

# Predicted structure of agonist-bound glucagon-like peptide 1 receptor, a class B G protein-coupled receptor

Andrea Kirkpatrick, Jiyoung Heo<sup>1</sup>, Ravinder Abrol<sup>2</sup>, and William A. Goddard III<sup>2</sup>

Materials and Process Simulation Center (139-74), California Institute of Technology, Pasadena, CA 91125

Contributed by William A. Goddard III, October 18, 2012 (sent for review July 9, 2012)

The glucagon-like peptide 1 receptor (GLP1R) is a G protein-coupled receptor (GPCR) involved in insulin synthesis and regulation; therefore, it is an important drug target for treatment of diabetes. However, GLP1R is a member of the class B1 family of GPCRs for which there are no experimental structures. To provide a structural basis for drug design and to probe class B GPCR activation, we predicted the transmembrane (TM) bundle structure of GLP1R bound to the peptide Exendin-4 (Exe4; a GLP1R agonist on the market for treating diabetes) using the MembStruk method for scanning TM bundle conformations. We used protein-protein docking methods to combine the TM bundle with the X-ray crystal structure of the 143-aa N terminus coupled to the Exe4 peptide. This complex was subjected to 28 ns of full-solvent, full-lipid molecular dynamics. We find 14 strong polar interactions of Exe4 with GLP1R, of which 8 interactions are in the TM bundle (2 interactions confirmed by mutation studies) and 6 interactions involve the N terminus (3 interactions found in the crystal structure). We also find 10 important hydrophobic interactions, of which 4 interactions are in the TM bundle (2 interactions confirmed by mutation studies) and 6 interactions are in the N terminus (6 interactions present in the crystal structure). Thus, our predicted structure agrees with available mutagenesis studies. We suggest a number of mutation experiments to further validate our predicted structure. The structure should be useful for guiding drug design and can provide a structural basis for understanding ligand binding and receptor activation of GLP1R and other class B1 GPCRs.

protein structure prediction | incretin receptors | peptide hormones

**G** protein-coupled receptors (GPCRs) are the largest family of integral membrane proteins within the human genome, and they are all characterized by seven transmembrane (TM) helices, with the N terminus on the extracellular (EC) side and the C terminus on the intracellular (IC). This family of proteins senses molecules outside of the cell and activates signal transduction pathways to cause cellular responses. Because of this vital role in cellular signaling networks, GPCRs are involved in many diseases, and they are the target of ~40% of all prescription pharmaceuticals on the market (1). Because of the importance of GPCRs as drug targets, it is vital to gain structural information for aiding in drug design. Unfortunately, GPCRs, like other membrane proteins, are difficult to crystallize. There are now X-ray crystal structures for more than 12 distinct receptors of the (at least) 800 human GPCRs (2–16). All of the crystallized receptors belong to the class A (rhodopsin-like) family of GPCRs. However, a phylogenetic analysis of GPCRs classifies them into different subfamilies: class A (rhodopsin-like), class B1 (secretin-like), class B2 (adhesion-like), class C (glutamate-like), and Frizzled/Taste2 (12).

Class B1 (secretin-like) GPCRs are activated by peptide hormones. The large ectodomain of these receptors interacts strongly with the C-terminal halves of their endogenous polypeptide agonists. The N terminus of their ligands putatively binds to the TM bundle and EC loops. The current structural understanding of ligand binding and activation of class B1 GPCRs comes from functional and ligand binding studies as well as the crystallized ectodomains of several class B1 GPCRs (17–19). These studies

have put forth speculations as to the plausible mechanisms for class B1 GPCR agonist binding and activation initiation, but in the absence of atomic-level structures of these receptors, it is difficult to understand, probe, and expand on these activation hypotheses.

In this work, we focus on the glucagon-like peptide 1 receptor (GLP1R). Activation of GLP1R by GLP1 stimulates the adenylyl cyclase pathway, which increases insulin synthesis and release of insulin in a glucose-dependent fashion (20). In addition, GLP1 reduces body weight by increasing satiety in the brain (21). Consequently, GLP1 would seem to be attractive for treating both type 2 diabetes and obesity. However, GLP1 is rapidly degraded by the serine protease dipeptidyl peptidase-IV in the body, resulting in its half-life of only 1–2 min (22, 23). Exendin-4 (Exe4), a 39-aa peptide isolated from the venom of the Gila monster, is a more stable analog of GLP1 with a half-life of 2.5 h in its marketed form (24–26). It has a 50% sequence homology with GLP1, and it is a full agonist with a stronger affinity and potency for GLP1R (27). Indeed, Exe4 is currently on the market for treatment of diabetes. Despite the success of Exe4 and its derivatives, there is still a need to develop small-molecule (nonpeptide) orally active agonists of GLP1R. This need is further supported by recent reports of oncogenic side effects of Exe4 (28). The process of novel drug design targeting GLP1R could be aided significantly if there was a structure of the full GLP1R bound to Exe4, which is the motivation of the research reported here. Our structure also provides testable hypotheses of GLP1R activation on ligand binding.

