Activation of free fatty acid receptor 1 (GPR40) by synthetic partial and full agonists occur via distinct allosteric sites. A crystal structure of GPR40-TAK-875 complex revealed the allosteric site for the partial agonist. Here we report the 2.76-Å crystal structure of human GPR40 in complex with a synthetic full agonist, compound 1, bound to the second allosteric site. Unlike TAK-875, which acts as a Gαq-coupled partial agonist, compound 1 is a dual Gαq and Gαs-coupled full agonist. Compound 1 binds in the lipid-rich region of the receptor near intracellular loop 2 (ICL2), in which the stabilization of ICL2 by the ligand is likely the primary mechanism for the enhanced G protein activities. The endogenous free fatty acid (FFA), γ-linolenic acid, can be computationally modeled in this site. Both γ-linolenic acid and compound 1 exhibit positive cooperativity with TAK-875, suggesting that this site could also serve as a FFA binding site.
World Health Organization reported in 2014 that diabetes is affecting an estimated 422 million adults globally, a nearly four-fold increase from 108 million cases in 1980. In the United States about 1.7 million new cases of diabetes are reported each year—a rate that could lead to one out of every three adults having diabetes by 2050. Type 2 diabetes mellitus (T2DM) accounts for ~90% of all diabetes, and the basis for treating T2DM is improvement of glycemic control. While numerous clinical treatments of T2DM are available today, many are associated with negative side effects such as hypoglycemia and weight gain; thus, there remains a demand for new anti-diabetic medicines with improved metabolic profiles. In the last decade, free fatty acid receptor 1 (FFAR1 or GPR40) has emerged as an attractive diabetes therapeutic target with both glucose-lowering and weight loss potential. A G-protein coupled receptor that responds to dietary long-chain free fatty acids (FFAs), GPR40 modulates FFA-stimulated insulin secretion in pancreatic β cells and incretin secretion in enteroendocrine cells in a glucose dependent manner. Among the clinically studied synthetic GPR40 agonists, TAK-875 advanced to phase II clinical trial before its termination due to toxicity. While clinical data demonstrated that TAK-875 has potent anti-diabetic effects with minimal incidence of hypoglycemia and weight gain, in vitro and in vivo studies showed that TAK-875 functioned as a partial agonist and minimally effected incretin release from enteroendocrine cell. Meanwhile, AM-1638 was discovered among a novel series of agonists that exerts full agonism on GPR40 and stimulates incretin secretion. This class of full agonists also acts allosterically with the endogenous FFAs, but binds to a topographically distinct site from TAK-875. Among the studied synthetic GPR40 agonists, TAK-875 advanced to phase II clinical trial before its termination due to toxicity. While clinical data demonstrated that TAK-875 has potent anti-diabetic effects with minimal incidence of hypoglycemia and weight gain, in vitro and in vivo studies showed that TAK-875 functioned as a partial agonist and minimally effected incretin release from enteroendocrine cell. Meanwhile, AM-1638 was discovered among a novel series of agonists that exerts full agonism on GPR40 and stimulates incretin secretion. This class of full agonists also acts allosterically with the endogenous FFAs, but binds to a topographically distinct site from TAK-875. Moreover, unlike TAK-875 which activates only the Gαs/Ca2+ pathway, these full agonists allow GPR40 to couple to both the Gαs/Ca2+ and Gαi/AC2+ signaling pathways, potentially explaining their unique activity as potent incretin secretagogues. Based upon the disparate pharmacology of these compounds, a widely held hypothesis suggests that aside from the orthosteric site for glucose metabolism (Fig. 2a-d). In intraperitoneal glucose tolerance test (IPGTT), compound 1 dose-dependently increases insulin plasma levels and improves glucose metabolism (Fig. 2a-d). The absence of increased insulin levels by compound 1 in GPR40 knock-out (KO) mice indicates that the ability of compound 1 to increase insulin plasma levels is GPR40-mediated (Fig. 2c). Compound 1 also exhibits similar dose-dependent reduction of blood glucose in oral glucose tolerance test (OGTT) that is GPR40-mediated (Fig. 2e-g). In

