STRUCTURAL BIOLOGY

Structural basis for the ligand recognition and signaling of free fatty acid receptors

Xuan Zhang[1], Abdul-Akim Guseinov[2]†, Laura Jenkins[3]†, Kunpeng Li[4], Irina G. Tikhonova[2]*, Graeme Milligan[3]*, Cheng Zhang[1]*

Free fatty acid receptors 1 to 4 (FFA1 to FFA4) are class A G protein–coupled receptors (GPCRs). FFA1 to FFA3 share substantial sequence similarity, whereas FFA4 is unrelated. However, FFA1 and FFA4 are activated by long-chain fatty acids, while FFA2 and FFA3 respond to short-chain fatty acids generated by intestinal microbiota. FFA1, FFA2, and FFA4 are potential drug targets for metabolic and inflammatory conditions. Here, we determined the active structures of FFA1 and FFA4 bound to docosahexaenoic acid, FFA4 bound to the synthetic agonist TUG-891, and butyrate-bound FFA2, each complexed with an engineered heterotrimeric G_q protein (miniG_q), by cryo–electron microscopy. Together with computational simulations and mutagenesis studies, we elucidated the similarities and differences in the binding modes of fatty acid ligands to their respective GPCRs. Our findings unveiled distinct mechanisms of receptor activation and G protein coupling. We anticipate that these outcomes will facilitate structure-based drug development and underpin future research on this group of GPCRs.

INTRODUCTION

Free fatty acids are bioactive lipids comprising a carboxylic acid head group and an aliphatic hydrocarbon chain with various lengths. In humans, and many other species, they can activate a group of G protein–coupled receptors (GPCRs) including free fatty acid receptors 1 to 4 (FFA1 to FFA4 receptors) and GPR84 to regulate metabolic homeostasis and immunity (*1*). Among them, FFA1 (GPR40) and FFA4 (GPR120) mainly sense long-chain fatty acids (LCFAs) with more than 12 carbons, while FFA2 (GPR43) and FFA3 (GPR41) primarily sense short-chain fatty acids (SCFAs) with less than six carbons (*1*, *2*). Representative LCFA ligands of FFA1 and FFA4 include ω-3 and ω-6 polyunsaturated fatty acids (PUFAs) (*3*, *4*). Meanwhile, SCFA ligands of FFA2 and FFA3 are mainly produced in the gut as products of microbiota-mediated fermentation and include acetate, propionate, and butyrate (*4*, *5*).

FFAs play critical roles in both immunity and metabolism (*1*, *2*, *5*). FFA1 signaling induced by LCFAs in pancreatic β cells can facilitate insulin secretion (*6*), making it a promising drug target for type 2 diabetes mellitus (T2D) (*7*, *8*). Although the FFA1 selective agonist TAK-875, also named fasiglifam, exhibited promising antidiabetic effects in clinical studies, it failed in phase 3 trials due to liver toxicity (*9*). However, other FFA1 activators are still being pursued for the treatment of T2D (*7*). FFA4, which has been described as the ω-3 PUFA receptor, is highly expressed in adipose tissue and macrophages. It mediates anti-inflammatory and other beneficial effects of ω-3 PUFAs such as docosahexaenoic acid (DHA) in those tissues and cells (*1*, *10*). FFA4 is also considered as an emerging drug target for diabetes (*11*). FFA4 selective or FFA1/FFA4 dual agonists (*12*–*14*) may hold the promise of becoming a new class of antidiabetic drugs with additional anti-inflammatory benefits (*11*, *14*). In addition to their functions in metabolism and immunity, both FFA1 and FFA4, particularly FFA4, have been suggested to function as lipid taste receptors (*15*). On the other hand, FFA2 and FFA3 are expressed in adipocytes and a range of immune cells. Their unique ligand preference of SCFAs produced by the fermentation of dietary fiber in the lower gut has led to intensive research on their roles at the interface of host and gut microbiota (*16*, *17*). Previous studies suggested that many beneficial effects of gut microbiota on the host, including the resolution of inflammation (*18*), suppression of fat accumulation (*19*), and protection from viral and bacterial pathogens (*20*, *21*), are mainly mediated by the SCFA-FFA2 signaling axis. Therefore, FFA2 and FFA3, especially FFA2, are considered as promising therapeutic targets for metabolic disorders, including obesity and diabetes and inflammatory diseases (*1*, *5*, *22*–*24*). However, compared to FFA1 and FFA4, fewer synthetic ligands have been reported for FFA2 and FFA3, which may suggest certain obstacles in developing small-molecule ligands for these two SCFA receptors.

Phylogenetic analysis suggests that FFA4 diverged early from other FFAs (fig. S1). As a result, while FFA1 to FFA3 are structurally related with high sequence similarity, FFA4 shares very little sequence similarity with FFA1 to FFA3 (*25*). This implies distinct ligand recognition and signaling mechanisms for FFA4 and other FFAs. Regarding G protein coupling, FFA1 is a highly promiscuous GPCR that is capable of coupling to all four G protein families: G_s, $G_{i/o}$, $G_{q/11}$, and $G_{12/13}$ (*25*, *26*). FFA2 and FFA4 both can signal through $G_{i/o}$ and $G_{q/11}$ (*25*, *26*). For FFA4, a human splice variant ($FFA4^{Long}$) has been identified with an additional 16–amino acid segment in intracellular loop 3 (ICL3). This isoform is unable to induce $G_{q/11}$ signaling but is capable of coupling to β-arrestins (*27*, *28*). Crystal structures of highly engineered forms of FFA1 bound to synthetic agonists including TAK-875 and positive allosteric modulators (PAMs) have been reported (*29*–*31*), and the receptor in those structures stayed in the inactive state. To understand the molecular mechanisms by which LCFAs and SCFAs act on and activate their FFAs, we determined active structures of DHA-bound FFA1, butyrate-bound FFA2, and DHA-bound FFA4 in complex with an engineered heterotrimeric G_q protein (miniG_q) (*32*, *33*) by cryo–electron microscopy (cryo-EM). To further examine how FFA4 recognizes different ligands, we also solved a cryo-EM structure of the FFA4-miniGq complex with the most widely used synthetic FFA4

[1]Department of Pharmacology and Chemical Biology, School of Medicine, University of Pittsburgh, Pittsburgh, PA 15261, USA. [2]School of Pharmacy, Medical Biology Centre, Queen's University Belfast, Belfast BT9 7BL, Northern Ireland, UK. [3]Centre for Translational Pharmacology, School of Molecular Biosciences, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow G12 8QQ, Scotland, UK. [4]Cryo-EM Core Facility, Case Western Reserve University, Cleveland, OH 44106, USA.
*Corresponding author. Email: chengzh@pitt.edu (C.Z.); i.tikhonova@qub.ac.uk (I.G.T.); graeme.milligan@glasgow.ac.uk (G.M.)
†These authors contributed equally to this work.

agonist, TUG-891 (*34*). These structures revealed diverse modes of ligand recognition by FFAs. Together with computational simulations and mutagenesis studies, these studies highlight similarities and differences in modes of binding of the fatty acid ligands of varying chain lengths to their corresponding GPCRs.

RESULTS

Overall structures of FFA signaling complexes

We used the wild-type human FFA1, FFA2, and FFA4 in our structural studies. For FFA4, there are two isoforms, and we chose the canonical short form since $FFA4^{Long}$ does not couple to $G_{q/11}$ (*28*). The $miniG_q$ protein contains an engineered $miniG_{\alpha q}$ subunit (*35*) with the N-terminal 35 amino acids replaced by their corresponding residues in $G_{\alpha i}$. The same $miniG_q$ protein has been successfully used to obtain cryo-EM structures of several other G_q-coupled GPCRs (*32*, *36*). To enhance the stability of FFA-$miniG_q$ complexes for high-resolution cryo-EM data collection and structure determination, we assembled the complexes using the NanoBit tethering strategy in insect Sf9 cells (*37*) together with an antibody fragment, scFv16, which has been developed previously to stabilize the G_i heterotrimer (*38*). The structures of DHA-bound FFA1 and FFA4, TUG-891-bound FFA4, and butyrate-bound FFA2 with $miniG_q$ were determined to overall resolutions of 3.4, 3.2, 3.1, and 3.1 Å, respectively (Fig. 1A, figs. S2 to S5, and tables S1 and S2).

Most residues in the three FFAs, $miniG_q$, and scFv16 were modeled on the basis of the robust cryo-EM density maps. We also modeled several cholesterol and lipid molecules to fit strong density maps surrounding the transmembrane domains (TMDs) of FFA2 and FFA4 (fig. S6). In the two structures of FFA4, the density of the intracellular region of transmembrane helix 4 (TM4) and ICL2 is relatively weak, indicating a high degree of flexibility. This part is not modeled in the final structures (Fig. 1B). In contrast, ICL2 forms a helical structure in both FFA1 and FFA2 (Fig. 1B). A large part of the extracellular loop 2 (ECL2) of FFA2 is also not modeled because of weak density. Noticeably, in the structures of all three receptors, no helix 8 after TM7 was modeled because of very weak density. This suggests that the C-terminal region after TM7 in all three receptors is highly mobile when coupled with G proteins.

As for the ligands, the density maps for TUG-891 and DHA in FFA4 were sufficiently clear to enable the modeling of the entire ligands (figs. S2 and S3). In addition, the density of butyrate in FFA2 was also strong (fig. S5). However, because of the small size of butyrate and the limited resolution of the cryo-EM map, functional data were necessary to complement cryo-EM map information for accurate ligand modeling. In the case of FFA1, we modeled DHA in the orthosteric site based on a partial density map (fig. S4). Further discussion on this topic will be provided in the subsequent content.

Binding of DHA and the synthetic agonist TUG-891 to FFA4

Both DHA and TUG-891 bind to a pocket formed among the extracellular regions of TM3 to TM7 of FFA4 (Fig. 2A). The N-terminal region of FFA4 preceding TM1 folds inside the TMD and directly interacts with the ligands (Fig. 2A), resulting in almost complete shielding of the ligand-binding pocket from the extracellular milieu (Fig. 2B). This is similar to the N-terminal region of DP2, a GPCR that binds the fatty acid ligand prostaglandin D_2 (PGD_2) and forms a well-folded structure that participates in ligand interactions (*39*, *40*).

The cryo-EM density is strong for the ω-3 unsaturated chain of DHA, while the density for the carboxylate group is comparatively weaker. Nevertheless, we were able to model the entire DHA molecule that well fits the density. Our structure revealed that the ω-3 unsaturated chain of DHA adopts an "L-shape" binding pose, which enables it to penetrate deeply into a binding pocket that is rich in aromatic residues (Fig. 2A). The carboxylate head group of DHA, on the other hand, extends outward toward the extracellular milieu (Fig. 2A). The six carbon-carbon double bonds in DHA are surrounded by, and potentially form, extensive π-π interactions with aromatic residues $F27^{N}$, $F88^{2.53}$, $F115^{3.29}$, $W198^{ECL2}$, $W207^{5.38}$, $F211^{5.42}$, $W277^{6.48}$, and $F303^{7.35}$ [superscripts represent Ballesteros-Weinstein numbering (*41*)] (Fig. 2C). Hydrophobic residues $I280^{6.51}$, $I284^{6.55}$, $I287^{6.58}$, and $L288^{6.59}$ lining one side of TM6 form additional hydrophobic interactions with DHA to further stabilize lipid binding (Fig. 2C).

TUG-891 (3-(4-((4-fluoro-4-methyl-[1,1-biphenyl]-2-yl)methoxy) phenyl)-propanoic acid) binds to FFA4 in a similar L-shape binding pose to DHA, overlapping extensively with the ω-3 chain of DHA and with the ortho-biphenyl moiety of TUG-891 defining the bottom of the binding pocket (Fig. 2A). This leads to the observation of similar sets of hydrophobic and π-π interactions between the three benzene rings of TUG-891 and FFA4 (Fig. 2D). Mutagenesis studies that we performed previously are in accord with these observations. For example, alteration of $F115^{3.29}$, $W207^{5.38}$, $F211^{5.42}$, and $W277^{6.48}$ (each to A) resulted in a complete lack of response to TUG-891, while altering $F88^{2.53}$ to A and $F303^{7.35}$ to H resulted in greater than 10-fold reduction in potency (*42*). We also show here that altering $F303^{7.35}$ to A almost abolished the TUG-891 response (Fig. 2E). The mutation W198A also lacked response to both TUG-891 and DHA (Fig. 2E and fig. S7). However, although mutations to Ala of $F27^{N}$ and $F28^{N}$ each lacked response to TUG-891 (Fig. 2E and fig. S7), these results could not be interpreted because, although well expressed following transient transfection in human embryonic kidney (HEK) 293 cells (fig. S7C), each of these mutants failed to reach the cell surface (fig. S7D). Mutation to Ala of either $I280^{6.51}$ or $I284^{6.55}$ also produced a greater than 100-fold (I280A) (*42*) or a more modest but still significant (I284A) reduction in potency for TUG-891 (Fig. 2E and fig. S7A). We previously observed a similar pattern of effects of these mutations for the ω-3 PUFA α-linolenic acid (*all-cis*-9,12,15-octadecatrienoic acid) (*42*). The I284A mutant did not alter DHA potency (Fig. 2E and fig. S7A) but notably reduced DHA efficacy, causing it to act as a partial agonist compared to TUG-891 at this mutant (fig. S7B). This effect was not observed for other mutants.

These initial experiments measured FFA4 interactions with arrestin-3 (also designated β-arrestin-2). Some GPCRs differ in arrestin-2 (also designated β-arrestin-1) versus arrestin-3 binding. However, additional arrestin-2 interaction studies also showed significantly reduced TUG-891 potency at I284A versus wild type FFA4 (fig. S8) and low efficacy of DHA. The cryo-EM structure contained an engineered $miniG_{\alpha q}$, and TUG-891 also showed reduced potency at I284A FFA4 in Ca^{2+} mobilization assays that reflect G_q activation (fig. S9).

An additional difference in the binding mode of TUG-891 and DHA is that the linking ether oxygen of TUG-891 forms a hydrogen bond with $T119^{3.33}$, which is absent in the DHA-bound FFA4 (Fig. 2, C and D). Notably, mutation of T119A significantly impaired the ability of TUG-891 to activate FFA4, in both arrestin-3 interaction (fig. S7A) and, particularly, G_q-mediated Ca^{2+} elevation assays (fig. S9) (*42*), indicating an important role of this hydrogen bond in TUG-891 binding and function. The effect of the T119A mutation on potency of TUG-891 was also observed in arrestin-2 interaction studies

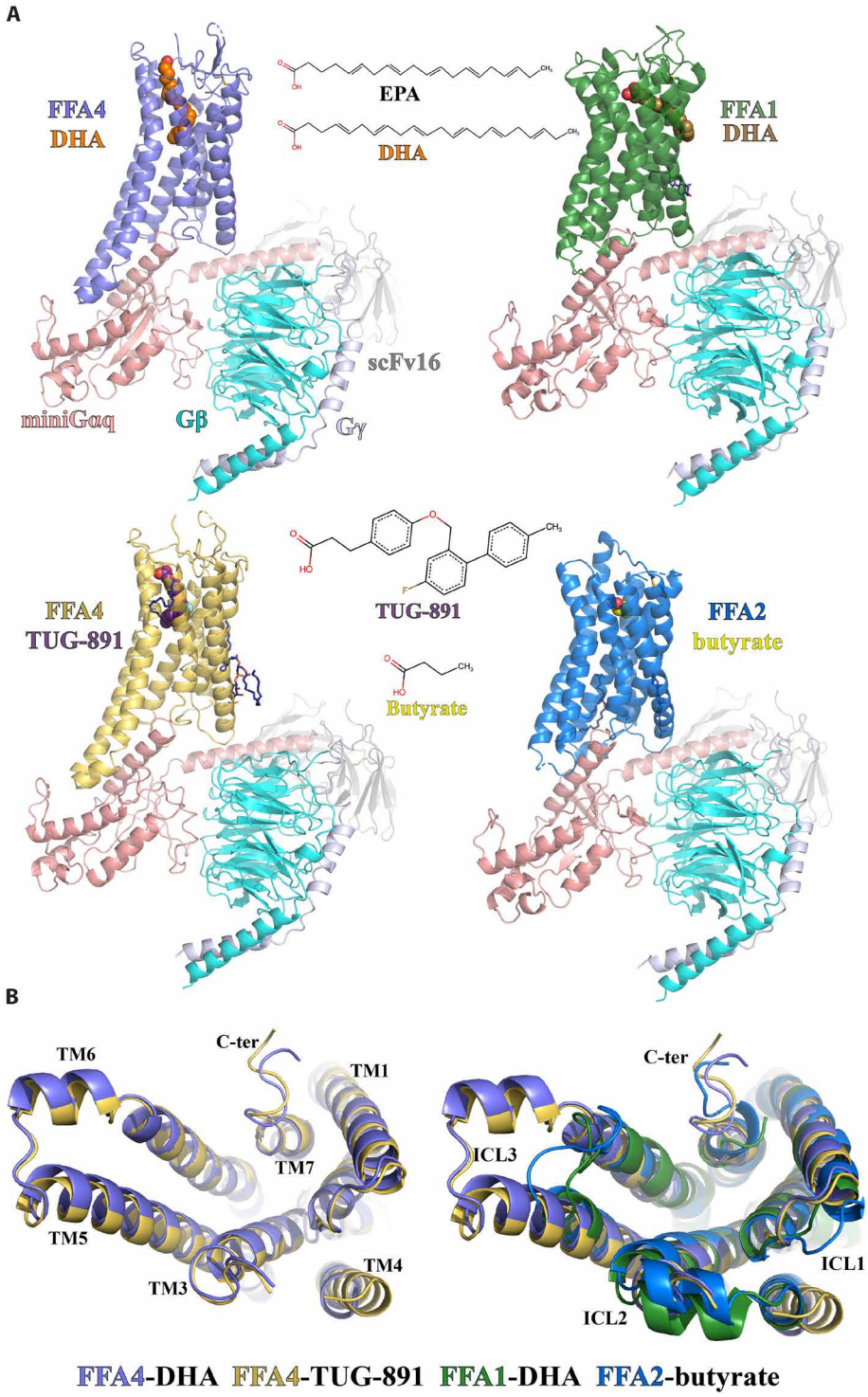

Fig. 1. Overall structures of FFA1, FFA2, and FFA4 signaling complexes. (**A**) Overall structures of DHA-FFA4 (slate), DHA-FFA1 (green), TUG-891-FFA4 (dark yellow), and butyrate-FFA2 (blue), each in complex with miniG_q, together with the chemical structures of the bound ligands. mini$G_{\alpha q}$, G_β, and G_γ subunits are colored salmon, cyan, and light blue, respectively. ScFv16 is colored gray. The LCFA EPA is also shown for comparison to DHA (see the main text for discussion). (**B**) Comparison of the above structures as seen from the intracellular face.