In the following sections, we present the predicted structure of the full membrane-bound GLP1R/Exe4 complex in the presence of water. The TM bundle was predicted using the MembStruk methodology (29). This bundle, which was attached to the GLP1R ectodomain crystallized with partial Exe4 (30), was inserted into a periodic membrane-water box and relaxed by molecular dynamics (MD). We find that this predicted structure is consistent with all available mutation data, and we suggest additional experimental tests to validate key aspects of our structure. We believe that this structural information presented should help the development of selective active small-molecule agonists for GLP1R and also aid in probing the activation of class B1 GPCRs.

## Methods

We have been developing methods for predicting the structures of GPCRs since the late 1990s. The earlier methodology, denoted as MembStruk, focused on sequential optimization of the seven TM helices starting from a homology template (29). More recently, we have developed a method, denoted as GENSeMBLE, that aims at a combinatorially complete set of helix rotations and tilts (31). The structure that we report here was built entirely using MembStruk several years ago, but it was not published. We applied our GENSeMBLE methodology to the older MembStruk structure, but we

Author contributions: A.K., J.H., R.A., and W.A.G. designed research; A.K. and J.H. performed research; and A.K., R.A., and W.A.G. wrote the paper.

The authors declare no conflict of interest.

<sup>1</sup>Present address: Biomedical Technology, Sangmyung University, Chungnam 330-720, Republic of Korea.

<sup>2</sup>To whom correspondence may be addressed. E-mail: abrol@wag.caltech.edu or wag@wag.caltech.edu.

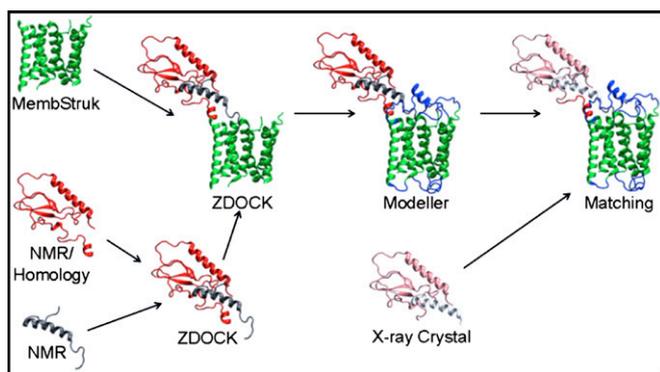
This article contains supporting information online at [www.pnas.org/lookup/suppl/doi:10.1073/pnas.1218051109/-DCSupplemental](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1218051109/-DCSupplemental).

obtained essentially the same packing of the seven TM helices (more details are in *SI Methods, section 8* and *Tables S1* and *S2*). Therefore, we decided to continue with our previous structure for the seven TM helix bundle and its connection to the N terminus, but we replaced the previously homology-built ectodomain of the N terminus of GLP1R (nGLP1R)/Exe4 part (where nGLP1R is the N terminus of GLP1R) with its crystal structure, which appeared recently (30).

A summary of the full procedures used to generate the GLP1R/Exe4 can be given in the following nine steps (the procedure is depicted in Fig. 1).

- i) The initial step in MembStruk (29) is to use comparative hydrophobicity analysis of GLP1R and related GPCRs to identify the seven likely TM domains and then, position ( $x, y, z$ ) the hydrophobic centers of these helices on a common plane with preselected tilts ( $\theta, \varphi$ ) and axis rotations ( $\eta$ ) based on some template (in this case, we used the predicted prostaglandin D2 (PGD2) receptor structure) (32).
- ii) This step is followed by sequential optimization of each TM domain by varying  $\eta$  over  $360^\circ$  in  $30^\circ$  increments and side chain optimization (SCREAM with modest minimization) (33) in a sequence that considers all TM domains multiple times.
- iii) Because of the large 143-aa N terminus of GLP1R, we originally built the structure for nGLP1R from homology to the NMR structure of the mouse Corticotropin-releasing factor (CRF) receptor 2 (Protein Data Bank ID code 2JND). This 131-aa region then underwent optimization of the side chains (34).
- iv) We used the ZDOCK procedure (35) to dock the NMR structure of the full Exe4 ligand (Protein Data Bank ID code 1JRJ) to the nGLP1R structure from step *iii* (36).
- v) We manually docked the nGLP1R/Exe4 complex from step *iv* to the TM bundle in such a way that the N terminus of the ligand could interact with the TM region from step *i*.
- vi) We connected the N terminus to the TM region (residues 131–145) and built the three EC and three IC loops using Modeller; this process was followed by SCREAM to position the side chains and then, energy minimization (37).
- vii) Then, we inserted the full GLP1R protein into a periodic membrane and water box ( $75 \times 75 \times 117 \text{ \AA}$ ), eliminating overlapping species to obtain a system with 61,119 atoms. This system was equilibrated at 300 K (first, the water and membrane and then, all atoms for 20 ns) using NAMD 2.6, with CHARMM22 charges for the protein and CHARMM27 charges for the lipids (38–40). The waters were modeled using the TIP3P potential function (41).
- viii) In the meantime, an X-ray crystal structure had appeared for nGLP1R bound to part of Exe4 (9–39; Protein Data Bank ID code 3C5T) (42). We matched this structure to our predicted structure and reoptimized (SCREAM, minimization, and MD for 18 ns).
- ix) After 18 ns of full-atom and full-solvent MD, we performed simulated annealing of the TM portion of the ligand binding site, and then, we carried 10 ns of full-atom and full-solvent MD at 300 K. A representative snapshot of the final 10 ns of MD was chosen for the discussion below.