![Fig. 1 Ligand binding and functional assays of compound 1.](image-url)
addition, oral administration of compound 1 demonstrates robust increase in incretins (GLP-1 and GIP) that is GPR40-mediated (Fig. 3a–d). Compound 1 does not stimulate peptide YY (PYY) release in vivo (Fig. 3e) suggesting that increased incretin secretion is mechanism mediated instead of degranulation. In contrast, TAK-875 does not compete with AM-1638 for the A2 site; it only activates the Gpr120 KO receptor and has very little incretin stimulation activity (Supplementary Fig. 1a–d).

Crystal structure of the GPR40-compound 1 complex. The structure described here was obtained using the same stabilized human GPR40 construct as the GPR40-TAK-875 structure (using the Protein Data Bank accession number 4PHU) in later discussion. This GPR40 receptor has a T4 lysozyme (T4L) protein inserted into the third intracellular loop (ICL3), as well as four point mutations that increase expression and thermal stability. These modifications likely constrain GPR40 in an inactive conformation, which could explain the reduced relative binding of [3H]-AM-1638 to the stabilized receptor vs. wild type (Supplementary Fig. 3b). The purified GPR40-compound 1 complex was crystallized in lipidic cubic phase. Only a single crystal was needed to obtain a complete X-ray data set to 2.76 Å using synchrotron radiation, with apparently minimal or no radiation damage as judged by minimal change in the relative B factors of diffraction frames over data collection time (Table 1; Supplementary Fig. 7).

The structure shows that compound 1 is highly lipid-exposed and does not participate in any packing interaction in the crystal (Supplementary Fig. 2a). This extra-helical allosteric binding site (A2) is a lipid-facing elongated hydrophobic pocket defined by transmembrane helices 3–5 (H3, H4, H5) and ICL2 (Figs. 4b and 5a). Some other examples of ligand binding to a location entirely outside of the helical bundle are provided by the P2Y1 receptor (P2Y1R) in complex with BPTU (Fig. 7). In contrast to the effect of compound 1 as a GPR40 full agonist, both BPTU and MK-0893 are allosteric antagonists for their respective receptors and do not share the same binding position as compound 1 (Fig. 6).

The A2 site occupied by compound 1 is distinct from the site occupied by TAK-875, and in 4PHU, the A2 site is occupied by a
monolein molecule, but the omit map electron density for monolein is weak and discontinuous (Fig. 4a, b). The A2 site is located in the non-polar region of the receptor, so it was unexpected to find compound 1, a ligand with a carboxylic acid moiety, binding at this location given the penalty of desolvation upon binding. However, the acidic moiety of compound 1 is fully satisfied by four hydrogen bond interactions: three from the side chains of Tyr 442.42, Tyr 114ICL2, Ser 1234.42 and one from water on the cytoplasmic side of the receptor (Fig. 5b) (superscripts indicate Ballesteros–Weinstein numbers21). Single-mutant study of Y44F, Y114F, and S123A in GPR40 shows that a simple change of removing the polar hydroxyl group in each of the three residues resulted in significant mitigation of the downstream Gαs stimulation by compound 1 (Fig. 6c) as a result of reduced compound 1 binding at the A2 site (Supplementary Fig. 3b, c). This demonstrates the importance of these three polar residues in engaging a ligand with acidic moiety in this hydrophobic environment. In addition, the section of compound 1 starting from the benzofuran moiety to the distal anisole (methoxybenzene) provides additional affinity by maintaining good van der Waals (vdW) contact against H3 (Fig. 5a).