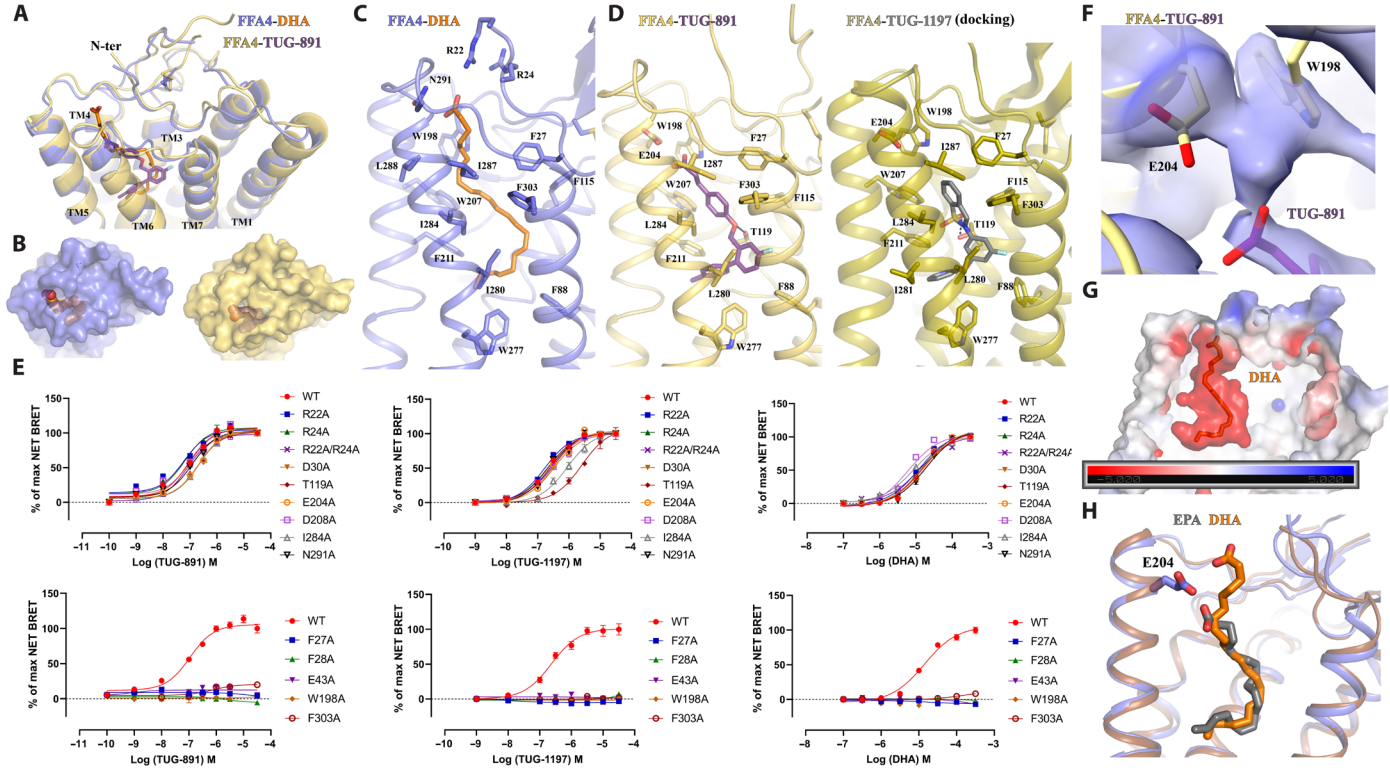

Fig. 2. Ligand binding in FFA4. (**A** to **D**) Details of the interactions of DHA (orange) and TUG-891 (purple) with FFA4. (A) illustrates the general positions of the two ligands, while (B) highlights the closed nature of the occupied ligand-binding pockets. Details of key residues of the binding pockets are highlighted for DHA (C) and TUG-891 [(D), left]. TUG-1197 docked into the FFA4 structure [(D), right] highlights the important role of T119 and the similarity of the binding mode of TUG-1197 and TUG-891 at the bottom of the pocket. (**E**) Various point mutants of FFA4 generated and assessed for the ability of each of TUG-891, TUG-1197, and DHA to promote interactions with arrestin-3. See fig. S7 for quantitation. In concert with the large-scale mutagenesis studies reported previously (*42*), this provides a comprehensive analysis of the orthosteric binding pocket of FFA4. (**F**) Continuous electron density observed between W198 and E204. The cryo-EM map is contoured at the level of 0.13. (**G**) Negative charge potential of the FFA4 binding pocket with DHA. (**H**) Additional length of DHA compared to EPA and the position of the DHA carboxylate above and beyond E204.

(fig. S8). No such effects of this mutation on the function of either DHA (Fig. 2E and figs. S7A to S9) or α-linolenic acid (*42*) was observed. A distinct group of sulfonamide-based FFA4 agonists have been reported (*25*, *43*). Among them, TUG-1197 (2-(3-(pyridin-2-yloxy) phenyl)-2,3-dihydrobenzo[*d*]isothiazole 1,1-dioxide) showed large, 5- to 20-fold, reductions in potency at the T119A mutant in each of arrestin-2 (fig. S8), arrestin-3 (fig. S7), and Ca^{2+} mobilization assays (fig. S9). In addition, TUG-1197 showed greatly reduced efficacy in comparison to TUG-891 in both arrestin assays (Fig. 2E and figs. S7, A and B, and S8). Docking of this ligand to the obtained structures of FFA4 suggests a similar binding pose to TUG-891 and a clear interaction of the sulfonamide, which overlaps in its location with the ether oxygen of TUG-891, with T119 (Fig. 2D).

Intriguingly, in the structure, the carboxylate group of the phenylpropanoic acid of TUG-891 is positioned in proximity to $E204^{5.35}$ and $W198^{ECL2}$ (Fig. 2D). The cryo-EM density between $E204^{5.35}$ and $W198^{ECL2}$ appears to be continuous (Fig. 2F), which raises the possibility that a water molecule may be present between these two residues to facilitate extensive polar interactions between the carboxylate group of TUG-891 and $E204^{5.35}$ and $W198^{ECL2}$ of FFA4. The mutant E204A modestly but significantly reduced potency of TUG-891 but not of DHA (fig. S7A), whereas a W198A mutant was not activated by either TUG-891 or DHA (Fig. 2E and fig. S7A).

The carboxylate group of DHA, which is associated with weak cryo-EM density, is modeled close to the polar residues $R22^{N}$ and $R24^{N}$ from the N-terminal region and $N291^{ECL3}$ from ECL3 (Fig. 2C). Note that the density of the carboxylate of DHA is not continuous with the density of its long chain (fig. S3G). As a result, the modeling may not be entirely accurate. Single mutations R22A and R24A or the double mutation R22A/R24A did not reduce the potency of DHA or either of the synthetic agonists TUG-891 and TUG-1197 in arrestin-3 assays (Fig. 2E and fig. S7A). We conclude that the binding of DHA to FFA4 is mainly driven by hydrophobic and π-π interactions. Nevertheless, the overall binding pocket of DHA exhibits a negatively charged potential, which may help to position the carboxylate group of DHA at the extracellular surface (Fig. 2G). A similar charge interaction-facilitated lipid recognition mechanism has also been suggested for other lipid GPCRs including prostaglandin E_2 (PGE_2) receptors and lysophospholipid GPCRs (*39*, *40*). However, although DHA adopts a similar binding orientation to PGE_2 and lysophospholipids, their binding sites are located differently (fig. S10A). The pockets of PGE_2 and lysophospholipids form among TM1-TM2-TM3-TM7 or TM2-TM3-TM5-TM6-TM7, while the pocket of DHA in FFA4 forms among TM3-TM4-TM5-TM6-TM7 (fig. S10A). To the best of our knowledge, no other lipid GPCRs have been shown to have lipid-binding pockets at locations similar to that of DHA in FFA4.

During the preparation of our manuscript, other research groups published structures of FFA4 bound to several LCFAs, including eicosapentaenoic acid (EPA), an ω-3 PUFA, and TUG-891 (*44*). The structure of EPA bound to FFA4 showed a highly similar binding pose to that of

DHA observed in our structure, especially with regard to their ω-3 chains (Fig. 2H). However, DHA is two carbons longer than EPA and contains an additional double bond. As a consequence, the carboxylate group of DHA extends further toward the extracellular surface above E204$^{5.35}$, while the carboxylate group of EPA is located below E204$^{5.35}$ (Fig. 2H). Furthermore, we observed a slightly different binding mode of TUG-891 in our cryo-EM structure, which is strongly supported by clear cryo-EM density, compared to the published structure (fig. S10B). The overall position of TUG-891 in the published structure is closer to the extracellular surface compared to that in our structure (fig. S10B). As a result, the carboxylate group of TUG-891 in our structure is too distant from N291^{ECL3} to form a hydrogen bond. Consistent with this, we did not observe an effect of the N291A mutant on the potency of TUG-891 (Fig. 2E and fig. S7A), and such a mutant was not reported in the published study (*44*). In addition, the hydrogen bond between the ether oxygen of TUG-891 and T119$^{3.33}$ in our structure is absent in the published structure (fig. S10B). The discrepancies in TUG-891 binding in the two structures may indicate a high degree of mobility of TUG-891 in FFA4. Nevertheless, despite these discrepancies, our mutagenesis studies demonstrated the important role of T119$^{3.33}$ in the action of TUG-891 (figs. S7A to S10B), and the positioning of the ortho-biphenyl, which is a key feature of TUG-891 and related synthetic FFA4 agonists, is entirely in accord with our earlier mutagenesis studies (*42*).

Distinct mechanisms of DHA recognition by FFA1 and FFA4

Despite their similar ligand recognition profiles to LCFA receptors, FFA1 and FFA4 exhibit little sequence similarity and a distant phylogenetic relationship (fig. S1). To investigate whether FFA1 uses a distinct mechanism to recognize DHA, we sought to obtain a cryo-EM structure of miniGq-coupled FFA1 bound to DHA. However, while the overall resolution of the structure reached 3.4 Å (fig. S4), the modeling of DHA proved to be challenging.

Previous structural studies on FFA1 with synthetic agonists and allosteric modulators (*29*–*31*) have identified two binding sites: "Site 1" or "'outer-leaflet pocket" is located in the extracellular region within the 7TM as the putative orthosteric site for synthetic agonists TAK-875 (*31*) and MK-8666 (*29*), while site 2 or inner-leaflet pocket is located on the surface of the 7TM above ICL2 as the allosteric site for the synthetic ago-PAMs AP8 (*29*) and compound 1 (3-benzyl-4-(cyclopropyl-(4-(2,5-dichlorophenyl)thiazol-2-yl)amino)-4-oxobutanoic acid) (*30*, *45*). Similar allosteric sites have also been identified for the C5a receptor (*46*) and β2-adrenergic receptor (*47*). Evidence from a previous study showed that TAK-875 exhibited positive cooperativity with the LCFA ligand of FFA1 γ-linolenic acid (γ-LA), suggesting that site 1 is not the primary site for γ-LA (*48*).

In our structure, we observed weak density in site 1 for DHA (Fig. 3A). We further performed local refinement focusing on the receptor to improve this density, which allowed us to model the part of DHA from C1 to C8 together with the carboxylate group in site 1 (Fig. 3A). The region of DHA from C9 to C22 was assigned with zero occupancy in the structure since it could not be well modeled. In our structural model, the carboxylate group of DHA forms salt bridges with R183$^{5.39}$ and R258$^{7.35}$, while the carbon chain from C1 to C8 forms hydrophobic and π-π interactions with surrounding residues F87$^{3.33}$, L138$^{4.57}$, F142$^{4.61}$, W174^{ECL2}, and L186$^{5.42}$ of FFA1 (Fig. 3A). A very recent study also reported a similar partial cryo-EM density for DHA in site 1 of FFA1 (*49*). The binding pose of DHA may not be entirely accurate because of the limitations of the weak density. Nevertheless, the density of most residues in site 1 after local refinement was sufficiently clear to allow unambiguous modeling, suggesting that the relatively weaker density of DHA was likely due to the high flexibility of DHA in this site.

Strong density in site 2 indicated the presence of a lipid molecule (Fig. 3B). It is possible that site 2 is another binding site for DHA in FFA1. LCFAs have been suggested to bind to site 2 as the inner-leaflet pocket of FFA1 (*45*). However, this site can also accommodate other lipid molecules, making it challenging to confirm it as the specific binding site for DHA. Nonetheless, since the cryo-EM sample contained a high concentration of DHA, it was the most prevalent lipid present. Thus, we fitted the density observed with a DHA molecule (Fig. 3B), and the binding pose of DHA in this site highly resembles that of the ago-PAM AP8 (*29*, *30*). The carboxylate group of DHA forms hydrogen bonds with two tyrosine residues Y44$^{2.42}$ and Y114^{ICL2}, while the carbon chain forms hydrophobic interactions with hydrophobic residues from TM3 to TM5 of FFA1 (Fig. 3B). We also observed a hydrogen bond between DHA and Y243 from the α5 of mini-$G_{\alpha q}$, which is the major receptor interaction site in mini-$G_{\alpha q}$ (Fig. 3B). To determine the roles of Y44$^{2.42}$ and Y114^{ICL2} in DHA binding, we generated Y44A and Y114A mutations (Fig. 3C). The Y44A mutation resulted in poor FFA1 expression, preventing further characterization (Fig. 3C). However, Y114A exhibited near wild-type expression level but significantly reduced DHA potency in arrestin-3 recruitment assays (Fig. 3C), indicating that Y114 is important for ligand function.

In the bile acid receptor GPBAR, the endogenous lipid ligand bile acid binds to a site formed between TM3 and TM4 above ICL2, similar to site 2 in FFA1 (*50*). However, bile acid can also bind to a more conventional orthosteric site located in the extracellular region of GPBAR (*50*). It is possible that DHA binds to FFA1 similarly to bile acid in GPBAR. Site 1 in FFA1 serves as the primary site for DHA, where DHA exhibits a high flexibility, while site 2 in FFA1 serves as the secondary site for DHA.