Experimental information was not used during any of the above steps, except that information known about GLP1 loop structures was used to select



**Fig. 1.** Creating the GLP1R/Exe4 bundle. The steps used to create our GLP1R/Exe4 structure are depicted along with the methods used at each point. After the steps shown here, the entire complex was optimized through 28 ns of MD.

loops from Modeller in step *vi*. The specific information used is discussed in *SI Methods, section 4*.

## Results and Discussion

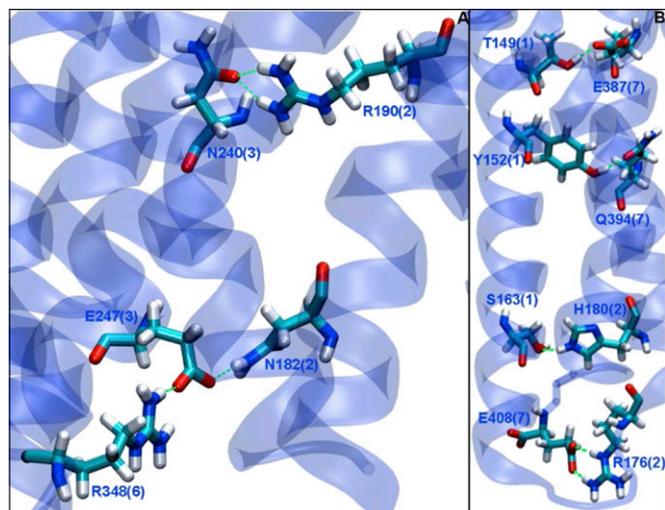
**Intraprotein Interactions.** For class A GPCRs, several conserved interhelical interactions, such as the 1-2-7 or 2-3-4 hydrogen bond networks, are present in most crystal structures. The amino acids involved in the 1-2-7 and 2-3-4 interactions are not conserved in class B GPCRs. However, we do find many interhelical interactions, some of which occur between residues conserved in class B1 GPCRs, which may be important for their structure and function. The GLP1R/Exe4 complex has 16 hydrogen bonds within its TM bundle (excluding helix-forming hydrogen bonds) or connecting loops, which is shown in *Table S3* and discussed in the following sections. The TM bundle interactions are pictured in Fig. 2.

**TM3-TM6 ionic lock.** We consider that the conserved E247(TM3)-R348(TM6) ionic lock is analogous to the R(3.50)-D/E(6.30) [using Ballesteros numbering of TM residues (43)] ionic lock of class A GPCRs known to stabilize the GPCR in an inactive form, although these donor and acceptor residue types are swapped in the class B1 version of the TM3-TM6 ionic lock (44). We find this ionic lock to be maintained through most of the dynamics. This interaction could play the role of maintaining the inactive form of class B1 GPCRs. Indeed, E247 is fully conserved among class B1 GPCRs, whereas R348 is either an R or K in all class B1 GPCRs. This hypothesis could be tested by mutations that break the ionic lock, which might lead to a constitutively active receptor (GLP1R does not exhibit constitutive activity) (45). This hypothesis is complicated in the light of GLP1R studies like the work by Heller et al. (46), which looked at the R348G mutant of GLP1R and found no activation for all concentrations of GLP1; in contrast, the work by Takhar et al. (47) found R348A to have no effect on binding or activation of GLP1. Perhaps, these large changes to alanine or glycine led to a modified TM bundle that changed the binding site, and/or this residue is critical for G-protein coupling and activation. A more subtle R348Q mutation would test the intricacies of this ionic lock in more detail.

Because the TM3-TM6 ionic lock is present in our structure, it is likely that our complex is not yet fully activated. However, our structure is bound to an agonist, and TM6 makes almost no interactions other than this ionic lock. Thus, it may be at least partially activated. This lack of TM6 interactions would allow TM6 (on breaking of the ionic lock) to immediately move away from TM3 to interact with the  $\alpha$ -subunit of the G protein, which it does in the active conformation of the G protein-bound  $\beta$ 2-adrenergic receptor structure (48).

**Coupling of TM2-TM3-TM6.** The 3-6 ionic lock is additionally coupled to TM2 by the conserved N182(TM2)-E247(TM3) hydrogen bond, resulting in a TM2-TM3-TM6 (2-3-6) hydrogen bond network [N182(TM2)-E247(TM3)-R348(TM6)]. This N182(TM2)-E247(TM3) has only minor fluctuations through the course of dynamics. Because N182 is an N, H, or Q in class B1 GPCRs, this 2-3-6 hydrogen bond network may be conserved in class B1 GPCRs. The 2-3-6 conserved network may be further stabilized by the similarly conserved R190(TM2)-N240(TM3) hydrogen bond, which is shown in conjunction with the 2-3-6 network in Fig. 2. We suggest that this interaction might be analogous to the 2.45(S/N/T)-3.42(S/N/T)-4.50(W) conserved interaction of class A GPCRs. Indeed, N240(TM3) is just 7 aa away from E247(TM3) compared with the 3.42 residue in class A GPCRs, which is 8 aa away from the 3.50 residue. We find the R190(TM2)-N240(TM3) interaction to be very constantly maintained throughout the dynamics. N240 is fully conserved among class B1 GPCRs, whereas R190 is R, K, or N, and therefore, this interaction is possible in any class B1 GPCR and may be a feature of this receptor class.