**Structural comparison with the GPR40-TAK-875 complex.** Comparing the A1 site occupied by TAK-875 in 4PHU, we observe that the entrance into the A1 site of our structure cannot accommodate TAK-875 due to movement of the upper half of H3 and H4 (Supplementary Fig. 6a–c), which is affected by crystal packing (Supplementary Fig. 2b); therefore, we believe that this slightly collapsed A1 site in our structure should not be interpreted as a conformation change resulting from the binding of compound 1 in the A2 site. In 4PHU, Arg 1835.39 and Arg 2587.35 are the key basic residues that bind the acidic moiety of TAK-875; in our structure, both arginines hydrogen bond to the residues on extracellular loop 2 (ECL2), thereby pulling the ECL2 towards H6 and H7 (Supplementary Fig. 6d, e).

Another significant difference between 4PHU and our structure lies in the conformation of ICL2. In our structure, ICL2 adopts a short helix, but it is disordered in 4PHU (Fig. 4a, b). This difference in ICL2 structure is likely due to the hydrogen

---

**Fig. 3 GLP-1 and GIP secretion by compound 1.**

**a** compound 1 significantly increases plasma total GLP-1 concentrations up to 8 h.

**b** compound 1 significantly increases plasma active GIP concentrations up to 8 h.

**c, d** compound 1 has no effect on plasma total GLP-1 (c) and active GIP (d) concentrations in GPR40 KO mice.

**e** Both compound 1 and TAK-875 exert no effect on plasma PYY level. Incretin values were log2 transformed and analyzed using ANOVA. *p < 0.05 vs vehicle by ANOVA with repeated measures followed by Dunnett’s comparison.
The interaction between ICL2 of GPR40 and Gαq. The stabilization of ICL2 into a short helix by compound 1 (and presumably other A2 binders) could explain the enhanced G-protein coupling to the Gαq/cAMP pathway by compound 1 (and presumably other A2 binders) could explain the enhanced G-protein coupling to the Gαq/cAMP pathway by compound 1 and Tyr 114ICL2 (Fig. 5b). In the absence of this polar interaction, this helical conformation of ICL2 is unfavorable due to the fact that the polar side chain of Tyr 114ICL2 cannot form a hydrogen bond in this hydrophobic region of site A2.

(briefly reviewed by Zheng et al.23). Therefore, We hypothesize that full agonist binding at the A2 site may impart the Gαq in addition to the Gαq signaling due to the stabilization of the ICL2 of GPR40 in the conformation favorable for Gαq coupling. However, since our structure is in the inactive conformation, other unobserved conformational changes directed by A2 binders may also influence the additional Gαq activity.

Discussion
To determine if the A2 site in GPR40 is also present in other GPCRs, we examined sequence alignments of four lipid GPCR subgroups (FFAR1–4, sphingosine-1-phosphate receptor 1–5, lysophosphatidic acid receptor 1–6, and prostaglandin E receptor 1–4). While none of these lipid receptors possess the same three amino acids in the equivalent positions of Tyr 442.42, Tyr 114ICL2, and Ser 1234.42 in GPR40, there are other polar residues present in this region that could still interact with a charged ligand (Supplementary Fig. 4a–c). Indeed, even though P2Y1R is not a lipid GPCR, in the X-ray crystal structure of P2Y1R complexed with BPTU19 (Protein Data Bank accession number 4XNV), a cholesterol hemisuccinate (CHS) molecule was modeled in this location with relatively weak omit map electron density (Supplementary Fig. 5c). In this P2Y1R structure, the hydrophobic portion of the CHS maintains some vdW contact in the pocket and its acidic moiety is partially satisfied by hydrogen bonds with the side chains of Tyr 892.42 and His 1483.49 (Supplementary Fig. 5a, b). It is less certain if there is any polar interaction from the receptor to the other oxygen atom of the ligand’s acidic moiety due to the fact that part of the ICL2 is disordered in the structure, and Ser 1513.52 (another nearby polar amino acid) is too far away from forming a productive hydrogen bond with the acidic moiety (~4 Å) (Supplementary Fig. 5b). Notably still, Tyr 892.42 of P2Y1R corresponds to Tyr 442.42 in GPR40, and His 1483.49 of P2Y1R corresponds to Gly 1033.49 in GPR40 or to Asp in the DRY-motif in most GPCRs (reviewed in ref. 29). This demonstrates that the relatively sparse
presence of polar amino acids in this non-polar region of the receptor may still provide specific polar interaction with a charged ligand. Further study is required to assess if such allosteric binding site exists for P2Y1R.