Neither site 1 nor site 2 in FFA1 is conserved in FFA4, providing further evidence of the distant phylogenetic relationship between FFA1 and FFA4 despite their similar ligand preferences. In FFA4, the carboxylate group of DHA is positioned near the ligand entrance at the extracellular surface (Fig. 2C). In our previous studies on DP2, we observed that the prostaglandin PGD_2 adopts a "polar-group-in" binding pose in DP2 with its carboxylate group buried deep within the binding pocket, while another prostaglandin, PGE_2, adopts a polar-group-out binding pose in PGE_2 receptors (EPs) with its carboxylate group positioned near the extracellular surface, and these two different binding poses of prostaglandins are facilitated by the distinct charge potentials of the binding pockets in DP2 and EPs (*39*, *40*). Similarly, DHA in FFA4 adopts a polar-group-out binding pose in a negatively charged environment (Fig. 2G), although the role of the charge potential in DHA binding is not clear.

Recognition of SCFAs by FFA2

As anticipated from the relatedness of FFA2 to FFA1 (*51*), the overall structures of these two receptors are similar. In addition, butyrate and the carboxylate head group of DHA in site 1 of FFA1 are very close if the structures of FFA1 and FFA2 are aligned (Fig. 4A). In the structure of FFA2-butyrate, the carboxylate group of butyrate is coordinated by a pair of adjacent arginine residues, R180$^{5.39}$ and R255$^{7.35}$ (Fig. 4B). These two arginine residues are highly conserved among FFA1 to FFA3 (*51*, *52*). The same pair of residues in FFA1

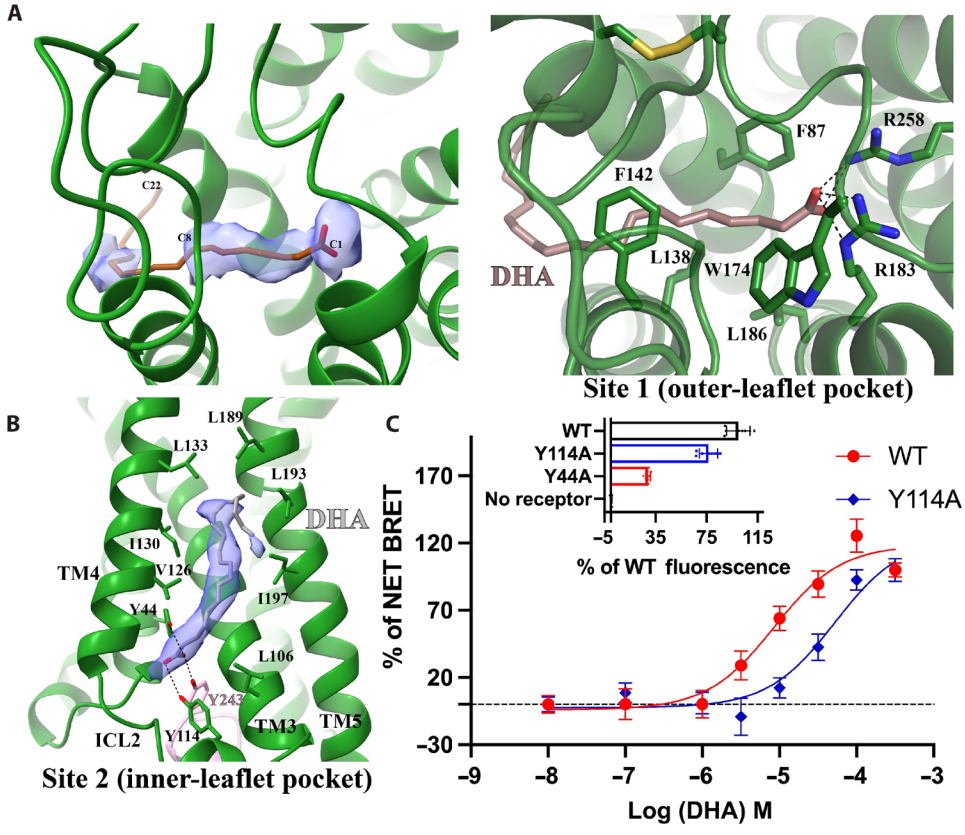

Fig. 3. Potential DHA binding sites in FFA1. (**A**) DHA binding in site 1. The partial cryo-EM density map of DHA colored light blue in the left panel is contoured at the level of 0.07. C1, C8, and C22 atoms of DHA are labeled. The occupancy of DHA C9-C22 was assigned as zero due to a lack of density. The details of interactions between DHA and FFA1 in site 1 are shown in the right panel. DHA is colored brown. (**B**) Putative DHA binding in site 2. The strong cryo-EM density map in this site is contoured at the level of 0.12. The modeled DHA molecule is colored gray. Polar interactions are shown as black dashed lines. FFA1 is colored green. (**C**) DHA concentration-response curves at the wild-type receptor (WT) and Y114A mutant measured by arrestin-3 interaction assays and relative expression of WT (black), Y114A (blue), and Y44A (red) FFA1 measured by eYFP fluorescence shown on top. Data are means ± SEM for $n = 3$ or more.

interacts with the carboxylate group of TAK-875 (*31*). In FFA2, mutating R180$^{5.39}$ to other amino acids, including Ala and Lys, as well as mutating R255$^{7.35}$ to Ala, eliminates the response to SCFAs (*52*). Furthermore, the mutation to Ala of H242$^{6.55}$, which interacts with R255$^{7.35}$ to organize the binding pocket for the carboxylate of SCFAs (Fig. 4B), also abolishes SCFA function (*52*). These mutations eliminate the binding of SCFAs rather than simply affecting ligand function, as evidenced by the fact that SCFAs are unable to compete for binding with an FFA2 orthosteric antagonist, the affinity of which is only slightly reduced compared to the wild-type receptor at each of the R180A, R255A, and H242A mutants of FFA2 (*53*). While R180$^{5.39}$ and R255$^{7.35}$ are conserved in FFA1, H242$^{6.55}$ in FFA2 is replaced by N244$^{6.55}$ in FFA1, which does not interact with R$^{5.39}$ and R$^{7.35}$.

In addition to R180$^{5.39}$ and R255$^{7.35}$, two other polar residues from ECL2, Q148^{ECL2}, and Y165 ^{ECL2} are also close to butyrate (Fig. 4B). Y165 ^{ECL2} forms a hydrogen bond with the carboxylate of butyrate, while Q148^{ECL2} interacts with and potentially stabilizes R180$^{5.39}$ (Fig. 4B). We have previously demonstrated that substituting Q148^{ECL2} to a glutamate residue results in a reduction of potency for SCFAs (*54*). It has an even more dramatic effect on larger synthetic FFA2 agonists including compound 1 (3-benzyl-4-(cyclopropyl-(4-(2,5-dichlorophenyl)thiazol-2-yl)amino)-4-oxobutanoic acid), where the agonist function is all but ablated (*54*). This may be due to the repulsion of negatively charged carboxylate groups in the mutated E148 residue and agonists. In addition, the mutation of Y165A resulted in a nearly 15-fold reduction of the potency of SCFAs (*54*), indicating the important role of this residue in ligand binding as well.

The opening between TM3 and TM4 of FFA1 that allows for binding of TAK-875 is closed in FFA2, resulting in a small orthosteric pocket in FFA2 with limited space to accommodate the carbon chains of fatty acids (Fig. 4C). This may explain the selectivity of FFA2 for SCFAs over LCFAs. When comparing the structures of FFA1 and FFA2 coupled with miniGq, it becomes apparent that while TM3 aligns well between the two, TM4 in FFA2 shifts toward TM3 in comparison to FFA1 (Fig. 4D). H140$^{4.56}$ in TM4 of FFA2 appears to play an important role in determining the selectivity of chain length of fatty acids. While alteration of this residue to Ala reduces the potency of SCFAs, it enables binding and function of the C6 fatty acid caproate and, to a more modest degree, C8 caprylate (*52*). In our structural model, H140$^{4.56}$ forms a hydrogen bond with E182$^{5.41}$ in FFA2 (Fig. 4D), which is replaced by S185$^{5.41}$ in FFA1. Such a bond may not be stable as suggested by the weak density of E182$^{5.41}$, likely due to the surrounding lipidic environment. Nevertheless, the longer side chain of E182$^{5.41}$ may force H140$^{4.56}$ together with TM4 to be positioned toward TM3, which would result in a narrower cleft between TM3 and TM4 when compared to FFA1

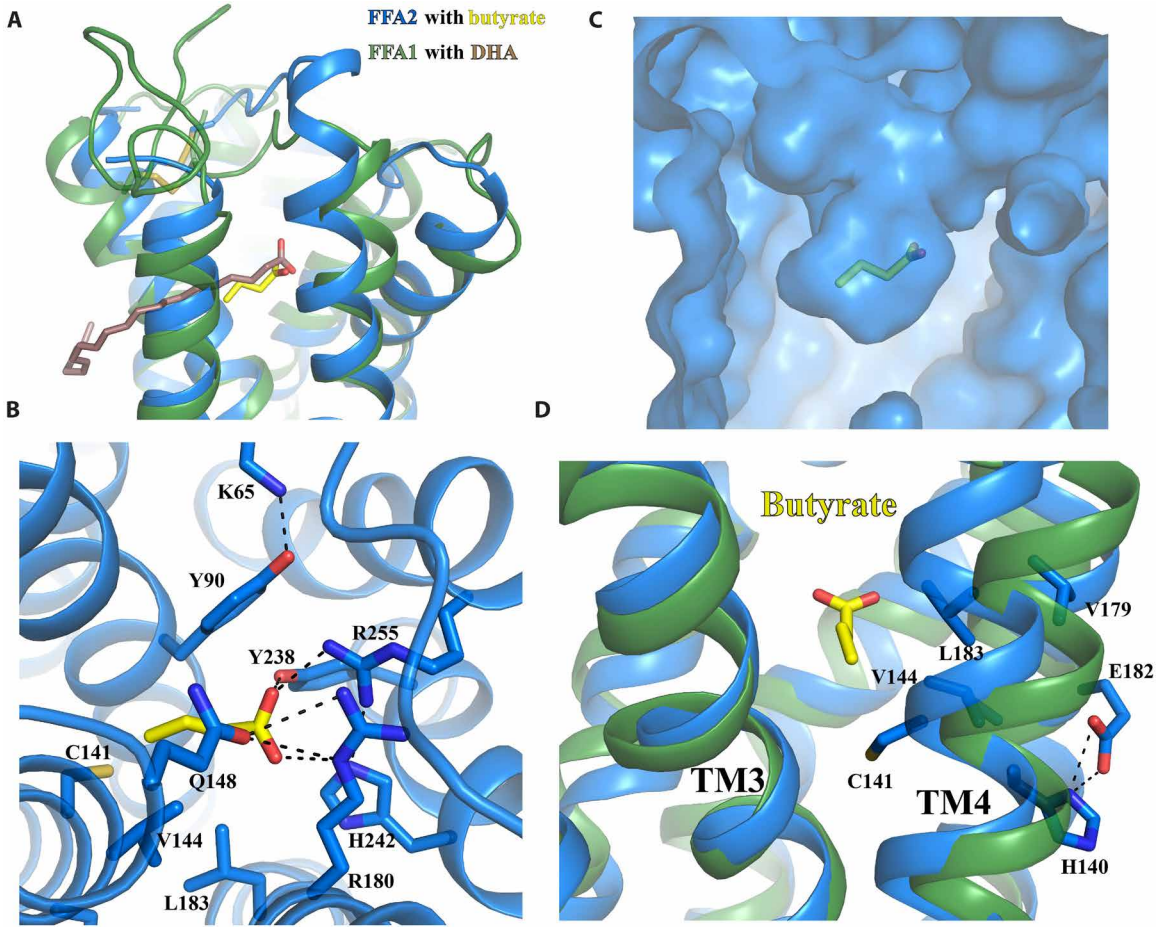

Fig. 4. Recognition of butyrate by FFA2. (**A**) Structural alignment of FFA2-butyrate and FFA1-DHA. The carboxylate of butyrate occupies the equivalent position to the carboxylate of DHA. (**B**) Details of interactions between butyrate and FFA2. (**C**) Overall shape of the butyrate binding pocket. (**D**) Differences in location of TM3 and TM4 in FFA1 and FFA2. In all panels, FFA1 and FFA2 are colored green and blue, respectively, while DHA and butyrate are colored brown and yellow, respectively. Polar interactions are shown as black dashed lines.

(Fig. 4D). The H140$^{4.56}$-E182$^{5.41}$ pair also defines the size of the hydrophobic pocket involving C141$^{4.57}$, V144$^{4.60}$, V179$^{5.38}$, and L183$^{5.42}$ in FFA2 that accommodates the hydrophobic tail of butyrate (Fig. 4D). Notably, the mutations V179A and L183A do not change the potency of SCFAs but cause ~10-fold reduction of the potency for larger synthetic FFA2 agonists compound 1 and compound 2 ((*R*)-3-(cyclopentylmethyl)-4-(cyclopropyl-(4-(2,6-dichlorophenyl) thiazol-2-yl)amino)-4-oxobutanoic acid) (*54*). C141$^{4.57}$ is located at the bottom of the SCFA binding pocket and forms van der Waals interactions with the last carbon of butyrate. This residue in bovine FFA2 is replaced by Gly. The C141G mutation in human FFA2 leads to an altered ligand preference for longer saturated and unsaturated C5 and C6 fatty acids, similar to that of bovine FFA2 (*55*).

In previous studies, two tyrosine residues in FFA2, Y90$^{3.29}$ and Y238$^{6.51}$, were suggested to participate in the binding of SCFAs (*53*, *54*). This hypothesis was supported by the interaction energy analysis based on our molecular dynamics (MD) simulations, and mutating each residue to alanine significantly reduced the potency of SCFAs (*53*, *54*). In the structure of FFA2-butyrate, Y90$^{3.29}$ forms direct hydrophobic interactions with the short chain of butyrate, while Y238$^{6.51}$ forms hydrogen-bonding and cation-π interactions with the critical arginine residue R255$^{7.35}$ (Fig. 4B), potentially stabilizing it in the appropriate position to interact with butyrate. In addition, Y90$^{3.29}$ is stabilized by a hydrogen bond with K65$^{2.60}$ (Fig. 4B), and mutations of K65$^{2.60}$ to either alanine or glutamate substantially reduce the potency of SCFAs (*56*). This indicates the important role of the K65$^{2.60}$-Y90$^{3.29}$ pair in SCFA binding. K65$^{2.60}$ of human FFA2 is replaced by R65$^{2.60}$ in mouse FFA2. Note that SCFAs including butyrate display lower potency at mouse FFA2 compared to human FFA2 (*57*). Consistent with this, a K65R mutation in human FFA2 resulted in a small but still significant reduction of potency of SCFAs (*56*).

Molecular determinants of ligand recognition by FFAs revealed by MD simulations

To evaluate the stability and dynamics of the receptor-ligand interactions in FFA2 and FFA4, we performed MD simulations of the two receptors in agonist-bound and unbound forms in a water-lipid bilayer. We also performed MD simulations on FFA1 bound to DHA. Guided by the protonation state prediction, D208$^{5.39}$ was protonated in the simulations of the FFA4 complex with TUG-891, which reduced the repulsion of the carboxyl groups of nearby

E204$^{5.35}$, D208$^{5.39}$, and TUG-891. While mutation of E204$^{5.35}$ to Ala reduced the potency of TUG-891 in each of arrestin-2/3 and Ca^{2+} mobilization assays, this alteration did not alter the potency of DHA. By contrast, while mutation of D208 to Ala was without effect on the potency of TUG-891, this alteration significantly increased the potency of DHA in both arrestin-2- and arrestin-3 interaction assays (figs. S7A and S8), consistent with differences in the detailed location of the carboxyl of the synthetic and fatty acid ligands (Fig. 5). The receptor and agonists had small fluctuations, with DHA having a higher mobility, in all the three 1-μs MD simulations (table S3 and fig. S11). Similar high mobility of DHA was also observed in our MD simulations on the FFA1-DHA complex (fig. S12), which is consistent with the poor cryo-EM density of DHA in site 1 of FFA1 (Fig. 3A). From the average ligand-residue interaction energy, DHA and TUG-891 showed electrostatic attraction to nearby FFA4 residues R24^{N} and R22^{N} and repulsion to E204$^{5.35}$ (Fig. 5, A and B). However, mutation of either or both R24^{N} and R22^{N} residues to alanine did not alter ligand potency (Fig. 2E and fig. S7A). In the simulations, in the absence of the agonists, R22^{N} and R24^{N} interact with E204$^{5.35}$ and D208$^{5.39}$, thus occluding ligand access to the interhelical hydrophobic cavity. TUG-891, with its carboxyl group located deeper in the binding cavity, is further stabilized by electrostatic interactions with W198^{ECL2}. The MD simulations allowed us to observe water clusters at the extracellular vestibule of the orthosteric binding site of FFA4 (Fig. 5, A and B), which mediate polar interactions between the agonists, E204$^{5.35}$, D208$^{5.39}$, W198^{ECL2}, R24^{N}, and R22^{N}. This supports the presence of water, as suggested by the cryo-EM density (Fig. 2F).