**TM1-TM2-TM7 interaction network.** The remaining two strong and conserved interactions between TM1 and TM7 [Y152(TM1)-Q394(TM7)] and TM2 and TM7 [R176(TM2)-E408(TM7)] might play an analogous role to the TM1.50(N)-TM2.50(D)-TM7.49(N) interactions conserved among class A GPCRs. Perhaps it is important to keep the TM1-TM2-TM7 (1-2-7) region rigid to control activation. Both interactions are stable during the course of the dynamics



**Fig. 2.** The (A) TM2-TM3-TM6 and (B) TM1-TM2-TM7 conserved hydrogen bond networks. (A) We believe that the E247(TM3)-R348(TM6) ionic lock may be associated with the unactivated GPCR structure (analogous to the R3.50-D/E6.30 interaction in class A GPCRs). TM3 is additionally coupled to TM2 through the conserved N182(TM2) to form a 2-3-6 interaction. This interaction is further stabilized by the R190(TM2)-N240(TM3) hydrogen bond, which may be analogous to the TM2.45(S/N/T)-3.42(S/N/T)-4.50(W) conserved interaction of class A GPCRs. (B) We also see several more transient couplings between TM1, -2, and -7, which are shown here. These four interactions help solidify the local structure of TMs 1, 2, and 7, and their disruption may be involved in activation.

simulation. These residues are also conserved in class B1 GPCRs: Y152 may be a Y or H, whereas Q394 may be a Q or H. R176 and E408 are fully conserved in all class B1 GPCRs. Note that the R176(TM2)-E408(TM7) interactions are located in the IC end of the TMs, and as such, they could also play a role in GPCR activation.

We find two additional interactions in the 1-2-7 region that involve nonconserved amino acids—T149(TM1)-E387(TM7) and H180(TM2)-S163(TM1)—that further stabilize the 1-2-7 coupling. We find that the T149(TM1)-E387(TM7) interaction is only transient during dynamics, being often mediated by the H1 residue of the ligand. Perhaps the agonist will eventually break this interaction as part of activation. The H180(TM2)-S163(TM1) interaction forms/breaks/reforms several times during the course of dynamics, indicating that it is less stable than the hydrogen bonds discussed previously. Despite their transience, these hydrogen bonds do help to stabilize the coupling of the 1-2-7 helices, and in conjunction with the with the more stable conserved interactions discussed earlier, they form a solid grouping of TMs 1, 2, and 7. The two conserved interactions discussed above, along with these two nonconserved interactions, are pictured in Fig. 2.

**EC loop couplings and N-terminal interactions.** The remaining four interhelical interactions are between the EC loops. The first three interactions are with D222(EC1) and adjoining EC2 residues R299, W297, and C296. In addition, there is a helical region present in EC1 from residue 215 to 225. It is the base of this helix that interacts with EC2. The final interloop hydrogen bond is between H374(EC3) and M303(EC2). The EC loops are clearly closely coupled, and they provide order to the flexible loop regions, which has been seen in other GPCR crystal structures (49). These stabilizing interactions would play a role in peptide binding, because they need to accommodate a peptide reaching from the N terminus past EC1 and into the TM bundle interior. Finally, GLP1R also has the TM3-EC2 disulfide coupling (C296-C226) conserved among class A GPCRs and class B1. There are no other Cys residues in the EC loops.

The overall architecture of the N terminus from the crystal structure remains stable during the course of the dynamics. There are still the three conserved disulfide bonds, two regions of antiparallel  $\beta$ -sheets, and an  $\alpha$ -helix adjacent to the ligand, which all

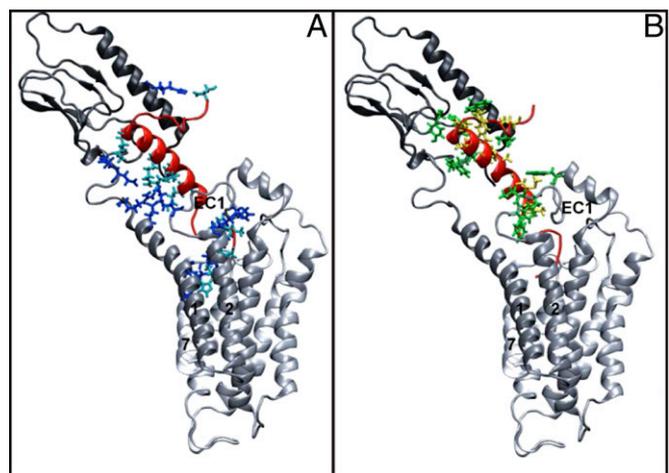
occur in all of the class B1 GPCR N-terminal crystal structures (50). No significant deviations from the crystal are observed.

Overall, our study of GLP1R shows that there are several conserved hydrogen bond networks that mimic those networks of class A GPCRs. The 3-6 ionic lock is similar to the lock of class A GPCRs, although it may play a somewhat different structural and functional role in G-protein activation. The 1-2, 1-7, and 2-7 networks mimic the 1-2-7 interaction motif of class A GPCRs. Finally, the loop structures, which may have direct bearing on ligand binding, are stabilized by several interloop interactions.