The total number of carbon atoms from the carboxyl group to the o-methyl of compound 1 and AM-1638 can be counted as 17 atoms, similar to the length of the 18-carbon FFA native ligand: γ-linolenic acid (γ-LA) (Fig. 7a–c). Supported by the observation of a lipid molecule (monoolein) in the same location in 4PHU (Fig. 4a), we hypothesized that the A2 site could also be a second lipid binding site. To assess the likelihood of an 18-carbon FFA being able to bind at site A2, we modeled both stearic acid (C18) and γ-LA (C18:3) into the site and estimated the free energy for each ligand going from the unbound to the bound state. While both ligands fit the A2 site, the bound state of γ-LA appeared to show more favorable enthalpy of binding, with two olefins forming hydrogen bonds with backbone oxygens of the protein in a very non-polar environment. In addition, γ-LA, as expected due to the partial unsaturation, appeared to access far fewer conformations in the unbound state than stearic acid, resulting in a much smaller entropy loss to adopt the bound state conformation. Binding of γ-LA, which has been reported to have a better GPR40 affinity than stearic acid, appears to originate from better enthalpy and entropy due to the partial unsaturation. γ-LA is a full agonist of GPR40 and has positive cooperativity with TAK-875 in the Ca2+ mobilization and insulin secretion assays. It is plausible that the allosteric activity observed between TAK-875 and γ-LA is mediated by cooperativity between sites A1 and A2 as γ-LA can be modeled in A2 (Fig. 7d). Moreover, we also observed positive allosteric cooperativity between TAK-875 and compound 1 (α = 2.1; β = 14) as well as enhanced thermal stability when both ligands are present (Fig. 8a, b). After several rounds of crystallizations trials of GPR40 complexed with both TAK-875 and compound 1, we were not able to obtain crystals. A reasonable model of this two-ligand ternary complex can be presented where there is no overlap of the two allosteric binding sites (Fig. 8c). Indeed, more studies remain needed to extensively characterize the allosteric relationship of the A1 and A2 sites and how their allostery modulates the downstream Gαq and Gαi signaling pathways. It is also worth speculating that GPR40 may not be a receptor that has a dedicated orthosteric site as previously hypothesized but a receptor with two allosteric binding sites that can both be accessed by endogenous FFAs.

Our findings, as well as other recently reported agonist-bound GPR40 structures provide structural evidence that agonists binding distinct allosteric sites of the same receptor generate biased G-protein signaling. This allosteric mechanism expands our understanding of basic pharmacology and could provide insight into the function of other members of the GPCR family. In addition, these results may be useful for the rational design of better GPR40 agonists. The possibility that oral administration of a GPR40 agonist can lead to increases in circulating levels of the therapeutic biomolecules insulin and GLP-1 has the potential to improve future therapies for T2DM.

**Methods**

**Ligand-binding assay of the wild type and mutant receptors.** Crude cell surface membranes were prepared from human embryonic kidney 293 (HEK 293) cells stably transfected (for the wild type) or transiently transfected (for the mutants and corresponding wild type) with full length recombinant human GPR40 (GPR40) cDNA, using differential centrifugation methods.