According to ligand fragment interaction energy calculations, the carboxyl group of the FFA4 agonists is further stabilized by electrostatic attraction to electron-deficient aromatic hydrogens of F25^{N} and F27^{N} and van der Waals interactions with their aromatic rings (Fig. 5A and table S4). The first two double bonds of DHA demonstrate high mobility and do not form persistent interactions with surrounding residues. Deeper inside the pocket, double bonds 3—5 form stable van der Waals interactions with F115$^{3.29}$, M118$^{3.32}$, and T119$^{3.33}$ (Fig. 5A). The last double bond of DHA is engaged in van der Waals interactions with F211$^{5.42}$, I280$^{6.51}$, and the "toggle switch," W277$^{6.48}$ (Fig. 5A). A similar picture is observed for TUG-891, with its last two aromatic rings, having interaction energy with these residues, along with I287$^{6.58}$ and I284$^{6.55}$ (Fig. 5B), the mutation of which to Ala, as noted earlier, reduces the potency of TUG-891 (figs. S7A to S9).

In FFA2, butyrate is stabilized by strong electrostatic interactions with R180$^{5.39}$ and R255$^{7.35}$, together with Y238$^{6.51}$, Y90$^{3.33}$, and Q148^{ECL2} (Fig. 5C), as suggested by the cryo-EM structure. In the simulations, we also observed electrostatic stabilization of the ligand by H242$^{6.55}$ (Fig. 5C). The hydrogen bond between H140$^{4.56}$ and E182$^{5.41}$ suggested by cryo-EM was found to be persistent throughout the MD simulations.

No large movements of the helices in FFA2 and FFA4 occurred upon the removal of the agonist and miniG_q protein during 1-μs simulations. However, we did observe the start of deactivation processes of the receptors. We saw an increase in mobility of the aromatic residues at positions 5.47, 6.44, and 6.48 (fig. S13) associated with GPCR activation (*58*, *59*). In addition, the formation of the "ionic lock" involving E$^{3.49}$ and R$^{3.50}$ of the ERY motif and the conformational change of the microswitch residue at position 7.53 of the NPxxY motif, both leading to an inactive state of GPCRs, were observed in the receptors (fig. S14). These changes were observed in all simulations lacking the miniG_q protein, usually to a greater extent in the systems without the agonist. In summary, the MD simulations support the importance of hydrophobic and aromatic contacts deep inside the interhelical cavity in FFA4 as opposed to the polar contacts at the extracellular cavity of FFA2.

Activation mechanisms of FFAs

As inactive structures of FFA4 have not been experimentally solved, we took an inactive structure model (FFA4-AF) obtained from the GPCRdb database (*60*), which was generated using an AlphaFold-based multistate prediction protocol (*61*), in our structural comparison analysis. Our analysis showed substantial conformational changes at the cytoplasmic region, including the outward and inward displacements of TM6 and TM7, respectively, as observed in the activation of other class A GPCRs (*58*, *59*, *62*), when comparing the active DHA-bound structure and the inactive structure of FFA4 (Fig. 6A). The extracellular region exhibits rather modest

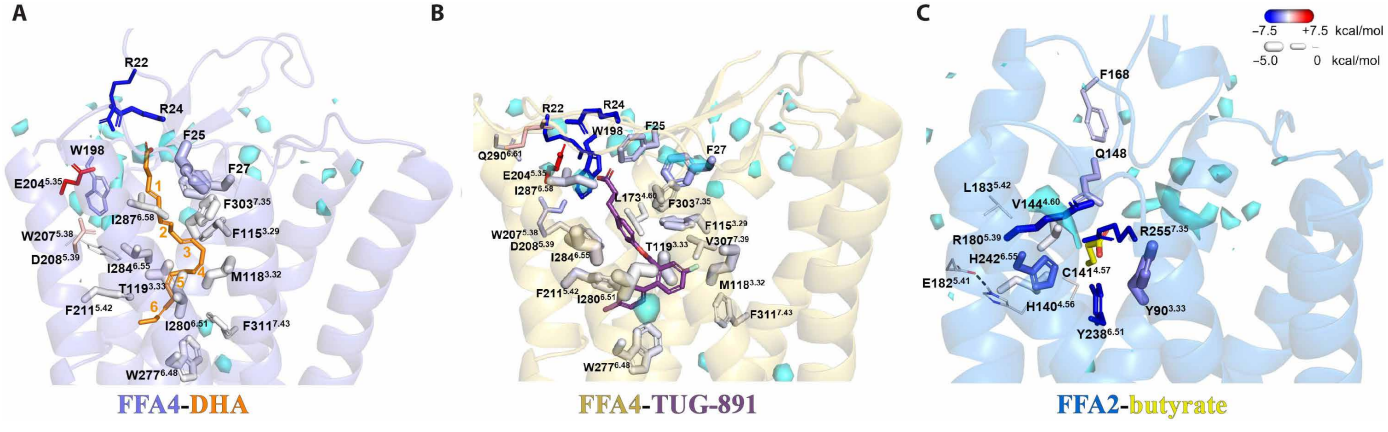

Fig. 5. Agonist recognition of FFA4 and FFA2 probed by MD simulations. (**A**) FFA4-DHA, (**B**) FFA4–TUG-891, and (**C**) FFA2-butyrate complexes. A representative frame is shown with key residues forming contacts with DHA (orange) with DHA carbon-carbon double bonds numbered 1 to 6, and TUG-891 (purple) and butyrate (yellow) in stick representation. The size and color of the residues correspond to the average strength of van der Waals and electrostatic interactions with the agonist, respectively. Water clusters observed in the MD simulations are shown in the cyan surface-like representation. The superscripts in the amino acid labels denote the Ballesteros–Weinstein generic GPCR residue numbering.

conformational differences except for the N-terminal region. In the FFA4-AF, the N-terminal region sticks out and does not interact with the rest of the receptor (Fig. 6A). This contrasts the "U-shape" N-terminal segment that is buried inside the 7-TM bundle in the active FFA4, which is likely stabilized by the interaction between the N-terminal residue F27 and DHA (Figs. 1C and 6A).

At the bottom of the ligand-binding pocket of inactive FFA4, a triad of aromatic residues F216$^{5.47}$, F274$^{6.44}$, and W277$^{6.48}$ form an aromatic cluster with extensive π-π interactions (Fig. 6B). In the active structure of FFA4-DHA, the long chain of DHA reaches this motif and causes rearrangements of the three residues (Fig. 6B). In many class A GPCRs, W$^{6.48}$ and F$^{6.44}$ constitute a conserved "activation switch" microdomain, and conformational rearrangement of this microdomain serves as a crucial step in the activation mechanism (*58*, *59*, *62*). The movement of F274$^{6.44}$ and W277$^{6.48}$ breaks the continuous helical structure of TM6 of FFA4 (Fig. 6B), resulting in an outward displacement of the cytoplasmic segment of TM6, a hallmark of GPCR activation. Note that the AlphaFold-predicted inactive GPCR structures may not be entirely precise. As elaborated in the following paragraphs, the inactive structures of FFA1 and FFA2 predicted by AlphaFold2 bear high resemblance to their active structures coupled with miniG_q. Therefore, our proposed FFA4 activation mechanism, which involves conformational changes of F216$^{5.47}$, F274$^{6.44}$, and W277$^{6.48}$, only represents a plausible hypothesis.

When we tried to compare the AlphaFold-predicted inactive structure of FFA2 obtained from the GPCRdb database (FFA2-AF) with our active structure of FFA2-butyrate, unexpectedly, we found that FFA2-AF closely resembles the active conformation of FFA2 with only subtle differences at the cytoplasmic region, indicative of an active conformation of FFA2-AF (Fig. 6C). This finding complicates the examination of conformational changes during receptor activation. Nonetheless, we observed inward movement of TM5 and TM7 at the extracellular region in the FFA2-butyrate structure compared to the ligand-free structure of FFA2-AF, which is likely due to interactions between butyrate and the two arginine residues R180$^{5.39}$ and R255$^{7.35}$ (Fig. 6C). Similar inward movement of the extracellular segment of TM5 was also observed in the β2-adrenergic receptor during receptor activation due to hydrogen bonds between agonists and two serine residues of TM5, leading to the rearrangement of the P$^{5.50}$/I$^{3.40}$/F$^{6.44}$ motif at the core region of the 7-TM bundle and outward movement of TM6 (*63*). The PIF motif functions as a molecular microswitch in the activation of some class A GPCRs (*58*, *59*, *62*). We hypothesize that FFA2 adopts a similar activation mechanism, where the agonist-induced inward movement of TM5 leads to the rearrangement of the core triad motif P191$^{5.50}$/T97$^{3.40}$/F231$^{6.44}$ and the outward movement of TM6.

For FFA1, the AlphaFold-predicted inactive structure (FFA1-AF) also displayed subtle conformational differences compared to the active miniG_q-coupled FFA1 in the intracellular region, suggesting that FFA1-AF may adopt an active-like conformation (Fig. 6D). In addition, the DHA binding site in FFA1 cannot be definitively resolved at this time, making it difficult to speculate on the mechanism of

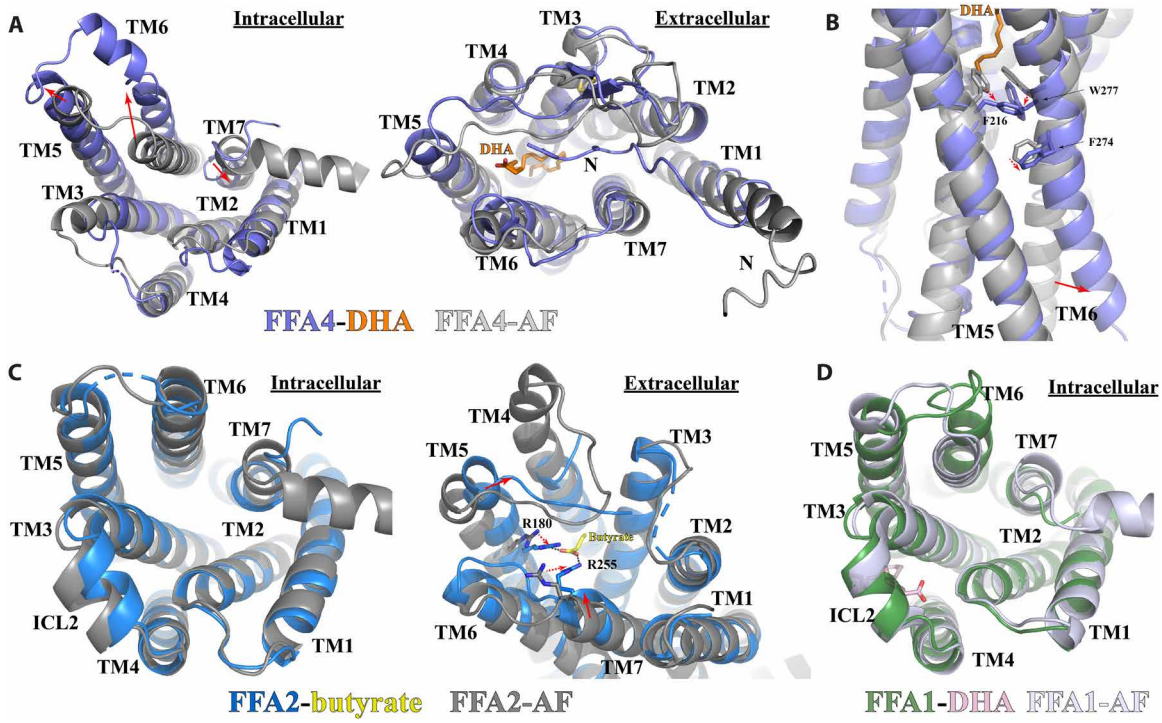

Fig. 6. Activation of FFAs. (**A**) Superimposition of the active DHA-bound FFA4 structure (slate) to the Alphafold-predicted inactive FFA4 structure FFA4-AF (light gray) viewed from the intracellular (left) and the extracellular (right) sides. (**B**) Residues involved in the receptor activation at the core region of FFA4. (**C**) Superimposition of the active butyrate-bound FFA2 structure (blue) to the Alphafold-predicted FFA2 structure FFA2-AF (dark gray) viewed from the intracellular (left) and the extracellular (right) sides. (**D**) Superimposition of the active DHA-bound FFA1 structure (green) to the Alphafold-predicted FFA1 structure FFA1-AF (light blue) viewed from the intracellular side. Red solid and dashed arrows represent conformational changes of TMs and individual residues, respectively, from the Alphafold-predicted structures to the active agonist-bound structures of FFA1, FFA2, and FFA4.

DHA-mediated activation of the receptor. If DHA binds to site 2, it is plausible that it stabilizes the helical structure of ICL2, similar to FFA1 ago-PAMs, to position it for interaction with G protein. Furthermore, DHA interacts directly with mini-$G_{\alpha q}$ at site 2 (Fig. 3B), indicating that it may also function to directly stabilize the interactions between FFA1 and G_q.

Shared and unique structural features of FFAs in the coupling of miniG$_q$

In our structural studies, we used the miniG$_q$ variant of the heterotrimeric G_q due to its greater propensity to form stable complexes with FFAs in our experiments. The C-terminal α5 helix of $G_{\alpha q}$, which serves as the major receptor binding site and a key factor in determining G protein coupling specificity (*64*), remains unchanged in miniG$_q$ (*33*). Therefore, we speculate that miniG$_q$ effectively mimics G_q in receptor coupling. The structural features of FFA1, FFA2, and FFA4 in their interactions with miniG$_q$, discussed below, may also be applicable to G_q.

In all of our structures, the orientations of α5 of miniG$_{\alpha q}$, the major binding site for FFAs, with respect to the three receptors were highly similar (Fig. 7A), with each receptor forming similar sets of interactions with the wavy hook region in α5 of mini-$G_{\alpha q}$ (fig. S15). As mentioned previously, we did not observe strong cryo-EM density for a putative helix 8 (H8) in any of the four structures. This suggests a disordered C-terminal region following TM7 in the three active receptors. However, we did observe direct interactions between the intracellular end of TM7 of each receptor and mini-$G_{\alpha q}$ (fig. S15), underscoring an important role of TM7 in G_q-coupling to each of the three receptors.

Although the interactions at the α5 helix of miniG$_{\alpha q}$ are similar across all three receptors, they engage in distinct interactions with other regions of miniG$_q$. Specifically, for FFA1 and FFA2, the ICL2 forms a helical structure to directly interact with miniG$_q$. In both receptors, a conserved tyrosine residue, Y114^{ICL2} in FFA1 and Y117^{ICL2} in FFA2, sticks toward the 7-TM core and forms a hydrogen bond with Y243 of mini-$G_{\alpha q}$ (Fig. 7B). In addition, on the opposite side of ICL2, V115^{ICL2} and L119^{ICL2} in FFA2, or L112^{ICL2} in FFA1, form hydrophobic interactions with mini-$G_{\alpha q}$ residues L34, V79, F228, and I235 (Fig. 7B). A similar set of hydrophobic interactions are also observed in the structures of the muscarinic acetylcholine receptor 1–G_{11} (*65*) and the serotonin receptor 5-HT_{2A}–miniG$_q$ complexes (*32*). However, in FFA4, a large part of ICL2 is disordered, and no similar hydrophobic interactions are observed (Fig. 1B). Nevertheless, R145^{ICL2} of FFA4 forms hydrogen bonds with the side chain of N239 and the main-chain carbonyl of I235 in the α5 helix of mini-$G_{\alpha q}$.

