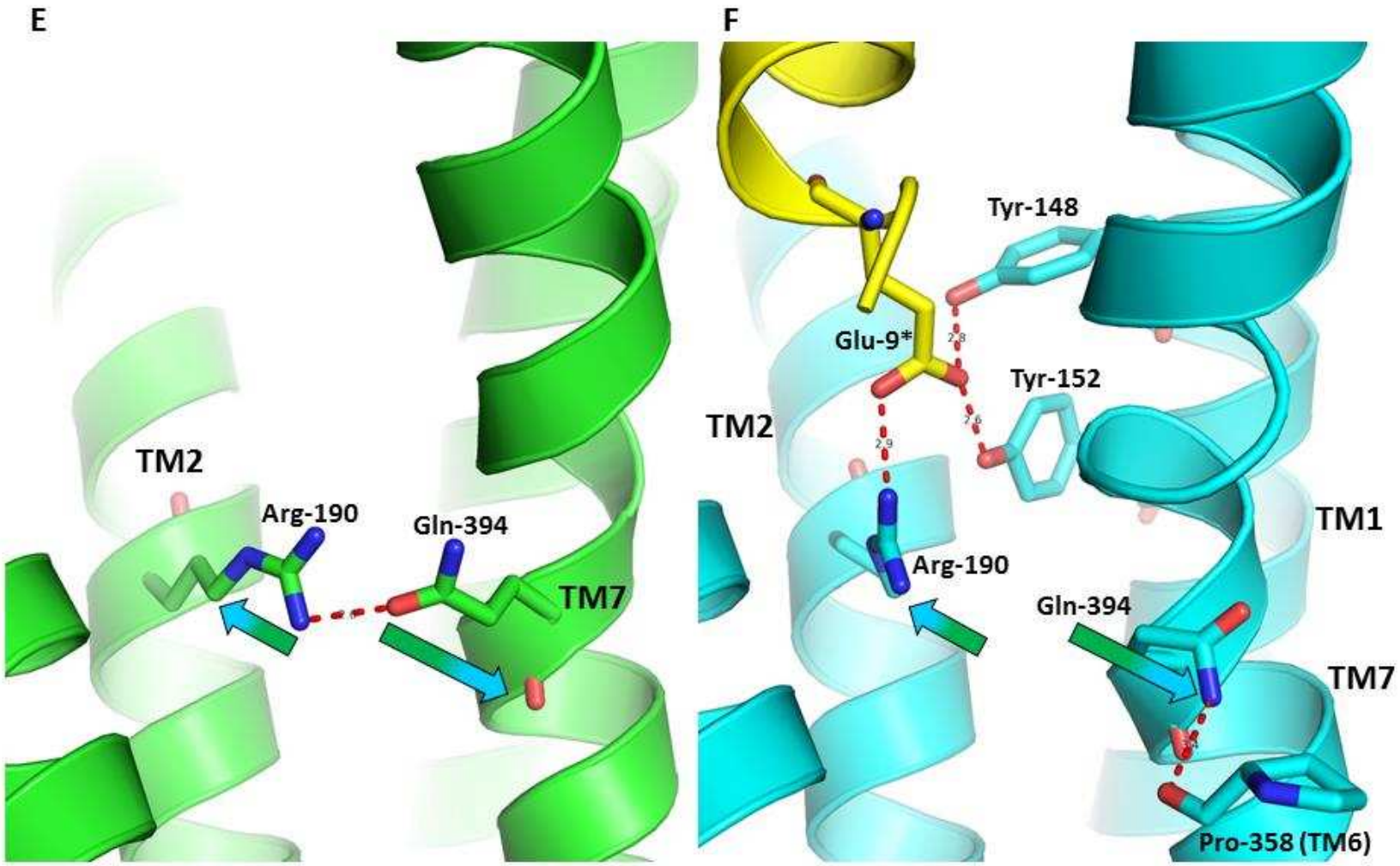

E
TM2
Arg-190
Gln-394
TM7
F
Tyr-148
Glu-9*
Tyr-152
TM2
TM1
Arg-190
Gln-394
TM7
Pro-358 (TM6)

Figure S6 G