**A1 Site Binding.** 10 µL of compound diluted in 100% DMSO and 90 µL of Assay Buffer (50 mM Tris-HCl, pH 7.5, 5 mM CaCl2, 5 mM MgCl2, 0.1% w/v fatty acid-free BSA) were added to a deep 96-well polypolypropylene assay plate (Beckman Coulter). In total 200 µL of [3H]-TAK-875 (52 Ci/mmol, Quotient Bioresearch Radiocchemicals, Ltd.; 5 nM final concentration) and 200 µL of GPR40 membranes (5 µg/well), both diluted in Assay Buffer, were added to the assay plate, followed by a 1 min shake and a 2 h incubation at room temperature (22 °C).
**Fig. 6** Functional study on GPR40 site A2 mutants and intracellular loop 2 (ICL2) mutants. **a** Structural superposition of GPR40-compound 1 with β2AR•G protein complex (PDB: 3SN6) in ribbon representation (Cα RMSD = 1.31 Å). GPR40-compound 1 structure is in cyan; compound 1 is shown as orange stick; β2AR is in red; Gαs is in yellow; Gβi is in purple; Gy is in green. The vdW interaction surface of the A2 binding site is shown as mesh in silver. The region circled is magnified in **b**. An important hydrophobic contact between the ICL2 of β2AR and Gαs. The ICL2 is labeled. vdW interaction surface formed by H41, V217, F219, F376, and C379 of the Gαs is shown as mesh and colored according to the electrostatic charge of the residues (red: negative; blue: positive; white: neutral). **c** cAMP accumulation assays in the presence of compound 1 (C) or TAK-875 (D). Levels of cAMP are measured as relative luminescence units (RLU). Y114F, S123A, Y44F are site A2 mutations. L112A and L112F are ICL2 mutations involved in G protein complex (PDB: 3SN6) in ribbon representation (Cα β circled is magnified). GPR40-compound 1 structure is in red; Gαs is in yellow; Gβi is in purple; Gy is in green. The vdW interaction surface of the A2 binding site is shown as mesh in silver. The region circled is magnified in **b**. An important hydrophobic contact between the ICL2 of β2AR and Gαs. The ICL2 is labeled. vdW interaction surface formed by H41, V217, F219, F376, and C379 of the Gαs is shown as mesh and colored according to the electrostatic charge of the residues (red: negative; blue: positive; white: neutral). **d** cAMP accumulation assays in the presence of compound 1 (C) or TAK-875 (D). Levels of cAMP are measured as relative luminescence units (RLU). Y114F, S123A, Y44F are site A2 mutations. L112A and L112F are ICL2 mutations involved in Gαs coupling. n = 3. EC50 values are shown in the parentheses. Error bars indicate s.e.m. **e** Measurements in WT extracted from **c** and **d** for ease of comparison.

A2 Site Binding. 10 µL of compound diluted in 100% DMSO and 90 µL of Assay Buffer (50 mM Tris-HCl, pH 7.5, 10 mM CaCl2, 10 mM MgCl2, 0.1% w/v BSA) were added to a deep 96-well polypropylene assay plate (Beckman Coulter) containing 200 µL of GPR40 membranes (5 µg/well), both diluted in Assay Buffer, were added to the assay plate, followed by a 1 min shake and a 1 h incubation at room temperature (22 °C).

Assays were terminated by filtration through GF/C glass fiber filters (Perkin Elmer) presoaked in 50 mM Tris–HCl, pH 7.5, using a Mach III cell harvester (TomTec). Filters were washed two times with 5 µL of ice-cold 50 mM Tris–HCl, pH 7.5 buffer, dried 1 h in a convection oven at 60 °C and embedded in Meltiex A solid scintillant (Perkin Elmer). Radioactivity was determined as counts per min (cpm) using a Trilux Microbeta plate scintillation counter (Perkin Elmer). The equilibrium dissociation constant (Kd) was calculated from the relative IC50 value based upon the equation Kd=IC50/(1+L/Kd), where L equals the concentration of radioligand used in the experiment and Kd equals the equilibrium binding affinity constant of the radioligand, determined from saturation analysis ([3H]-TAK-875 Kd=0.2 nM; [3H]-AM-1638 Kd=1.2 nM).