ICL3 is another region in the three receptors that interacts differently with miniG$_q$. In FFA1 and FFA2, ICL3 is positioned close to the wavy hook region of mini-$G_{\alpha q}$ (Fig. 7C). S212^{ICL3} in FFA1 forms a hydrogen bond with D233 of miniG$_{\alpha q}$, while R217^{ICL3} in FFA2 forms hydrogen bonds with D233 and Q237 of miniG$_{\alpha q}$ (Fig. 7C). However, in FFA4, TM5 extends by two helical turns compared to that of FFA1 and FFA2, resulting in ICL3 of FFA4 being positioned near the α4 helix and the α4-β6 loop of mini-$G_{\alpha q}$ (Fig. 7C). In this position, a segment of ICL3 in FFA4 adopts a helical structure. In the structure of FFA4-miniG$_q$ with DHA, residues Y247^{ICL3} and R254^{ICL3} of FFA4 form hydrogen bonds with the carbonyl groups of the mini-$G_{\alpha q}$ residues C211 and T202 backbone, respectively, while H251^{ICL3} of FFA4 forms a hydrogen bond with D199 of mini-$G_{\alpha q}$ (Fig. 7C). In addition,

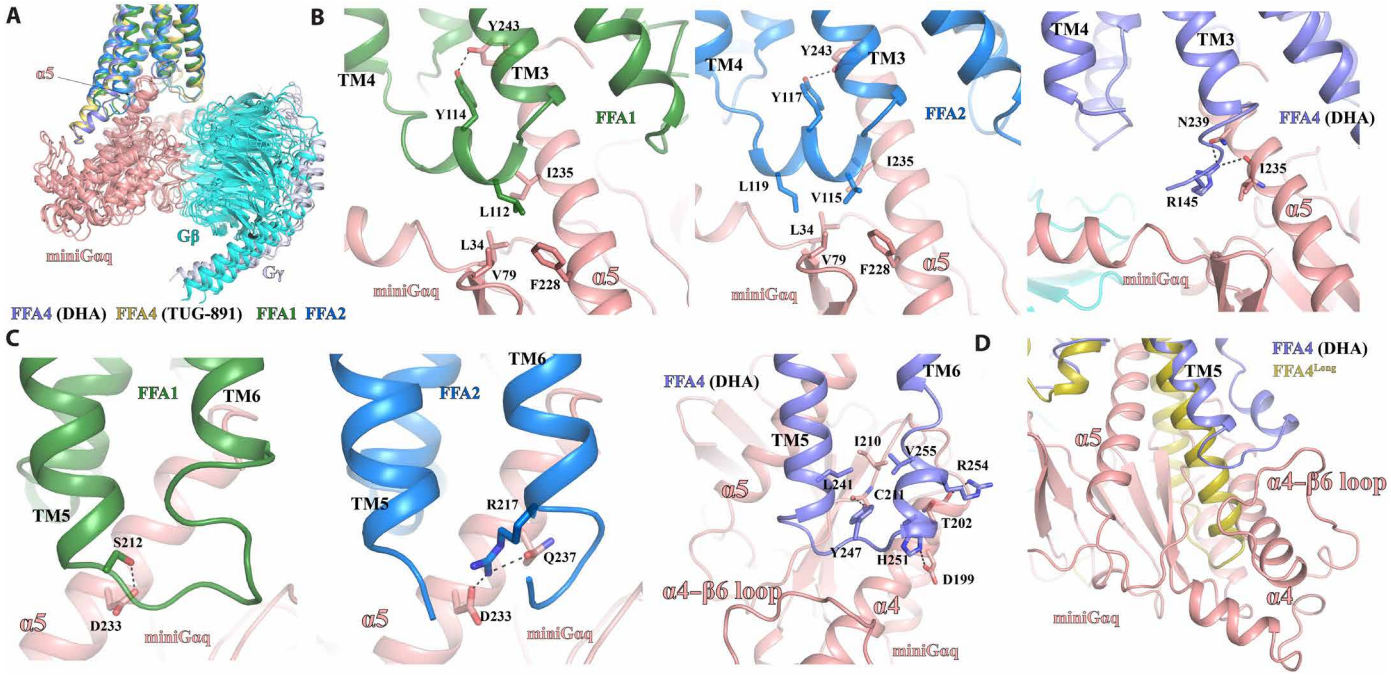

Fig. 7. Differences in the coupling of miniG$_q$ to FFAs. (**A**) Alignment of the structures of FFA1, FFA2, and FFA4 coupled with miniG$_q$ based on the receptors. (**B**) Differences in the interactions between miniG$_{\alpha q}$ and ICL2 of FFA1, FFA2, and FFA4. (**C**) Differences in the interactions between miniG$_{\alpha q}$ and ICL3 of FFA1, FFA2, and FFA4. (**D**) Superimposition of the AlphaFold-predicted structure of FFA4^{Long} to the structure of DHA-bound FFA4 coupled with miniG$_q$. MiniG$_{\alpha q}$, G_β, and G_γ subunits are colored salmon, cyan, and light blue, respectively. The colors of receptors and ligands are indicated in each panel.

V255^{ICL3} and L241$^{5.72}$ of FFA4 pack against I210 of mini-$G_{\alpha q}$ to form hydrophobic interactions (Fig. 7C).

The conformation of TM5 and ICL3 may account for the inability of FFA4^{Long}, the long form of FFA4, to induce $G_{q/11}$ signaling (*28*, *44*). FFA4^{Long} contains an insertion of 16 additional amino acids after Q232$^{5.63}$, which would further extend TM5 and potentially cause a severe steric clash with $G_{\alpha q}$. This clash is clearly visible if we superimpose the AlphaFold-predicted structure of FFA4^{Long} onto FFA4 in our structure (Fig. 7D).

DISCUSSION

Lipid recognition by FFAs

The different binding pockets in FFA2 and FFA4 for their endogenous SCFA and LCFA ligands clearly explain their preferences of free fatty acids (*25*, *26*). For FFA2, the small size of the binding pocket can only accommodate fatty acids with very short carbon chains. For FFA4, the binding pocket is much larger to accommodate LCFAs. It seems that the binding of LCFAs to FFA4 is largely driven by hydrophobic and π-π interactions since the mutations of polar residues near the carboxylate group of DHA did not significantly reduce the potency. SCFAs and medium-chain fatty acids with shorter carbon chains would result in less contacts with FFA4 and thus weaker potency for this receptor. On the other hand, despite the small observed binding pocket of FFA2, a number of larger synthetic FFA2 selective and orthosteric agonists have been identified and studied (*54*, *66*). It will be interesting in time to explore how FFA2 recognizes those large ligands.

The molecular basis for the lipid recognition of FFA1 is not readily clear based on our structural data. Site 1 (outer-leaflet pocket) and site 2 (inner-leaflet pocket) are both potential binding sites for DHA. Site 1 but not site 2 is rich in aromatic residues, which would favor the binding of PUFAs to FFA1. In addition, the two arginine residues R158$^{5.39}$ and R258$^{7.35}$ that are highly conserved in FFA1 to FFA3 form salt bridges with the carboxylate group of DHA (Fig. 3A). Mutations of these two residues significantly reduced the potency of DHA and another LCFA, γ-LA (*49*), further suggesting that site 1 is the primary binding site for DHA. However, the strong cryo-EM density in site 2 in our structure implies potential binding of DHA at this site as well, which may serve as the secondary binding site for DHA. Mutating Y114 to Ala in site 2 reduced the potency of DHA in arrestin-3 recruitment assays (Fig. 3C). However, we cannot exclude the possibility that Y114A does not affect DHA binding but rather the receptor activation. In addition, neither site 1 nor site 2 is large enough to accommodate the entire LCFAs with 20 carbons or more. Therefore, whether site 2 serves as a critical site for DHA binding in FFA1 and the mechanism for the selectivity of FFA1 for LCFAs over SCFAs still need further investigation.

Insight into drug development on FFAs

Over the past decade, a growing body of research has established the critical roles of FFAs in regulating metabolism and immunity. Studies have also provided evidence suggesting that agonists of FFA1, FFA2, and FFA4 have the potential to treat metabolic and inflammatory diseases. However, only the FFA1 agonist TAK-875 has completed all three phases of clinical trials. While TAK-875 showed promising results in improving glycemic control with a low risk of hypoglycemia in a phase 3 trial for diabetics, its further development was halted because of liver toxicity concerns (*67*). The underlying mechanism behind TAK-875's liver toxicity is still not entirely clear, with studies suggesting both direct hepatoxicity in an FFA1-dependent manner and metabolite-induced inhibition of hepatic transporters and mitochondrial respiration (*68*, *69*). To advance future drug development, it is crucial to determine whether TAK-875's liver toxicity is related to the activation of FFA1. The two well-defined ligand-binding sites in FFA1 offer opportunities for designing chemically diverse FFA1 ligands. If TAK-875's liver toxicity is caused by its metabolites, developing FFA1 agonists with new chemical scaffolds may provide a solution to this issue.

Compared to FFA1, the development of synthetic agonists for FFA2 and FFA4 toward the clinic has been limited, with only a small number of FFA2 ligands having been reported (*25*, *55*, *66*). By contrast, although the chemical diversity of synthetic FFA4 agonists has been rather limited (*11*), many analogs of the TUG-891 phenyl-propionic scaffold have been generated to improve the drug-like characteristics of this ligand for further assessment of their effects on the regulation of glucose homeostasis and other disease indications. An unexpected feature of the observed binding of TUG-891 and DHA within FFA4 was the absence of interaction between the carboxylate of agonists and R99$^{2.64}$. This had been widely anticipated on the basis of earlier modeling and mutagenesis studies. Initial modeling studies linked to the development of TUG-891 predicted an ionic interaction with this residue (*34*, *70*), and subsequent mutagenesis to R99Q confirmed the importance of this residue as the agonists were unable to activate this mutant. However, R99$^{2.64}$ does not interact directly with agonists but rather forms polar interactions with D30^{N} and E43$^{1.35}$ and a cation-π interaction with F304$^{7.36}$ (fig. S16). By doing so, R99$^{2.64}$ likely stabilizes the aromatic network of F25^{N}, F27^{N}, F28^{N}, F115$^{3.29}$, and F303$^{7.35}$ that forms a lid to the hydrophobic pocket. Mutation of R99$^{2.64}$, therefore, could break this aromatic network and lead to the exposure of the hydrophobic pocket to water destabilizing agonist binding.

Our work and another recent structural study (*44*) show that the carboxyl group of orthosteric ligands does not form specific interactions with FFA4; rather than that, the binding and ligand location is driven through hydrophobic interactions. The observed orthosteric binding pocket of FFA4, as revealed by our structures, offers valuable templates for designing new agonists for this receptor using computer-aided and artificial intelligence–driven drug design (CADD and AIDD) approaches. Another group of FFA4 agonists indeed do not contain the carboxyl group but have a sulfonamide or amide moiety as the central core linking the aromatic rings to form an L shape (*11*, *25*). When docking one such ligand, TUG-1197, into our FFA4 structure, the sulfonamide formed an H bond with T119$^{3.33}$, and its position overlapped with the O-linker of TUG-891 (Fig. 2D). TUG-1197 displayed markedly reduced potency at a T119A mutant of FFA4, and, although more modest in extent, the potency of TUG-891 was also reduced at T119A.

In contrast to FFA4, key residues of the orthosteric binding pocket of FFA2 were highly aligned with those predicted by previous mutagenesis studies (*52*). Given the challenges of developing potent and selective orthosteric FFA2 activators, there has been interest in the availability and design of selective allosteric agonists of FFA2 (*11*, *25*). Although nothing is currently known about their mode of binding, structural studies akin to those reported herein, and the ability to use computational tools to predict allosteric binding sites (*71*), it is likely that rapid progress could be made. This will also allow the development of more "drug-like" allosteric regulators of FFA2 and potentially also of the other SCFA receptor FFA3 where no high-potency

synthetic ligands are currently available (*25*), even as tool compounds, to better explore the biology and potential patho-physiological functions of this receptor. In both FFA2 and FFA4, we observed strong cryo-EM density in sites similar to the site 2 observed in FFA1, where we modeled phosphatidylinositol-4-phosphate and palmitic acid to fit the density (fig. S6). These observations suggest the possibility of developing allosteric modulators for FFA2 and FFA4 targeting this site. In the case of FFA2, we also observed a large cavity at the extracellular region right above the butyrate binding pocket. This cavity represents another potential allosteric site, which we refer to as site 3. If the structure of FFA2 is aligned with the structure of the muscarinic receptor M2R bound to a PAM named LY2119620, the allosteric site of LY2119620 highly overlaps with the putative site 3 in FFA2. As noted earlier, there are a number of FFA2 PAMs reported (*26*). They may target the two potential allosteric sites, site 2 and site 3, revealed by our structures. Further structural studies are needed to fully understand the molecular mechanisms of allosteric modulation in FFAs.

MATERIALS AND METHODS

Protein expression and purification

Human FFA1, FFA2, and FFA4 were cloned into the pFastBac vector (Thermo Fisher Scientific) with the LargeBit protein (*37*) fused to the C terminus of each receptor. The mini$G_{\alpha q}$ subunit (*35*) was cloned into the pFastBac vector. Human $G_{\beta 1}$ was fused with an N-terminal His_6-tag and a C-terminal HiBiT subunit connected with a 15–amino acid linker, which was cloned into pFastBac dual vector (Thermo Fisher Scientific) together with human $G_{\gamma 2}$.

The scFv16 was expressed in High Five cells using Bac-to-Bac expression system. To purify the protein, the cell supernatant was collected and loaded onto Ni–nitrilotriacetic acid resins. Following nickel affinity chromatography, the protein was further purified by size exclusion chromatography using a Superdex 200 Increase 100/300 GL column (GE Healthcare). The purified scFv16 fractions were pooled, concentrated, and stored at −80°C until use.

Free fatty acid receptors, mini$G_{\alpha q}$ and $G_{\beta 1\gamma 2}$ were coexpressed in Sf9 insect cells using Bac-to-Bac method. Sf9 cells were infected with three types of viruses at the ratio of 1:1:1 for 48 hours at 27°C. After being cultured for 48 hours, the cells were harvested and frozen at −80°C for further protein purification. Cell pellets were thawed in lysis buffer containing 20 mM Hepes (pH 7.5), 50 mM NaCl, 10 mM $MgCl_2$, 5 mM $CaCl_2$, leupeptin (2.5 μg/ml), and benzamidine (300 μg/ml). The complexes of DHA-bound or TUG-891-bound FFA4 with miniG_q were assembled on cell membranes by the addition of 10 μM DHA or 10 μM TUG-891. For the FFA1-miniG_q and FFA2-miniG_q complexes, 10 μM DHA and 1 mM butyrate were added to stimulate the formation of signaling complexes. To facilitate complex formation, apyrase (25 mU/ml; NEB) and 100 μM tris(2-carboxyethyl)phosphine (TCEP) were added and incubated at room temperature for 2 hours. The cell membranes were isolated by centrifugation at 25,000*g* for 40 min and then resuspended in solubilization buffer containing 20 mM Hepes (pH 7.5), 100 mM NaCl, 0.5% (w/v) lauryl maltose neopentylglycol (LMNG, Anatrace), 0.1% (w/v) cholesteryl hemisuccinate (CHS, Anatrace), 10% (v/v) glycerol, 10 mM $MgCl_2$, 5 mM $CaCl_2$, apyrase (12.5 mU/ml), 10 μM or 1 mM ligands, leupeptin (2.5 μg/ml), benzamidine (300 μg/ml), and 100 μM TECP at 4°C for 2 hours. The supernatant was collected by centrifugation at 25,000*g* for 1 hour and incubated with nickel Sepharose resin (GE Healthcare) at 4°C overnight. The resin was washed with buffer A containing 20 mM Hepes (pH 7.5), 100 mM NaCl, 0.05% (w/v) LMNG, 0.01% (w/v) CHS, 20 mM imidazole, and 10 μM or 1 mM ligands, leupeptin (2.5 μg/ml), benzamidine (300 μg/ml), and 100 μM TECP. The complex was eluted with buffer A containing 400 mM imidazole. The eluate was supplemented with 2 mM $CaCl_2$ and loaded onto anti-Flag M1 antibody resin. After wash, the complex was eluted in buffer A containing 5 mM EDTA and FLAG peptide (200 μg/ml) and concentrated using an Amicon Ultra Centrifugal Filter (molecular weight cut-off, 100 kDa). Last, a 1.3 M excess of scFv16 was added to the elution. The sample was then loaded onto a Superdex 200 Increase 10/300 column (GE Healthcare) preequilibrated with buffer containing 20 mM Hepes (pH 7.5), 100 mM NaCl, 0.00075% (w/v) LMNG, 0.00025% (w/v) glyco-diosgenin, 0.00015% (w/v) CHS, 10 μM or 1 mM ligands, and 100 μM TECP. Peak fractions of the complex were collected and concentrated to 5 to 10 mg/ml for cryo-EM experiments.

Cryo-EM sample preparation and data acquisition

For cryo-EM grid preparation of the DHA-FFA1-miniG_q complex, the butyrate-FFA2-miniG_q complex, and the TUG-891-FFA4-miniG_q complex, 3 μl of the protein complex was applied onto 300-mesh R1.2/1.3 UltrAuFoil Holey gold support films (Quantifoil). For cryo-EM grid preparation of the DHA-FFA4-miniG_q complex, 3 μl of the purified complex was applied to glow-discharged holey carbon grids (Quantifoil, Au300 R1.2/1.3). Grids were plunge-frozen in liquid ethane using Vitrobot Mark IV (Thermo Fischer Scientific).