**Protein-Ligand Interactions. Structure overview.** The GLP1R/Exe4 binding site involves interactions throughout the N terminus, TM regions, and EC loops with the primary interactions occurring with the N terminus, TMs 1, 2, and 7, and EC1 (Fig. 3). We find Exe4 to be helical from residues 9–29. The strongest interactions with the N terminus include six polar interactions (three of which were present in the crystal structure) and six hydrophobic interactions (all were present in the crystal structure). The TM bundle features eight polar interactions (two of which were confirmed by mutation studies) and four hydrophobic interactions (two of which were confirmed by mutation studies). The full-energy analysis for these interactions is shown in Table S4. We will discuss the binding site in three parts: hydrogen bonds, nonpolar interactions, and detailed comparison with mutation data.

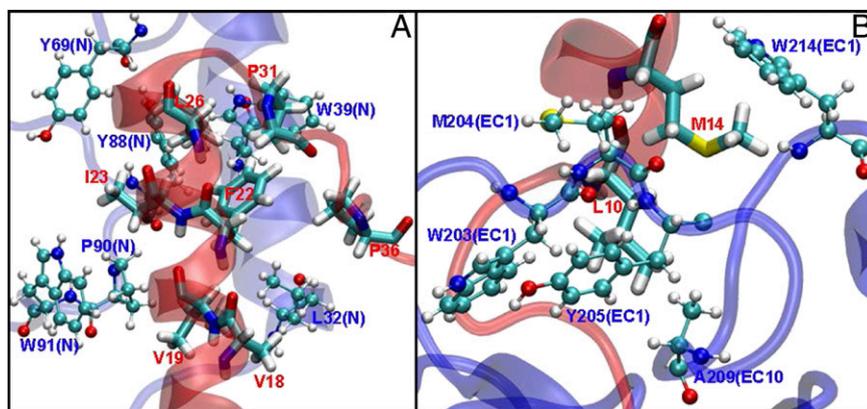
**Polar interactions.** We find 14 polar interactions (hydrogen bonds or salt bridges) between GLP1R and Exe4 (Fig. 4 and Table 1). There are eight polar interactions within the TM region, which reflects the primary areas of interaction between the ligand and TM bundle: TMs 1, 2, and 7 as well as EC1. The TM region polar interactions are particularly focused at the first few residues of the ligand—specifically H1 and E3 but also, T5 and F6. Our TM bundle interactions are also validated by site-directed mutagenesis studies for residues T149, E387, T391, and K197, which we find to interact with H1 and E3 of the ligand (14, 51) as shown in Table 1.

Our five N-terminal interactions include the two crystal salt bridges: E128(N)-R20 and E127(N)-K27 (42). These two residues' importance has also been shown through mutation studies on E127 and E128 (52). In addition, we find three very strong interactions of the N terminus, of which two interactions are in the flexible region between the structured N terminus and TM1 and the third interaction is at the C terminus of the ligand. The remaining crystal hydrogen bond between R121(N) and the backbone of K27 alternates during the MD from water-mediated



**Fig. 3.** Overview of the ligand binding site's hydrophobic (A) and hydrophilic (B) interactions. GLP1R is shown with a color transition from black to white as the protein goes from the N terminus to the C terminus. Exe4 is shown in red. Hydrophilic interactions are shown in A, with protein residues in blue and ligand residues in turquoise. Hydrophobic interactions are shown in B, with protein residues in green, and ligand residues in yellow.





**Fig. 5.** Hydrophobic interactions between GLP1R and Exe4 in the N terminus (A) and EC1 (B). Protein residues are shown in blue (CPK drawing method), whereas ligand residues are shown in red (licorice drawing method). We find 10 important hydrophobic interactions, of which 6 interactions are in the N terminus (A; all confirmed by X-ray crystal structure) and 4 interactions are in the TM bundle (B; of which 2 interactions have been confirmed by mutation studies).

F6 and required M204, Y205, D215, and R227 to be close to some part of Exe4).

Finally, we note that our structure preserved all interactions found in the crystal structure—both hydrophilic and hydrophobic—over the course of the dynamics. This finding provides additional validation of our structure, because these interactions would be expected to be stable. In addition, all residues indicated in the literature to be potentially important for binding or activation are found in our structure to point into the TM bundle, or they are otherwise accessible to the ligand. As such, they can either interact with the ligand itself or have structural effects. Overall, our GLP1R structure strongly agrees with the available experimental data, making it valuable for further structural and activation studies.

**Testing the Predicted Binding Site.** Our predicted structure of GLP1R and its binding site for Exe4 suggest many mutation studies for its further validation. Indeed, we constructed structures for 14 such mutations (Table S7 shows their effects on energies). The procedure was to use SCREAM to introduce the desired mutation and then minimize the protein to 0.5 kcal/mol per angstrom rms force (33). These calculations assumed that the overall backbone structure remains intact. Of 14 mutations, 2 mutations are predicted to improve binding, whereas 12 mutations are predicted to decrease binding.