Ca2+ mobilization assay (FLIPR). HEK293 cells stably overexpressing full length human GPR40 were plated (50K cells/well) into 96-well microtiter plates using Dulbecco’s Modified Eagle’s Medium (DMEM) supplemented with 10% fetal bovine serum (FBS), 20 mM HEPES, pH 7.4, 1 mM Sodium Pyruvate, 100 mM Penicillin, 100 µg/mL Streptomycin, and 800 µg/mL Geneticin. The cells were incubated overnight at 37 °C and 5% CO2. Calcium 4 dye (Molecular Devices) diluted in assay buffer was added (100 µL per well) to the cell plates, followed by a 1 h incubation in the dark at 25 °C. Test compounds were serially diluted three-fold in 100% DMSO and immediately diluted in assay buffer. Diluted compounds were immediately added to the cell plates using the liquid handling capabilities of a FLIPR to achieve final top test compound concentrations of 10 µM for TAK-875 or 5 µM for compound 1 (10-point concentration response curve) at a final DMSO concentration of 1%. Receptor activation was immediately measured as an increase in intracellular calcium using the FLIPR over 3 min. To determine agonist responses, relative fluorescence units (RFUs) over 60 reads were calculated per well and used to calculate percent stimulation relative to 500 µM of the natural ligand, linoleic acid, response. EC50 values were calculated using a four-parameter logistic curve fitting equation.

cAMP accumulation assay. HEK293 cells transiently transfected with Promega’s cAMP response element (CRE/luc2P) and a GPR40 construct were plated (40K cells/well) into 96-well microtiter plates using Dulbecco’s Modified Eagle’s Medium (DMEM) supplemented with 10% FBS, 20 mM HEPES, pH 7.4, 100 µM Penicillin, 100 µg/mL Streptomycin, and 800 µg/mL Geneticin. Test compounds were serially diluted three-fold in 100% DMSO and immediately diluted in media. Diluted compounds were added to the cell plates to achieve final top test compound concentration of 100 µM for compound 1 (10-point concentration response curve) and incubated for 4 h. cAMP levels were indirectly measured using the Promega’s Bright-Glo Luciferase Assay System. EC50 values were calculated using a four-parameter logistic curve fitting equation.

Statistics and reagents of in vitro assays. All in vitro assays were developed according to the procedures outlined in the Assay Guidance Manual28. Technical replicates were performed on separate days and the number of replicates required for statistical significance was based on Minimum Significant Ratios29 established during assay development. Extreme outliers were excluded when they were distant from other observations in the same data set and no pre-established criteria existed.
Animal cell lines were tested and found to be mycoplasma negative (MycoAlert®, Lonza Rockland, Inc).

**Intraperitoneal glucose tolerance test (IPGTT) in lean mice.** Nine to ten week old male C57BL/6 mice and KO mice (Taconic, Hudson, NY) were used. All animals were singly housed in a temperature-controlled (24 °C) facility using a 12-h light/dark cycle (light on 0600) and had ad libitum access to food and water unless otherwise stated. All in vivo experiments were performed in compliance with the policies of the Animal Care and Use Committee of Eli Lilly and Company, in conjunction with the American Association for the Accreditation of Laboratory Animal Care-approved guidelines.

**Incretin secretion assays.** Nine week old male C57Bl/6 mice (Envigo, Indianapolis, IN) were used. The night before the incretin assay, mice were transferred to clean cages and fasted overnight. On the morning of the assay, the mice were weighed and randomized by block randomization into groups based on body weight (N = 6) so each group had similar mean and standard deviation of body weight. Sample size calculation suggests N = 6 per group is sufficient to achieve a power of 80% to detect 50% change in the GLP-1 level. Animals were dosed with vehicle or the test article by oral gavage. Group allocations were not blinded from the experimenters who collected the data. The mice were sacrificed at 0.5, 1, 2, 4, or 8 h post the compound treatment. Blood was taken by cardiac puncture into EDTA tubes containing DPP4 inhibitor and aprotinin cocktail after CO2 euthanization. The final DPP4 inhibitor concentration was 50 μM and aprotinin was 250 K IU/mL. Blood samples were centrifuged at 3000×g for 5 min and plasma transferred into 96 well plates. Plasma total GLP-1 and active GIP levels were measured by ELISA assays that were developed in-house. Data were expressed as mean±s.e.m. Incretin values were log2 transformed as suggested by Box-Cox transformation and analyzed using ANOVA. Cook’s distance was also calculated from linear regression. All animals were included in the analysis with their individual Cook’s distance < 1. Significance is denoted at p < 0.05.