All cryo-EM data were collected using a Titan Krios transmission electron microscope, equipped with a Gatan K3 Summit direct electron detector and an energy filter. For the TUG-891-FFA4-miniG_q complex and the DHA-FFA4-miniG_q complex, cryo-EM movie stacks were recorded with a nominal defocus setting in the range of −1.0 to −1.8 μm using SerialEM (*72*) with beam-tilt image-shift data collection strategy with a 3 × 3 pattern and three shots per hole. A total of 4968 movies for the dataset of the TUG-891-FFA4-miniG_q complex and 10,040 movies for three datasets of the DHA-FFA4-miniG_q complex were collected in the correlated double sampling super-resolution mode of the K3 camera at a nominal magnification of 105,000×, yielding a physical pixel size of 0.828 Å. Each stack was dose-fractionated to 52 frames with a total dose of 55 $e^-/Å^2$. For the DHA-FFA1-miniG_q complex and the butyrate-FFA2-miniG_q complex, 12,349 movies and 15,371 movies were collected with a nominal magnification of 105,000× using the SerialEM software running a 3 × 3 image shift pattern and three shots per hole. Micrographs were recorded using a super-resolution mode at a calibrated pixel size of 0.826 Å and a defocus range of −0.8 to −2.5 μm. Each stack was dose-fractionated to 50 frames with a total dose of 61.6 $e^-/Å^2$.

Data processing, 3D reconstruction, and modeling building

Image stacks were subjected to patch motion correction using cryoSPARC (*73*). The contrast transfer function (CTF) parameters were calculated using the patch CTF estimation tool in cryoSPARC.

For the TUG-891-FFA4-miniG_q complex, a total of 4,928,436 particles were auto-picked and then subjected to two-dimensional (2D) classification to discard poorly defined particles. After ab initio reconstruction and heterogeneous refinement, 391,203 particles were subjected to nonuniform refinement and local refinement, which generated a map with an indicated global resolution of 3.06 Å at a Fourier shell correlation (FSC) of 0.143. For the DHA-FFA4-miniG_q

complex, a threshold of CTF fit resolution of more than 4 Å was used to exclude low-quality micrographs. Each dataset was processed separately with autopicking and 2D classification. After ab initio reconstruction and heterogeneous refinement, 380,284 particles were subjected to nonuniform refinement and local refinement, which generated a map with an indicated global resolution of 3.14 Å at an FSC of 0.143. For the DHA-FFA1-miniG_q complex and the butyrate-FFA2-miniG_q complex, a threshold of CTF fit resolution of more than 4 Å was used to exclude low-quality micrographs, respectively. A total of 7,942,319 particles and 12,559,721 particles were auto-picked and then subjected to 2D classification to discard bad particles. After ab initio reconstruction and heterogeneous refinement, 305,318 particles and 393,952 particles were subjected to nonuniform refinement and local refinement, which generated a map with an indicated global resolution of 3.39 and 3.07 Å at an FSC of 0.143. To further improve the quality of the receptor of FFA1, local refinement focusing on the receptor was performed in cryoSPARC. Local resolutions of density maps were estimated in cryoSPARC.

The Alphafold-predicted structures of FFA1, FFA2, and FFA4 served as initial models for model rebuilding and refinement against the EM map. The model was initially docked into the EM density map using Chimera (*74*). This step was followed by iterative manual adjustment and rebuilding in COOT (*75*). Real space refinement and Rosetta refinement were then carried out using Phenix programs (*76*). To validate the model, MolProbity was used for assessing its structural statistics (*77*).

For the preparation of structural figures, Chimera and PyMOL (https://pymol.org/2/) were used. The final refinement statistics can be found in table S1. To evaluate the degree of overfitting during the refinement process, the final model was refined against one of the half-maps. The resulting map versus model FSC curves were compared with both half-maps and the full model. Surface coloring of the density map was achieved using UCSF Chimera (*74*).

Mutagenesis and BRET-based arrestin-2 and -3 recruitment assays

All cell culture reagents and 3,3',5,5' tetramethylbenzidine (TMB substrate) solution were from Thermo Fisher Scientific (Loughborough, UK). Polyethylenimine (PEI) [linear poly(vinyl alcohol) (molecular weight ~25,000 Dalton)] was from Polysciences (Warrington, PA). Molecular biology enzymes and reagents were from Promega (Southampton, UK). TUG-891 and DHA were from Tocris Biosciences (Bristol, UK). TUG-1197 was a gift from T. Ulven (University of Copenhagen).

A plasmid encoding the short isoform of human FFA4 fused at its C terminus to enhanced yellow fluorescent protein (eYFP) and containing an N-terminal FLAG epitope tag was generated as described previously (*78*). Mutations were introduced into the FFA4 sequence using the QuikChange method (Stratagene), and, in all cases, the presence of the mutation was verified through sequencing. The NanoLuc luciferase coding sequence was subcloned after polymerase chain reaction amplification (using primers designed to add Xba I and Not I sites) into either an arrestin-2 or arrestin-3-pcDNA3 plasmid (arrestin-2/3-NLuc).

All FFA4 constructs were expressed in HEK293T cells, which were maintained in Dulbecco's modified Eagle's medium (DMEM) supplemented with L-glutamine (0.292 g/liter), 1% penicillin-streptomycin mixture, and 10% heat-inactivated fetal bovine serum at 37°C in a 5% CO_2 humidified atmosphere. To express receptors, cells were transfected using PEI. The day before transfection, 2×10^6 cells were plated into 10-cm dishes. Plasmid DNA was then combined with PEI (in 1:6 ratio) in 500 µl of 150 mM NaCl, thoroughly mixed, then incubated for 10 min at room temperature. Cell medium was changed, and the DNA–PEI mixture was added to the medium in a dropwise manner.

For bioluminescence resonance energy transfer (BRET) assays, HEK293T cells were seeded in 10-cm^2 dishes and transiently co-transfected with wild type or each of the indicated FFA4 mutants, each with a FLAG epitope tag engineered into the N-terminal domain and eYFP fused at its C terminus, and arrestin-2-NLuc or arrestin-3-NLuc at a 50:1 ratio, respectively, using PEI. Control cells were transfected with arrestin-2-NLuc or arrestin-3-NLuc only. After 24 hours, cells were detached by incubating with trypsin-EDTA and seeded at 5×10^4 cells per well in poly-D-lysine–coated white 96-well plates and then incubated overnight at 37°C. Cells were washed once with prewarmed (37°C) Hanks' balanced salt solution (HBSS) and incubated in HBSS for 30 to 60 min at 37°C. The luciferase substrate coelenterazine-h (Prolume) was added to a final concentration of 5 µM and the plate was incubated for 10 min at 37°C protected from light. Agonists were added at the relevant concentrations in triplicate and the plate was incubated for a further 5 min at 37°C; then, the emissions at 475 and 535 nm were read on a PHERAstar FS. The net BRET ratio was calculated as follows: [(signal 535 nm/signal 475 nm) − (signal nanoluc luciferase only 535 nm/signal nanoluc luciferase only 475 nm)] * 1000. Direct measures of eYFP fluorescence determined total expression levels of FFA4 receptor variants.

Generation of Flp-In T-REx 293 cell lines expressing wild type and mutated FFA4 variants

Flp-In T-REx 293 cells (Invitrogen) were maintained in DMEM without sodium pyruvate, supplemented with 10% (v/v) fetal bovine serum, 1% penicillin-streptomycin mixture, and blasticidin (10 µg/ml; Invivogen, Toulouse, France) at 37°C in a 5% CO_2 humidified atmosphere. To generate Flp-In T-REx cells able to express in an inducible manner the various FFA4-eYFP receptor constructs, cells were transfected with a mixture containing the desired cDNA in pcDNA5/FRT/TO vector and pOG44 vector (1:9) by using PEI (1 mg/ml) (MW 25,000). Cells were plated until 60 to 80% confluent then transfected with 8 µg of required plasmid DNA and PEI (ratio 1:6 DNA/PEI), diluted in 150 mM NaCl (pH 7.4). After incubation at room temperature for 10 min, the mixture was added to cells. After 48 hours, the medium was changed to medium supplemented with hygromycin B (200 µg/ml; Invivogen, Toulouse, France) to initiate the selection of stably transfected cells. After isolation of resistant cells, expression of the appropriate construct from the Flp-In TREx locus was induced by treatment with doxycycline (100 ng/ml) for 24 hours.

Ca^{2+} mobilization assays

Flp-In T-REx 293 cells engineered to express the desired form of FFA4 (see above) were plated 75,000 cells per well in black 96-well plates with clear bottoms. Cells were then treated with doxycycline (100 ng/ml) to induce receptor expression and maintained overnight at 37°C and 5% CO_2 before their use. For the assay, cells were first prelabeled for 45 min with the calcium-sensitive dye Fura2-AM (Sigma-Aldrich) then washed and maintained in HBSS. Fura-2 fluorescent emission at 510 nm resulting from excitation at 340 or 380 nm was then

monitored using a FlexStation II plate reader (Molecular Devices). Basal fluorescence was measured for 16 s; test compounds were then added, and fluorescence was measured for an additional 74 s. The background subtracted peak 340/380 ratio obtained following compound addition was then used to plot concentration-response data.

Cell surface enzyme-linked immunosorbent assays

Cell surface expression of receptors was quantified by live-cell enzyme-linked immunosorbent assay. The same cotransfected cells used for BRET studies were seeded at 5×10^4 cells per well in poly-D-lysine–coated clear 96-well plates and incubated overnight at 37°C. Cells were incubated with primary antibody (rabbit polyclonal anti-FLAG 1:1000) in culture medium for 1 hour at 37°C and then washed three times with DMEM-Hepes and incubated with secondary antibody (horseradish peroxidase–sheep anti-rabbit immunoglobulin G, 1:10,000) for 1 hour at 37°C protected from light. Cells were then washed three times with warmed (37°C) phosphate-buffered saline (PBS). Last, PBS was removed, and room temperature TMB substrate (100 µl per well) was added. The plate was incubated for 15 min at room temperature protected from light, and then the absorbance at 620 nm was read on a POLARStar Omega.

MD simulations

The cryo-EM structures of FFA4 bound to TUG-891 or DHA, FFA1 bound to DHA, and FFA2 bound to butyrate were prepared using Schrodinger Maestro 2021-3 (*79*).The input for membrane simulations was prepared using the CHARMM-GUI server (*80*–*88*). The receptor was placed in the 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine bilayer membrane sized 100 Å by 100 Å, and the 22.5-Å solvent layer on each side of the membrane.The minimization, equilibration, and production were done using the PMEMD program from the Amber20 package (AMBER 20) (*89*–*91*). The production was done in three replicas of 1 µs for each biosystem. The snapshots of the simulations were saved every 50,000 steps, or 0.1 ns, in simulations.The results of simulations were analyzed using CPPTRAJ from the Amber20 package and MDAnalysis (*92*, *93*). The residue-ligand interaction energy was calculated using the "namdenergy.tcl" script v1.6 for VMD (*94*) and the NAMD2 program with cutoff and switch parameters of 9 and 7.5 Å, respectively (*95*, *96*). Forcefield parameters were taken from the AMBER parameter file used for simulations. Energy calculation was done for every 10th snapshot, or every 1.0 ns of simulation. Details of modeling methods can be found in the Supplementary Materials.

Supplementary Materials

This PDF file includes:

Supplemental Methods
Figs. S1 to S16
Tables S1 to S4
References

REFERENCES AND NOTES

1. E. Alvarez-Curto, G. Milligan, Metabolism meets immunity: The role of free fatty acid receptors in the immune system. *Biochem. Pharmacol.* **114**, 3–13 (2016).
2. I. Kimura, A. Ichimura, R. Ohue-Kitano, M. Igarashi, Free fatty acid receptors in health and disease. *Physiol. Rev.* **100**, 171–210 (2020).
3. D. Y. Oh, S. Talukdar, E. J. Bae, T. Imamura, H. Morinaga, W. Fan, P. Li, W. J. Lu, S. M. Watkins, J. M. Olefsky, GPR120 is an omega-3 fatty acid receptor mediating potent anti-inflammatory and insulin-sensitizing effects. *Cell* **142**, 687–698 (2010).
4. C. P. Briscoe, M. Tadayyon, J. L. Andrews, W. G. Benson, J. K. Chambers, M. M. Eilert, C. Ellis, N. A. Elshourbagy, A. S. Goetz, D. T. Minnick, P. R. Murdock, H. R. Sauls Jr., U. Shabon, L. D. Spinage, J. C. Strum, P. G. Szekeres, K. B. Tan, J. M. Way, D. M. Ignar, S. Wilson, A. I. Muir, The orphan G protein-coupled receptor GPR40 is activated by medium and long chain fatty acids. *J. Biol. Chem.* **278**, 11303–11311 (2003).
5. J. L. McCarville, G. Y. Chen, V. D. Cuevas, K. Troha, J. S. Ayres, Microbiota metabolites in health and disease. *Annu. Rev. Immunol.* **38**, 147–170 (2020).
6. Y. Itoh, Y. Kawamata, M. Harada, M. Kobayashi, R. Fujii, S. Fukusumi, K. Ogi, M. Hosoya, Y. Tanaka, H. Uejima, H. Tanaka, M. Maruyama, R. Satoh, S. Okubo, H. Kizawa, H. Komatsu, F. Matsumura, Y. Noguchi, T. Shinohara, S. Hinuma, Y. Fujisawa, M. Fujino, Free fatty acids regulate insulin secretion from pancreatic β cells through GPR40. *Nature* **422**, 173–176 (2003).
7. Z. Li, Z. Zhou, L. Zhang, Current status of GPR40/FFAR1 modulators in medicinal chemistry (2016-2019): A patent review. *Expert Opin. Ther. Pat.* **30**, 27–38 (2020).
8. A. D. Mancini, V. Poitout, The fatty acid receptor FFA1/GPR40 a decade later: How much do we know? *Trends Endocrinol. Metab.* **24**, 398–407 (2013).
9. A. D. Mancini, V. Poitout, GPR40 agonists for the treatment of type 2 diabetes: Life after 'TAKing' a hit. *Diabetes Obes. Metab.* **17**, 622–629 (2015).
10. G. Milligan, E. Alvarez-Curto, B. D. Hudson, R. Prihandoko, A. B. Tobin, FFA4/GPR120: Pharmacology and therapeutic opportunities. *Trends Pharmacol. Sci.* **38**, 809–821 (2017).
11. G. Carullo, S. Mazzotta, M. Vega-Holm, F. Iglesias-Guerra, J. M. Vega-Perez, F. Aiello, A. Brizzi, GPR120/FFAR4 pharmacology: Focus on agonists in type 2 diabetes mellitus drug discovery. *J. Med. Chem.* **64**, 4312–4332 (2021).
12. S. Halder, S. Kumar, R. Sharma, The therapeutic potential of GPR120: A patent review. *Expert Opin. Ther. Pat.* **23**, 1581–1590 (2013).
13. G. Bianchini, C. Nigro, A. Sirico, R. Novelli, I. Prevenzano, C. Miele, F. Beguinot, A. Aramini, A new synthetic dual agonist of GPR120/GPR40 induces GLP-1 secretion and improves glucose homeostasis in mice. *Biomed. Pharmacother.* **139**, 111613 (2021).
14. X. Zhang, M. J. Macielag, GPR120 agonists for the treatment of diabetes: A patent review (2014 present). *Expert Opin. Ther. Pat.* **30**, 729–742 (2020).
15. C. Cartoni, K. Yasumatsu, T. Ohkuri, N. Shigemura, R. Yoshida, N. Godinot, J. le Coutre, Y. Ninomiya, S. Damak, Taste preference for fatty acids is mediated by GPR40 and GPR120. *J. Neurosci.* **30**, 8376–8382 (2010).
16. M. H. Kim, S. G. Kang, J. H. Park, M. Yanagisawa, C. H. Kim, Short-chain fatty acids activate GPR41 and GPR43 on intestinal epithelial cells to promote inflammatory responses in mice. *Gastroenterology* **145**, 396–406.e1–10 (2013).
17. P. D. Cani, A. Everard, T. Duparc, Gut microbiota, enteroendocrine functions and metabolism. *Curr. Opin. Pharmacol.* **13**, 935–940 (2013).
18. K. M. Maslowski, A. T. Vieira, A. Ng, J. Kranich, F. Sierro, D. Yu, H. C. Schilter, M. S. Rolph, F. Mackay, D. Artis, R. J. Xavier, M. M. Teixeira, C. R. Mackay, Regulation of inflammatory responses by gut microbiota and chemoattractant receptor GPR43. *Nature* **461**, 1282–1286 (2009).
19. I. Kimura, K. Ozawa, D. Inoue, T. Imamura, K. Kimura, T. Maeda, K. Terasawa, D. Kashihara, K. Hirano, T. Tani, T. Takahashi, S. Miyauchi, G. Shioi, H. Inoue, G. Tsujimoto, The gut microbiota suppresses insulin-mediated fat accumulation via the short-chain fatty acid receptor GPR43. *Nat. Commun.* **4**, 1829 (2013).
20. Y. Zhao, F. Chen, W. Wu, M. Sun, A. J. Bilotta, S. Yao, Y. Xiao, X. Huang, T. D. Eaves-Pyles, G. Golovko, Y. Fofanov, W. D'Souza, Q. Zhao, Z. Liu, Y. Cong, GPR43 mediates microbiota metabolite SCFA regulation of antimicrobial peptide expression in intestinal epithelial cells via activation of mTOR and STAT3. *Mucosal Immunol.* **11**, 752–762 (2018).
21. K. H. Antunes, J. L. Fachi, R. de Paula, E. F. da Silva, L. P. Pral, A. A. Dos Santos, G. B. M. Dias, J. E. Vargas, R. Puga, F. Q. Mayer, F. Maito, C. R. Zarate-Blades, N. J. Ajami, M. R. Sant'Ana, T. Candreva, H. G. Rodrigues, M. Schmiele, M. T. P. Silva Clerici, J. L. Proenca-Modena, A. T. Vieira, C. R. Mackay, D. Mansur, M. T. Caballero, J. Marzec, J. Li, X. Wang, D. Bell, F. P. Polack, S. R. Kleeberger, R. T. Stein, M. A. R. Vinolo, A. P. D. de Souza, Microbiota-derived acetate protects against respiratory syncytial virus infection through a GPR43-type 1 interferon response. *Nat. Commun.* **10**, 3273 (2019).
22. L. B. Bindels, E. M. Dewulf, N. M. Delzenne, GPR43/FFA2: Physiopathological relevance and therapeutic prospects. *Trends Pharmacol. Sci.* **34**, 226–232 (2013).
23. S. Kim, Y. M. Kim, Y. S. Kwak, A novel therapeutic target, GPR43; where it stands in drug discovery. *Arch. Pharm. Res.* **35**, 1505–1509 (2012).
24. I. Kimura, D. Inoue, K. Hirano, G. Tsujimoto, The SCFA receptor GPR43 and energy metabolism. *Front. Endocrinol. (Lausanne)* **5**, 85 (2014).
25. G. Milligan, B. Shimpukade, T. Ulven, B. D. Hudson, Complex pharmacology of free fatty acid receptors. *Chem. Rev.* **117**, 67–110 (2017).
26. M. Grundmann, E. Bender, J. Schamberger, F. Eitner, Pharmacology of free fatty acid receptors and their allosteric modulators. *Int. J. Mol. Sci.* **22**, 1763 (2021).
27. K. Moore, Q. Zhang, N. Murgolo, T. Hosted, R. Duffy, Cloning, expression, and pharmacological characterization of the GPR120 free fatty acid receptor from cynomolgus monkey: Comparison with human GPR120 splice variants. *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* **154**, 419–426 (2009).
28. S.-J. Watson, A. J. Brown, N. D. Holliday, Differential signaling by splice variants of the human free fatty acid receptor GPR120. *Mol. Pharmacol.* **81**, 631–642 (2012).
29. J. Lu, N. Byrne, J. Wang, G. Bricogne, F. K. Brown, H. R. Chobanian, S. L. Colletti, J. Di Salvo, B. Thomas-Fowlkes, Y. Guo, D. L. Hall, J. Hadix, N. B. Hastings, J. D. Hermes, T. Ho,