The first set of mutations was chosen with the goal of validating our predictions of the strongest interactions between GLP1R and Exe4. Ten cases were aimed at disrupting the binding site by breaking interactions discussed previously: R134A/Q, K202A/Q, and K383A/Q break salt bridges or hydrogen bonds, whereas W203N/Y and W214N/Y disrupt the EC1–ligand hydrophobic interactions. We also suggest two ligand mutations that would decrease binding: K12A and M14Q. The K12A mutation would lose the E139 salt bridge, whereas the M14Q would disrupt the hydrophobic network that the M14 has in the EC1 area. In each of the 12 cases, our predicted change in binding agrees with expectation. Finally, we predicted two mutations to improve ligand binding. The first mutation was S11W of Exe4, which allows a new interaction to be formed with W214 on the receptor. The second change is Q213K of GLP1R, which forms a new hydrogen bond with D9 on Exe4. Both of these changes improve our predicted interaction energies between the receptor and ligand.

**Discussion of Ligand Binding and Protein Activation.** Our GLP1R/Exe4 structure suggests several general features of ligand binding to GLP1R and potentially, class B1 proteins as a whole. One feature is that we find that the TM region polar interactions are particularly focused at the first few residues of the ligand—specifically, H1 and E3 but also, T5 and F6. This finding is in accordance with the known importance of the N terminus (specifically residues 1–7) of class B1 agonists for protein activation (50). In the specific case of Exe4, if the ligand is truncated by eight residues, it becomes a competitive antagonist for GLP1R, because it can still bind the receptor but no longer cause activation (27).

Next, we find that the binding pocket of Exe4 shows strong polar and hydrophobic interactions with EC1. Experimental studies showed that mutations of EC2 residues to alanine dramatically decreased binding of GLP1 but had no effect on the binding of Exe4 (57), consistent with our structure. We suggest that the reason for the importance of the loops in the peptide binding is to align the ligand in the correct conformation for TM bundle entry. In the two-domain model of class B1 GPCR protein binding, the N-terminal ectodomain plays the role of recognizing the ligand and supports the initial binding (58). We believe that the next step is for the flexible N terminus/ligand complex to align itself to the TM bundle by loop interactions followed by final insertion of the head region of the ligand into the TM bundle itself.

Also, we note that the flexible N terminus of the ligand is nestled in the TM1-TM2-TM7 binding pocket, which leaves the TM3-TM4-TM6 region largely open, making this area available for binding of small molecules serving as ago-allosteric modulators (59). In addition, the residues of the ligand inserted themselves between hydrogen bonds of the apo GLP1R [for example, T149(TM1)-H1-E387(TM7)]. This finding suggests that part of the effect of Exe4 binding may be to break some of the TM1-TM2-TM7 strong interactions, giving the structure the flexibility to achieve its active conformation.

Any discussion of GPCR activation would be incomplete without mention of the TM3-TM6 ionic lock, of which we find a variation in our structure. Instead of the conserved R(3.50)-D/E(6.30) ionic lock of class A GPCRs, we find an analogous conserved E247(TM3)-R348(TM6) ionic lock. Breaking this interaction may be crucial for GLP1R activation. To test this hypothesis, one could mutate one of the charged residues to a polar residue, such as glutamine; and therefore, the overall hydrophilicity of the region could be preserved, but the interaction would be broken.

Finally, it has been suggested that an N-terminal helix capping motif of a peptide agonist is a key element underlying class B1 GPCR activation (18). This structural feature is theorized to consist of a hydrophobic interaction between residues 6 and 10, and it is stabilized by a salt bridge between residues 7 and 10 of the ligand. The result of these interactions is that the ligand forms an L-shape at its N terminus. Although we do not see the exact 7–10 and 6–10 interactions—instead, we find that residues Phe6 and Ser8 form a backbone hydrogen bond—this alternate interaction causes the ligand to adopt a slightly more loose L-conformation. This structural constraint may be important for the rational drug design of peptide agonists targeting class B1 receptors.

## Conclusion

We present here the predicted TM bundle for GLP1R (residues 146–408), which we combined with the crystal structure for the N terminus (residues 28–145) and Exe4. The resulting structure was then equilibrated in an explicit membrane and water environment for 28 ns. We find strong agreement with available experimental data, most of which played no role in the predictions. We suggest 14 mutations to provide strong tests of our predicted binding site. This structure can now form the basis for the

rational design of other peptide ligands and the greatly needed small-molecule ligands.

The model that we present here can be used to explore the mechanisms of class B1 GPCR binding and activation (52). In addition, we expect that this structure will provide a basis for the design and optimization of new small-molecule ligands that bind selectively and specifically to GLP1R.

Finally, one of the grand challenges in understanding GPCRs is to elucidate the mechanism of activation. This study does not address this issue directly; however, we do identify a TM3-TM6

ionic lock that is conserved across class B1 GPCRs that we believe may play a very similar or more complex role in GLP1R activation compared with the TM3-TM6 ionic lock conserved across a subset of class A GPCRs. Mutation studies on this ionic lock will help test such speculations and provide structural signatures of active and inactive receptor conformations.

**ACKNOWLEDGMENTS.** The MembStruk studies by J.H. were funded by Allozyne (Ken Grabstein). The final optimization of the structure and analysis (A.K.) was funded by Sanofi-Aventis (Ken Wertman).