**Computational modeling of AM-1638 and γ-linolenic acid (γ-LA).** γ-LA was modeled in site A2 of GPR40 by minimization in the binding site with two manually imposed hydrogen bond distance constraints from the electron-rich olefin carbons at positions 6 and 12 of the hydrocarbon tail to backbone carbonyl oxygens of residues 99 and 95 on GPR40, respectively. AM-1638 was modeled in Site A2 by alignment onto the bound-state of compound 1, followed by fully unrestrained minimizations. In order to estimate ligand strain for the predicted bound state of γ-LA, compared to the free state, the FreeForm utility within the SYBYL81 application was used. All molecular mechanics minimizations were run using the AMBER10:EHT force field81,82, with a reaction field model treatment of non-bonded electrostatics83,84 using interior and exterior dielectric constants of 1 and 4, respectively, as implemented in the MOE 2014.09 software85.

Fig. 7 Site A2 as a possible free fatty acid binding site. Chemical structure of: a compound 1, b AM-1638, c γ-linolenic acid (γ-LA). Some of the atoms are numbered to show length similarity between compound 1, AM-1638, and γ-LA. d Structural superposition of 4PHU with the GPR40-compound 1 complex (Cα RMSD = 0.99 Å) in ribbon representation. 4PHU is in white with TAK-875 shown in green as space-filling representation. GPR40-compound 1 structure is in cyan, and compound 1 in site A2 is replaced by a model of γ-LA in yellow as space-filling representation. The A2 binding pocket is rendered as Connolly analytic surface86,87, colored based on electrostatic potential: red (negative); blue (positive); white (neutral).
Cloning and baculovirus expression. Human GPR40 encompassing amino acids 1–211+L42A/F88A/G103A/Y202F and amino acids 214–300 with a T4 lysozyme internal fusion (GPR40 numbered relative to reference sequence NP_005294) was TOPO cloned using forward primer: ATGGATTACAAGGATGATGATGA and reverse primer: ATTAATGGTGATGGTGATGATGTA and reverse primer. Expression of GPR40-compound 1 was induced using the Bac-to-Bac system protocol (Life Technologies) in combination with the DH10EMBacY bacmid (Geneva Biotech) was used to generate virus. Fermentations were run at 28 °C for 48 h, harvested by centrifugation, and pellets were stored at -80 °C.

Purification. Frozen cell pellets were resuspended in a lysis buffer containing 20 mM HEPES, pH 7.5, 20 mM NaCl, 0.05% CHS, and 0.05% Triton X-100 (Triton X-100). Cell membranes were isolated by ultracentrifugation in a type 80 Ti rotor (Beckman Coulter) at ~224,000×g for 2 h to allow the receptor–ligand complex to form. A solution of 200 mM n-dodecyl-beta-D-maltoside (DDM; Anatrace) and 2% cholesterol hemisuccinate (CHS; Anatrace) was added to a final concentration of 10 mM DDM and 0.1% CHS, and mixed for 30 min at 4 °C with gentle stirring. The insoluble material was removed by ultracentrifugation, and the supernatant was incubated with TALON resin (Clontech) for 2 h. Resin was packed in a S200 10/300 (GE Healthcare) size exclusion column in 20 mM HEPES, pH 7.5, 300 mM NaCl, 0.01% LMNG, 0.001% CHS, protease inhibitor cocktail, and 10 μM compound 1. The binding pose of TAK-875 from 4PHU was then superimposed in site A1. Crystal packing artifact in GPR40-compound 1 looks to have stabilized a somewhat collapsed site A1 that appears incompatible with TAK-875 binding. To model this ligand’s binding pose, we removed residues 72–86 from transmembrane helix 3 and extracellular loop 1 in GPR40-compound 1, and modeled those residues back using the corresponding residues in 4PHU.