A. D. Howard, H. Josien, M. Kornienko, K. J. Lumb, M. W. Miller, S. B. Patel, B. Pio, C. W. Plummer, B. S. Sherborne, P. Sheth, S. Souza, S. Tummala, C. Vonrhein, M. Webb, S. J. Allen, J. M. Johnston, A. B. Weinglass, S. Sharma, S. M. Soisson, Structural basis for the cooperative allosteric activation of the free fatty acid receptor GPR40. *Nat. Struct. Mol. Biol.* **24**, 570–577 (2017).
30. J. D. Ho, B. Chau, L. Rodgers, F. Lu, K. L. Wilbur, K. A. Otto, Y. Chen, M. Song, J. P. Riley, H. C. Yang, N. A. Reynolds, S. D. Kahl, A. P. Lewis, C. Groshong, R. E. Madsen, K. Conners, J. P. Lineswala, T. Gheyi, M. D. Saflor, M. R. Lee, J. Benach, K. A. Baker, C. Montrose-Rafizadeh, M. J. Genin, A. R. Miller, C. Hamdouchi, Structural basis for GPR40 allosteric agonism and incretin stimulation. *Nat. Commun.* **9**, 1645 (2018).
31. A. Srivastava, J. Yano, Y. Hirozane, G. Kefala, F. Gruswitz, G. Snell, W. Lane, A. Ivetac, K. Aertgeerts, J. Nguyen, A. Jennings, K. Okada, High-resolution structure of the human GPR40 receptor bound to allosteric agonist TAK-875. *Nature* **513**, 124–127 (2014).
32. K. Kim, T. Che, O. Panova, J. F. DiBerto, J. Lyu, B. E. Krumm, D. Wacker, M. J. Robertson, A. B. Seven, D. E. Nichols, B. K. Shoichet, G. Skiniotis, B. L. Roth, Structure of a hallucinogen-activated Gq-coupled 5-HT_{2A} serotonin receptor. *Cell* **182**, 1574–1588.e19 (2020).
33. R. Nehmé, B. Carpenter, A. Singhal, A. Strege, P. C. Edwards, C. F. White, H. Du, R. Grisshammer, C. G. Tate, Mini-G proteins: Novel tools for studying GPCRs in their active conformation. *PLOS ONE* **12**, e0175642 (2017).
34. B. Shimpukade, B. D. Hudson, C. K. Hovgaard, G. Milligan, T. Ulven, Discovery of a potent and selective GPR120 agonist. *J. Med. Chem.* **55**, 4511–4515 (2012).
35. Q. Wan, N. Okashah, A. Inoue, R. Nehme, B. Carpenter, C. G. Tate, N. A. Lambert, Mini G protein probes for active G protein-coupled receptors (GPCRs) in live cells. *J. Biol. Chem.* **293**, 7466–7473 (2018).
36. J. G. Meyerowitz, M. J. Robertson, X. Barros-Alvarez, O. Panova, R. M. Nwokonko, Y. Gao, G. Skiniotis, The oxytocin signaling complex reveals a molecular switch for cation dependence. *Nat. Struct. Mol. Biol.* **29**, 274–281 (2022).
37. J. Duan, D. D. Shen, X. E. Zhou, P. Bi, Q. F. Liu, Y. X. Tan, Y. W. Zhuang, H. B. Zhang, P. Y. Xu, S. J. Huang, S. S. Ma, X. H. He, K. Melcher, Y. Zhang, H. E. Xu, Y. Jiang, Cryo-EM structure of an activated VIP1 receptor-G protein complex revealed by a NanoBiT tethering strategy. *Nat. Commun.* **11**, 4121 (2020).
38. A. Koehl, H. Hu, S. Maeda, Y. Zhang, Q. Qu, J. M. Paggi, N. R. Latorraca, D. Hilger, R. Dawson, H. Matile, G. F. X. Schertler, S. Granier, W. I. Weis, R. O. Dror, A. Manglik, G. Skiniotis, B. K. Kobilka, Structure of the μ-opioid receptor-Gi protein complex. *Nature* **558**, 547–552 (2018).
39. L. Wang, D. Yao, R. Deepak, H. Liu, Q. Xiao, H. Fan, W. Gong, Z. Wei, C. Zhang, Structures of the human PGD(2) receptor CRTH2 reveal novel mechanisms for ligand recognition. *Mol. Cell* **72**, 48–59.e4 (2018).
40. H. Liu, R. Deepak, A. Shiriaeva, C. Gati, A. Batyuk, H. Hu, U. Weierstall, W. Liu, L. Wang, V. Cherezov, H. Fan, C. Zhang, Molecular basis for lipid recognition by the prostaglandin D_2 receptor CRTH2. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2102813118 (2021).
41. J. A. Ballesteros, H. Weinstein, [19] Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G protein-coupled receptors. *J. Neurosci. Methods* **25**, 366–428 (1995).
42. B. D. Hudson, B. Shimpukade, G. Milligan, T. Ulven, The molecular basis of ligand interaction at free fatty acid receptor 4 (FFA4/GPR120). *J. Biol. Chem.* **289**, 20345–20358 (2014).
43. C. M. Azevedo, K. R. Watterson, E. T. Wargent, S. V. Hansen, B. D. Hudson, M. A. Kepczynska, J. Dunlop, B. Shimpukade, E. Christiansen, G. Milligan, C. J. Stocker, T. Ulven, Non-acidic free fatty acid receptor 4 agonists with antidiabetic activity. *J. Med. Chem.* **59**, 8868–8878 (2016).
44. C. Mao, P. Xiao, X. N. Tao, J. Qin, Q. T. He, C. Zhang, S. C. Guo, Y. Q. Du, L. N. Chen, D. D. Shen, Z. S. Yang, H. Q. Zhang, S. M. Huang, Y. H. He, J. Cheng, Y. N. Zhong, P. Shang, J. Chen, D. L. Zhang, Q. L. Wang, M. X. Liu, G. Y. Li, Y. Guo, H. E. Xu, C. Wang, C. Zhang, S. Feng, X. Yu, Y. Zhang, J. P. Sun, Unsaturated bond recognition leads to biased signal in a fatty acid receptor. *Science* **380**, eadd6220 (2023).
45. M. Lückmann, M. Trauelsen, T. M. Frimurer, T. W. Schwartz, Structural basis for GPCR signaling by small polar versus large lipid metabolites-discovery of non-metabolite ligands. *Curr. Opin. Cell Biol.* **63**, 38–48 (2020).
46. H. Liu, H. R. Kim, R. Deepak, L. Wang, K. Y. Chung, H. Fan, Z. Wei, C. Zhang, Orthosteric and allosteric action of the C5a receptor antagonists. *Nat. Struct. Mol. Biol.* **25**, 472–481 (2018).
47. X. Liu, J. Kaindl, M. Korczynska, A. Stossel, D. Dengler, M. Stanek, H. Hubner, M. J. Clark, J. Mahoney, R. A. Matt, X. Xu, K. Hirata, B. K. Shoichet, R. K. Sunahara, B. K. Kobilka, P. Gmeiner, An allosteric modulator binds to a conformational hub in the β2 adrenergic receptor. *Nat. Chem. Biol.* **16**, 749–755 (2020).
48. C. Yabuki, H. Komatsu, Y. Tsujihata, R. Maeda, R. Ito, K. Matsuda-Nagasumi, K. Sakuma, K. Miyawaki, N. Kikuchi, K. Takeuchi, Y. Habata, M. Mori, A novel antidiabetic drug, fasiglifam/TAK-875, acts as an ago-allosteric modulator of FFAR1. *PLOS ONE* **8**, e76280 (2013).
49. P. Kumari, A. Inoue, K. Chapman, P. Lian, D. M. Rosenbaum, Molecular mechanism of fatty acid activation of FFAR1. *Proc. Natl. Acad. Sci. U.S.A.* **120**, e2219569120 (2023).
50. F. Yang, C. Mao, L. Guo, J. Lin, Q. Ming, P. Xiao, X. Wu, Q. Shen, S. Guo, D. D. Shen, R. Lu, L. Zhang, S. Huang, Y. Ping, C. Zhang, C. Ma, K. Zhang, X. Liang, Y. Shen, F. Nan, F. Yi, V. C. Luca, J. Zhou, C. Jiang, J. P. Sun, X. Xie, X. Yu, Y. Zhang, Structural basis of GPBAR activation and bile acid recognition. *Nature* **587**, 499–504 (2020).
51. L. A. Stoddart, N. J. Smith, G. Milligan, International Union of Pharmacology. LXXI. Free fatty acid receptors FFA1, -2, and -3: Pharmacology and pathophysiological functions. *Pharmacol. Rev.* **60**, 405–417 (2008).
52. L. A. Stoddart, N. J. Smith, L. Jenkins, A. J. Brown, G. Milligan, Conserved polar residues in transmembrane domains V, VI, and VII of free fatty acid receptor 2 and free fatty acid receptor 3 are required for the binding and function of short chain fatty acids. *J. Biol. Chem.* **283**, 32913–32924 (2008).
53. E. Sergeev, A. H. Hansen, S. K. Pandey, A. E. MacKenzie, B. D. Hudson, T. Ulven, G. Milligan, Non-equivalence of key positively charged residues of the free fatty acid 2 receptor in the recognition and function of agonist versus antagonist ligands. *J. Biol. Chem.* **291**, 303–317 (2016).
54. B. D. Hudson, M. E. Due-Hansen, E. Christiansen, A. M. Hansen, A. E. Mackenzie, H. Murdoch, S. K. Pandey, R. J. Ward, R. Marquez, I. G. Tikhonova, T. Ulven, G. Milligan, Defining the molecular basis for the first potent and selective orthosteric agonists of the FFA2 free fatty acid receptor. *J. Biol. Chem.* **288**, 17296–17312 (2013).
55. B. D. Hudson, E. Christiansen, I. G. Tikhonova, M. Grundmann, E. Kostenis, D. R. Adams, T. Ulven, G. Milligan, Chemically engineering ligand selectivity at the free fatty acid receptor 2 based on pharmacological variation between species orthologs. *FASEB J.* **26**, 4951–4965 (2012).
56. E. Sergeev, A. H. Hansen, D. Bolognini, K. Kawakami, T. Kishi, J. Aoki, T. Ulven, A. Inoue, B. D. Hudson, G. Milligan, A single extracellular amino acid in free fatty acid receptor 2 defines antagonist species selectivity and G protein selection bias. *Sci. Rep.* **7**, 13741 (2017).
57. B. D. Hudson, I. G. Tikhonova, S. K. Pandey, T. Ulven, G. Milligan, Extracellular ionic locks determine variation in constitutive activity and ligand potency between species orthologs of the free fatty acid receptors FFA2 and FFA3. *J. Biol. Chem.* **287**, 41195–41209 (2012).
58. S. Filipek, Molecular switches in GPCRs. *Curr. Opin. Struct. Biol.* **55**, 114–120 (2019).
59. Q. Zhou, D. Yang, M. Wu, Y. Guo, W. Guo, L. Zhong, X. Cai, A. Dai, W. Jang, E. I. Shakhnovich, Z. J. Liu, R. C. Stevens, N. A. Lambert, M. M. Babu, M. W. Wang, S. Zhao, Common activation mechanism of class A GPCRs. *eLife* **8**, e50279 (2019).
60. G. Pandy-Szekeres, J. Caroli, A. Mamyrbekov, A. A. Kermani, G. M. Keseru, A. J. Kooistra, D. E. Gloriam, GPCRdb in 2023: State-specific structure models using AlphaFold2 and new ligand resources. *Nucleic Acids Res.* **51**, D395–D402 (2023).
61. L. Heo, M. Feig, Multi-state modeling of G-protein coupled receptors at experimental accuracy. *Proteins* **90**, 1873–1885 (2022).
62. W. I. Weis, B. K. Kobilka, The molecular basis of G protein-coupled receptor activation. *Annu. Rev. Biochem.* **87**, 897–919 (2018).
63. S. G. Rasmussen, H. J. Choi, J. J. Fung, E. Pardon, P. Casarosa, P. S. Chae, B. T. Devree, D. M. Rosenbaum, F. S. Thian, T. S. Kobilka, A. Schnapp, I. Konetzki, R. K. Sunahara, S. H. Gellman, A. Pautsch, J. Steyaert, W. I. Weis, B. K. Kobilka, Structure of a nanobody-stabilized active state of the $β_2$ adrenoceptor. *Nature* **469**, 175–180 (2011).
64. W. M. Oldham, H. E. Hamm, Heterotrimeric G protein activation by G-protein-coupled receptors. *Nat. Rev. Mol. Cell Biol.* **9**, 60–71 (2008).
65. S. Maeda, Q. Qu, M. J. Robertson, G. Skiniotis, B. K. Kobilka, Structures of the M1 and M2 muscarinic acetylcholine receptor/G-protein complexes. *Science* **364**, 552–557 (2019).
66. A. H. Hansen, E. Sergeev, D. Bolognini, R. R. Sprenger, J. H. Ekberg, C. S. Ejsing, C. J. McKenzie, E. Rexen Ulven, G. Milligan, T. Ulven, Discovery of a potent thiazolidine free fatty acid receptor 2 agonist with favorable pharmacokinetic properties. *J. Med. Chem.* **61**, 9534–9550 (2018).
67. K. Kaku, K. Enya, R. Nakaya, T. Ohira, R. Matsuno, Efficacy and safety of fasiglifam (TAK-875), a G protein-coupled receptor 40 agonist, in Japanese patients with type 2 diabetes inadequately controlled by diet and exercise: A randomized, double-blind, placebo-controlled, phase III trial. *Diabetes Obes. Metab.* **17**, 675–681 (2015).
68. M. Kim, G. J. Gu, Y. S. Koh, S. H. Lee, Y. R. Na, S. H. Seok, K. M. Lim, Fasiglifam (TAK-875), a G protein-coupled receptor 40 (GPR40) agonist, may induce hepatotoxicity through reactive oxygen species generation in a GPR40-dependent manner. *Biomol. Ther. (Seoul)* **26**, 599–607 (2018).
69. M. A. Otieno, J. Snoeys, W. Lam, A. Ghosh, M. R. Player, A. Pocai, R. Salter, D. Simic, H. Skaggs, B. Singh, H. K. Lim, Fasiglifam (TAK-875): Mechanistic investigation and retrospective identification of hazards for drug induced liver injury. *Toxicol. Sci.* **163**, 374–384 (2018).
70. I. G. Tikhonova, Application of GPCR structures for modelling of free fatty acid receptors. *Handb. Exp. Pharmacol.* **236**, 57–77 (2017).
71. A. Ciancetta, A. K. Gill, T. Ding, D. S. Karlov, G. Chalhoub, P. J. McCormick, I. G. Tikhonova, Probe confined dynamic mapping for G protein-coupled receptor allosteric site prediction. *ACS Cent. Sci.* **7**, 1847–1862 (2021).