- Filmore D (2004) It's a GPCR world. *Modern Drug Discovery* 7(11):24–28.
- Palczewski K, et al. (2000) Crystal structure of rhodopsin: A G protein-coupled receptor. *Science* 289(5480):739–745.
- Cherezov V, et al. (2007) High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor. *Science* 318(5854):1258–1265.
- Warne T, et al. (2008) Structure of a beta1-adrenergic G-protein-coupled receptor. *Nature* 454(7203):486–491.
- Jaakola VP, et al. (2008) The 2.6 angstrom crystal structure of a human A2A adenosine receptor bound to an antagonist. *Science* 322(5905):1211–1217.
- Chien EY, et al. (2010) Structure of the human dopamine D3 receptor in complex with a D2/D3 selective antagonist. *Science* 330(6007):1091–1095.
- Wu B, et al. (2010) Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists. *Science* 330(6007):1066–1071.
- Shimamura T, et al. (2011) Structure of the human histamine H1 receptor complex with doxepin. *Nature* 475(7354):65–70.
- Haga K, et al. (2012) Structure of the human M2 muscarinic acetylcholine receptor bound to an antagonist. *Nature* 482(7386):547–551.
- Kruse AC, et al. (2012) Structure and dynamics of the M3 muscarinic acetylcholine receptor. *Nature* 482(7386):552–556.
- Hanson MA, et al. (2012) Crystal structure of a lipid G protein-coupled receptor. *Science* 335(6070):851–855.
- Lagerström MC, Schiöth HB (2008) Structural diversity of G protein-coupled receptors and significance for drug discovery. *Nat Rev Drug Discov* 7(4):339–357.
- Wu HX, et al. (2012) Structure of the human  $\kappa$ -opioid receptor in complex with JDTic. *Nature* 485(7398):327–332.
- Manglik A, et al. (2012) Crystal structure of the  $\mu$ -opioid receptor bound to a morphinan antagonist. *Nature* 485(7398):321–326.
- Granier S, et al. (2012) Structure of the  $\delta$ -opioid receptor bound to naltrindole. *Nature* 485(7398):400–404.
- Thompson AA, et al. (2012) Structure of the nociceptin/orphanin FQ receptor in complex with a peptide mimetic. *Nature* 485(7398):395–399.
- Parthier C, Reetz-Runge S, Rudolph R, Stubbs MT (2009) Passing the baton in class B GPCRs: Peptide hormone activation via helix induction? *Trends Biochem Sci* 34(6):303–310.
- Neumann JM, et al. (2008) Class-B GPCR activation: Is ligand helix-capping the key? *Trends Biochem Sci* 33(7):314–319.
- Pal K, Melcher K, Xu HE (2012) Structure and mechanism for recognition of peptide hormones by Class B G-protein-coupled receptors. *Acta Pharmacol Sin* 33(3):300–311.
- Schmidt WE, Siegel EG, Creutzfeldt W (1985) Glucagon-like peptide-1 but not glucagon-like peptide-2 stimulates insulin release from isolated rat pancreatic islets. *Diabetologia* 28(9):704–707.
- Kieffer TJ, Habener JF (1999) The glucagon-like peptides. *Endocr Rev* 20(6):876–913.
- Kieffer TJ, McIntosh CH, Pederson RA (1995) Degradation of glucose-dependent insulinotropic polypeptide and truncated glucagon-like peptide 1 in vitro and in vivo by dipeptidyl peptidase IV. *Endocrinology* 136(8):3585–3596.
- Mentlein R, Gallwitz B, Schmidt WE (1993) Dipeptidyl-peptidase IV hydrolyses gastric inhibitory polypeptide, glucagon-like peptide-1(7–36)amide, peptide histidine methionine and is responsible for their degradation in human serum. *Eur J Biochem* 214(3):829–835.
- Eng J, Kleinman WA, Singh L, Singh G, Raufman JP (1992) Isolation and characterization of exendin-4, an exendin-3 analogue, from *Heloderma suspectum* venom. Further evidence for an exendin receptor on dispersed acini from guinea pig pancreas. *J Biol Chem* 267(11):7402–7405.
- Parkes DG, Pittner R, Jodka C, Smith P, Young A (2001) Insulinotropic actions of exendin-4 and glucagon-like peptide-1 in vivo and in vitro. *Metabolism* 50(5):583–589.
- Bray GM (2006) Exenatide. *Am J Health Syst Pharm* 63(5):411–418.
- Göke R, et al. (1993) Exendin-4 is a high potency agonist and truncated exendin-(9–39)-amide an antagonist at the glucagon-like peptide 1-(7–36)-amide receptor of insulin-secreting beta-cells. *J Biol Chem* 268(26):19650–19655.
- Elashoff M, Matveyenko AV, Gier B, Elashoff R, Butler PC (2011) Pancreatitis, pancreatic, and thyroid cancer with glucagon-like peptide-1-based therapies. *Gastroenterology* 141(1):150–156.
- Vaidehi N, et al. (2002) Prediction of structure and function of G protein-coupled receptors. *Proc Natl Acad Sci USA* 99(20):12622–12627.
- Runge S, Thøgersen H, Madsen K, Lau J, Rudolph R (2008) Crystal structure of the ligand-bound glucagon-like peptide-1 receptor extracellular domain. *J Biol Chem* 283(17):11340–11347.
- Abrol R, Griffith AR, Bray JK, Goddard WA, 3rd (2012) Structure prediction of G protein-coupled receptors and their ensemble of functionally important conformations. *Methods Mol Biol* 914:237–254.