72. D. N. Mastronarde, Automated electron microscope tomography using robust prediction of specimen movements. *J. Struct. Biol.* **152**, 36–51 (2005).
73. A. Punjani, J. L. Rubinstein, D. J. Fleet, M. A. Brubaker, cryoSPARC: Algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).
74. E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin, UCSF Chimera—A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
75. P. Emsley, K. Cowtan, Coot: Model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).
76. P. D. Adams, P. V. Afonine, G. Bunkoczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L. W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, P. H. Zwart, PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010).
77. V. B. Chen, W. B. Arendall III, J. J. Headd, D. A. Keedy, R. M. Immormino, G. J. Kapral, L. W. Murray, J. S. Richardson, D. C. Richardson, MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 12–21 (2010).
78. B. D. Hudson, B. Shimpukade, A. E. Mackenzie, A. J. Butcher, J. D. Pediani, E. Christiansen, H. Heathcote, A. B. Tobin, T. Ulven, G. Milligan, The pharmacology of TUG-891, a potent and selective agonist of the free fatty acid receptor 4 (FFA4/GPR120), demonstrates both potential opportunity and possible challenges to therapeutic agonism. *Mol. Pharmacol.* **84**, 710–725 (2013).
79. Schrödinger Release 2022-3: Maestro (Schrödinger LLC, 2022).
80. S. Jo, T. Kim, W. Im, Automated builder and database of protein/membrane complexes for molecular dynamics simulations. *PLOS ONE* **2**, e880 (2007).
81. S. Jo, J. B. Lim, J. B. Klauda, W. Im, CHARMM-GUI membrane builder for mixed bilayers and its application to yeast membranes. *Biophys. J.* **97**, 50–58 (2009).
82. E. L. Wu, X. Cheng, S. Jo, H. Rui, K. C. Song, E. M. Davila-Contreras, Y. Qi, J. Lee, V. Monje-Galvan, R. M. Venable, J. B. Klauda, W. Im, CHARMM-GUI membrane builder toward realistic biological membrane simulations. *J. Comput. Chem.* **35**, 1997–2004 (2014).
83. J. Lee, X. Cheng, J. M. Swails, M. S. Yeom, P. K. Eastman, J. A. Lemkul, S. Wei, J. Buckner, J. C. Jeong, Y. Qi, S. Jo, V. S. Pande, D. A. Case, C. L. Brooks III, A. D. MacKerell Jr., J. B. Klauda, W. Im, CHARMM-GUI input generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM simulations using the CHARMM36 additive force field. *J. Chem. Theory Comput.* **12**, 405–413 (2016).
84. S. Kim, J. Lee, S. Jo, C. L. Brooks III, H. S. Lee, W. Im, CHARMM-GUI ligand reader and modeler for CHARMM force field generation of small molecules. *J. Comput. Chem.* **38**, 1879–1886 (2017).
85. J. Lee, D. S. Patel, J. Stahle, S. J. Park, N. R. Kern, S. Kim, J. Lee, X. Cheng, M. A. Valvano, O. Holst, Y. A. Knirel, Y. Qi, S. Jo, J. B. Klauda, G. Widmalm, W. Im, CHARMM-GUI membrane builder for complex biological membrane simulations with glycolipids and lipoglycans. *J. Chem. Theory Comput.* **15**, 775–786 (2019).
86. J. Lee, M. Hitzenberger, M. Rieger, N. R. Kern, M. Zacharias, W. Im, CHARMM-GUI supports the amber force fields. *J. Chem. Phys.* **153**, 035103 (2020).
87. S. Jo, T. Kim, V. G. Iyer, W. Im, CHARMM-GUI: A web-based graphical user interface for CHARMM. *J. Comput. Chem.* **29**, 1859–1865 (2008).
88. B. R. Brooks, C. L. Brooks III, A. D. Mackerell Jr., L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A. R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R. W. Pastor, C. B. Post, J. Z. Pu, M. Schaefer, B. Tidor, R. M. Venable, H. L. Woodcock, X. Wu, W. Yang, D. M. York, M. Karplus, CHARMM: The biomolecular simulation program. *J. Comput. Chem.* **30**, 1545–1614 (2009).
89. R. Salomon-Ferrer, A. W. Gotz, D. Poole, S. Le Grand, R. C. Walker, Routine microsecond molecular dynamics simulations with AMBER on GPUs. 2. Explicit solvent particle mesh ewald. *J Chem Theory Comput* **9**, 3878–3888 (2013).
90. A. W. G. Scott Le Grand, R. C. Walker, SPFP: Speed without compromise—A mixed precision model for GPU accelerated molecular dynamics simulations. *Comput. Phys. Commun.* **184**, 374–380 (2013).
91. K. B. D. A. Case, I. Y. Ben-Shalom, S. R. Brozell, D. S. Cerutti, T. E. Cheatham, III, V. W. D. Cruzeiro,, R. E. D. T. A. Darden, G. Giambasu, M. K. Gilson, H. Gohlke, A. W. Goetz,R Harris, S. Izadi, S. A. Izmailov, K. Kasavajhala, A. Kovalenko, R. Krasny, T. Kurtzman, T. S. Lee, S. LeGrand, P. Li, C. Lin, J. Liu,, R. L. T. Luchko, V. Man, K. M. Merz, Y. Miao, O. Mikhailovskii, G. Monard, H. Nguyen, A. Onufriev, F., S. P. Pan, R. Qi, D. R. Roe, A. Roitberg, *C. Sagui*, S. Schott-Verdugo, J. Shen, C. L. Simmerling, N. R., J. S. Skrynnikov, J. Swails, R. C. Walker, J. Wang, L. Wilson, R. M. Wolf, X. Wu, Y. Xiong, Y. Xue,, P. A. K. D. M. York, *AMBER 2020* (University of California, San Francisco, 2020).
92. N. Michaud-Agrawal, E. J. Denning, T. B. Woolf, O. Beckstein, MDAnalysis: A toolkit for the analysis of molecular dynamics simulations. *J. Comput. Chem.* **32**, 2319–2327 (2011).
93. R. Gowers, M. Linke, J. Barnoud, T. Reddy, M. Melo, S. Seyler, J. Domański, D. Dotson, S. Buchoux, I. Kenney, O. Beckstein, in *Proceedings of the 15th Python in Science Conference* SciPy 2016, Austin, Texas, USA, pp. 98–105.
94. W. Humphrey, A. Dalke, K. Schulten, VMD: Visual molecular dynamics. *J. Mol. Graph.* **14**, 27–38 (1996).
95. J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale, K. Schulten, Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26**, 1781–1802 (2005).
96. J. C. Phillips, D. J. Hardy, J. D. C. Maia, J. E. Stone, J. V. Ribeiro, R. C. Bernardi, R. Buch, G. Fiorin, J. Henin, W. Jiang, R. McGreevy, M. C. R. Melo, B. K. Radak, R. D. Skeel, A. Singharoy, Y. Wang, B. Roux, A. Aksimentiev, Z. Luthey-Schulten, L. V. Kale, K. Schulten, C. Chipot, E. Tajkhorshid, Scalable molecular dynamics on CPU and GPU architectures with NAMD. *J. Chem. Phys.* **153**, 044130 (2020).
97. G. Pandy-Szekeres, C. Munk, T. M. Tsonkov, S. Mordalski, K. Harpsoe, A. S. Hauser, A. J. Bojarski, D. E. Gloriam, GPCRdb in 2018: Adding GPCR structure models and ligands. *Nucleic Acids Res.* **46**, D440–D446 (2018).
98. Rambaut, A. (2010) FigTree v1.3.1. Institute of Evolutionary Biology, University of Edinburgh, Edinburgh. http://tree.bio.ed.ac.uk/software/figtree/.
99. R. A. Friesner, J. L. Banks, R. B. Murphy, T. A. Halgren, J. J. Klicic, D. T. Mainz, M. P. Repasky, E. H. Knoll, M. Shelley, J. K. Perry, D. E. Shaw, P. Francis, P. S. Shenkin, Glide: A new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J. Med. Chem.* **47**, 1739–1749 (2004).
100. T. A. Halgren, R. B. Murphy, R. A. Friesner, H. S. Beard, L. L. Frye, W. T. Pollard, J. L. Banks, Glide: A new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. *J. Med. Chem.* **47**, 1750–1759 (2004).
101. R. A. Friesner, R. B. Murphy, M. P. Repasky, L. L. Frye, J. R. Greenwood, T. A. Halgren, P. C. Sanschagrin, D. T. Mainz, Extra precision glide: Docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. *J. Med. Chem.* **49**, 6177–6196 (2006).
102. Schrodinger Release 2022-3: Glide, Schrodinger, LLC, New York, NY, 2022.
103. Schrodinger Release 2022-3: Prime, Schrodinger, LLC, New York, NY, 2022.
104. M. P. Jacobson, D. L. Pincus, C. S. Rapp, T. J. Day, B. Honig, D. E. Shaw, R. A. Friesner, A hierarchical approach to all-atom protein loop prediction. *Proteins* **55**, 351–367 (2004).
105. X. He, V. H. Man, B. Ji, X. Q. Xie, J. Wang, Calculate protein-ligand binding affinities with the extended linear interaction energy method: Application on the Cathepsin S set in the D3R Grand Challenge 3. *J. Comput. Aided Mol. Des.* **33**, 105–117 (2019).
106. C. UniProt, UniProt: The Universal Protein Knowledgebase in 2023. *Nucleic Acids Res.* **51**, D523–D531 (2023).
107. M. H. Olsson, C. R. Sondergaard, M. Rostkowski, J. H. Jensen, PROPKA3: Consistent treatment of internal and surface residues in empirical pKa predictions. *J. Chem. Theory Comput.* **7**, 525–537 (2011).
108. C. R. Søndergaard, M. H. Olsson, M. Rostkowski, J. H. Jensen, Improved treatment of ligands and coupling effects in empirical calculation and rationalization of pKa values. *J. Chem. Theory Comput.* **7**, 2284–2295 (2011).
109. M. A. Lomize, I. D. Pogozheva, H. Joo, H. I. Mosberg, A. L. Lomize, OPM database and PPM web server: Resources for positioning of proteins in membranes. *Nucleic Acids Res.* **40**, D370–D376 (2012).
110. C. Tian, K. Kasavajhala, K. A. A. Belfon, L. Raguette, H. Huang, A. N. Migues, J. Bickel, Y. Wang, J. Pincay, Q. Wu, C. Simmerling, ff19SB: Amino-acid-specific protein backbone parameters trained against quantum mechanics energy surfaces in solution. *J. Chem. Theory Comput.* **16**, 528–552 (2020).
111. C. J. Dickson, R. C. Walker, I. R. Gould, Lipid21: Complex lipid membrane simulations with amber. *J. Chem. Theory Comput.* **18**, 1726–1736 (2022).
112. X. He, V. H. Man, W. Yang, T. S. Lee, J. Wang, A fast and high-quality charge model for the next generation general AMBER force field. *J. Chem. Phys.* **153**, 114502 (2020).
113. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935 (1983).
114. H. J. C. Berendsen, J. P. M. Postma, W. F. Vangunsteren, A. Dinola, J. R. Haak, Molecular-dynamics with coupling to an external bath. *J. Chem. Phys.* **81**, 3684–3690 (1984).
115. R. W. Pastor, B. R. Brooks, A. Szabo, An analysis of the accuracy of Langevin and molecular dynamics algorithms. *Mol. Phys.* **65**, 1409–1419 (1988).

Acknowledgments: We thank the cryo-EM facility at the University of Pittsburgh partly supported by grants S10 OD025009 (Krios) and S10 OD019995 (Falcon 2/3 camera) from the National Institutes of Health (NIH) in the United States. We thank S. Chakrapani for oversight of the cryo-EM core facility at Case Western Reserve University. **Funding:** This work was supported by the NIH grant R35GM128641 to C.Z., the Medical Research Council (UK) grant MR/X010198/1 to G.M., and the Biotechnology and Biological Sciences Research Council (UK) grants BB/R001480/1 and BB/S000453/1 to G.M. and BB/R007101/1 to I.G.T. A.-A.G.'s PhD study is supported by the MSCA COFUND CITI-GENS Programme funded by the E.U. Horizon 2020 research and innovation programme (grant agreement no. 945231). This project made use of

computational time on Kelvin-2 supported by Engineering and Physical Sciences Research Council (EPSRC) (grant no. EP/T022175/1 and EP/W03204X/1) and ARCHER2 granted via the U.K. High-End Computing Consortium for Biomolecular Simulation, HECBioSim (www.hecbiosim.ac.uk), supported by EPSRC (grant no. EP/R029407/1 and EP/W03204X/1). **Author contributions:** Conceptualization: C.Z., G.M., I.G.T., and X.Z. Methodology: X.Z., A.-A.G., L.J., C.Z., G.M., and I.G.T. Investigation: X.Z., A.-A.G., L.J., and K.L. Visualization: X.Z., A.-A.G., L.J., and C.Z. Supervision: C.Z., G.M., and I.G.T. Writing—original draft: C.Z. Writing—review and editing: C.Z., G.M., I.G.T., X.Z., and A.-A.G. **Competing interests:** G.M. is cofounder and a director of both Caldan Therapeutics (https://caldantherapeutics.com/) and KelticPharmaTherapeutics (https://keltic-pharma.com/), which both have interests in the development of FFA4 activators. All other authors declare that they have no competing interests. **Data and materials availability:** The 3D cryo-EM density maps of FFA signaling complexes have been deposited in the Electron Microscopy Data Bank under the accession numbers EMD-41013 and EMD-41014 for the DHA-FFA1-miniGq complex, EMD-41010 for the butyrate-FFA2-miniGq complex, and EMD-41007 and EMD-41008 for the DHA- and TUG-891-FFA4-miniGq complexes, respectively. Atomic coordinates for the atomic models have been deposited in the Protein Data Bank (PDB) under the accession numbers 8T3V, 8T3S, 8T3Q, and 8T3Q for the DHA-FFA1-miniGq complex, the butyrate-FFA2-miniGq complex, and the DHA- and TUG-891-FFA4-miniGq complexes, respectively. The plasmids used in the structural biology studies can be provided by C.Z. pending scientific review and a completed material transfer agreement. Requests for the plasmids should be submitted to C.Z. (chengzh@pitt.edu). All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials.

Submitted 14 June 2023
Accepted 5 December 2023
Published 10 January 2024
10.1126/sciadv.adj2384

ScienceAdvances

Structural basis for the ligand recognition and signaling of free fatty acid receptors

Xuan Zhang, Abdul-Akim Guseinov, Laura Jenkins, Kunpeng Li, Irina G. Tikhonova, Graeme Milligan, and Cheng Zhang

Sci. Adv. **10** (2), eadj2384. DOI: 10.1126/sciadv.adj2384

Science Advances (ISSN 2375-2548) is published by the American Association for the Advancement of Science. 1200 New York Avenue NW, Washington, DC 20005. The title *Science Advances* is a registered trademark of AAAS.