- Li Y, et al. (2007) Prediction of the 3D structure and dynamics of human DP G-protein coupled receptor bound to an agonist and an antagonist. *J Am Chem Soc* 129(35):10720–10731.
- Kam VWT, Goddard WA (2008) Flat-bottom strategy for improved accuracy in protein side-chain placements. *J Chem Theory Comput* 4(12):2160–2169.
- Grace CR, et al. (2007) Structure of the N-terminal domain of a type B1 G protein-coupled receptor in complex with a peptide ligand. *Proc Natl Acad Sci USA* 104(12):4858–4863.
- Chen R, Li L, Weng ZP (2003) ZDOCK: An initial-stage protein-docking algorithm. *Proteins* 52(1):80–87.
- Neidigh JW, Fesinmeyer RM, Prickett KS, Andersen NH (2001) Exendin-4 and glucagon-like-peptide-1: NMR structural comparisons in the solution and micelle-associated states. *Biochemistry* 40(44):13188–13200.
- Eswar N, et al. (2006) Comparative protein structure modeling using Modeller. *Current Protocols in Bioinformatics* (Wiley, New York), Unit 5.6.
- Phillips JC, et al. (2005) Scalable molecular dynamics with NAMD. *J Comput Chem* 26(16):1781–1802.
- MacKerell AD, et al. (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J Phys Chem B* 102(18):3586–3616.
- Brooks BR, et al. (2009) CHARMM: The biomolecular simulation program. *J Comput Chem* 30(10):1545–1614.
- Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) Comparison of simple potential functions for simulating liquid water. *J Chem Phys* 79(2):926–935.
- Runge VM (2008) Notes on "Characteristics of gadolinium-DTPA complex: A potential NMR contrast agent" *AJR Am J Roentgenol* 190(6):1433–1434.
- Ballesteros A, Weinstein H (1995) Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G protein-coupled receptors. *Methods Neurosci* 25:366–428.
- Vogel R, et al. (2008) Functional role of the "ionic lock"—an interhelical hydrogen-bond network in family A heptahelical receptors. *J Mol Biol* 380(4):648–655.
- Fortin JP, Schroeder JC, Zhu Y, Beinborn M, Kopin AS (2010) Pharmacological characterization of human incretin receptor missense variants. *J Pharmacol Exp Ther* 332(1):274–280.
- Heller RS, Kieffer TJ, Habener JF (1996) Point mutations in the first and third intracellular loops of the glucagon-like peptide-1 receptor alter intracellular signaling. *Biochem Biophys Res Commun* 223(3):624–632.
- Takhar S, et al. (1996) The third cytoplasmic domain of the GLP-1[7–36 amide] receptor is required for coupling to the adenylyl cyclase system. *Endocrinology* 137(5):2175–2178.
- Rasmussen SG, et al. (2011) Crystal structure of the  $\beta$ 2 adrenergic receptor-Gs protein complex. *Nature* 477(7366):549–555.
- Katritch V, Cherezov V, Stevens RC (2012) Diversity and modularity of G protein-coupled receptor structures. *Trends Pharmacol Sci* 33(1):17–27.
- Couvaineau A, Laburthe M (2012) The family B1 GPCR: Structural aspects and interaction with accessory proteins. *Curr Drug Targets* 13(1):103–115.
- Beinborn M, Worrall CI, McBride EW, Kopin AS (2005) A human glucagon-like peptide-1 receptor polymorphism results in reduced agonist responsiveness. *Regul Pept* 130(1–2):1–6.
- Coopman K, et al. (2011) Residues within the transmembrane domain of the glucagon-like peptide-1 receptor involved in ligand binding and receptor activation: Modelling the ligand-bound receptor. *Mol Endocrinol* 25(10):1804–1818.
- Underwood CR, et al. (2010) Crystal structure of glucagon-like peptide-1 in complex with the extracellular domain of the glucagon-like peptide-1 receptor. *J Biol Chem* 285(11):723–730.
- López de Maturana R, Treece-Birch J, Abidi F, Findlay JB, Donnelly D (2004) Met-204 and Tyr-205 are together important for binding GLP-1 receptor agonists but not their N-terminally truncated analogues. *Protein Pept Lett* 11(1):15–22.
- Al-Sabah S, Donnelly D (2003) 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(2):339–346.
- López de Maturana R, Donnelly D (2002) The glucagon-like peptide-1 receptor binding site for the N-terminus of GLP-1 requires polarity at Asp198 rather than negative charge. *FEBS Lett* 530(1–3):244–248.
- Koole C, et al. (2012) Second extracellular loop of human glucagon-like peptide-1 receptor (GLP-1R) has a critical role in GLP-1 peptide binding and receptor activation. *J Biol Chem* 287(6):3642–3658.
- Hoare SR (2005) Mechanisms of peptide and nonpeptide ligand binding to Class B G-protein-coupled receptors. *Drug Discov Today* 10(6):417–427.
- Knudsen LB, et al. (2007) Small-molecule agonists for the glucagon-like peptide 1 receptor. *Proc Natl Acad Sci USA* 104(3):937–942.