Thermal stability assay. 7-Diethylamino-3-(4-maleimidylphenyl)-4-methylcoumarin (CMPO dye; Adipogene) thermal denaturation experiments were carried out in 0.1 mM PCR strip tubes and measured with a Rotor-Gene Q real-time PCR instrument (Qiagen, Model 6-Plx). Assay was performed in a total volume of 20 μL using 5 μM purified unliganded GPR40 in 10 mM HEPES, pH 7.5, 150 mM NaCl, 1 mM DDM, 0.01% CHS, 50 μM ligand, 25 μM CMPO dye. For the single-channel recordings, the membrane resuspension was incubated with 10 μM of compound 1 for 2 h to allow the receptor–ligand complex to form. A solution of 200 mM n-dodecyl-beta-D-maltoside (DDM; Anatrace) and 2% cholesterol hemisuccinate (CHS; Anatrace) was added to a final concentration of 10 mM DDM and 0.1% CHS, and mixed for 30 min at 4 °C with gentle stirring. The insoluble material was removed by ultracentrifugation, and the supernatant was incubated with TALON resin (Clontech) for 2 h. Resin was packed in a S200 10/300 (GE Healthcare) size exclusion column in 20 mM HEPES, pH 7.5, 300 mM NaCl, 0.01% LMNG, 0.001% CHS, protease inhibitor cocktail, and 10 μM compound 1. The binding pose of TAK-875 from 4PHU was then superimposed in site A1. Crystal packing artifact in GPR40-compound 1 looks to have stabilized a somewhat collapsed site A1 that appears incompatible with TAK-875 binding. To model this ligand’s binding pose, we removed residues 72–86 from transmembrane helix 3 and extracellular loop 1 in GPR40-compound 1, and modeled those residues back using the corresponding residues in 4PHU.

Prior to solubilization, the membrane resuspension was incubated with 10 μM of compound 1 for 2 h to allow the receptor–ligand complex to form. A solution of 200 mM n-dodecyl-beta-D-maltoside (DDM; Anatrace) and 2% cholesterol hemisuccinate (CHS; Anatrace) was added to a final concentration of 10 mM DDM and 0.1% CHS, and mixed for 30 min at 4 °C with gentle stirring. The insoluble material was removed by ultracentrifugation, and the supernatant was incubated with TALON resin (Clontech) for 2 h. Resin was packed in a S200 10/300 (GE Healthcare) size exclusion column in 20 mM HEPES, pH 7.5, 300 mM NaCl, 0.01% LMNG, 0.001% CHS, protease inhibitor cocktail, and 10 μM compound 1. The binding pose of TAK-875 from 4PHU was then superimposed in site A1. Crystal packing artifact in GPR40-compound 1 looks to have stabilized a somewhat collapsed site A1 that appears incompatible with TAK-875 binding. To model this ligand’s binding pose, we removed residues 72–86 from transmembrane helix 3 and extracellular loop 1 in GPR40-compound 1, and modeled those residues back using the corresponding residues in 4PHU.

Prior to solubilization, the membrane resuspension was incubated with 10 μM of compound 1 for 2 h to allow the receptor–ligand complex to form. A solution of 200 mM n-dodecyl-beta-D-maltoside (DDM; Anatrace) and 2% cholesterol hemisuccinate (CHS; Anatrace) was added to a final concentration of 10 mM DDM and 0.1% CHS, and mixed for 30 min at 4 °C with gentle stirring. The insoluble material was removed by ultracentrifugation, and the supernatant was incubated with TALON resin (Clontech) for 2 h. Resin was packed in a S200 10/300 (GE Healthcare) size exclusion column in 20 mM HEPES, pH 7.5, 300 mM NaCl, 0.01% LMNG, 0.001% CHS, protease inhibitor cocktail, and 10 μM compound 1. The binding pose of TAK-875 from 4PHU was then superimposed in site A1. Crystal packing artifact in GPR40-compound 1 looks to have stabilized a somewhat collapsed site A1 that appears incompatible with TAK-875 binding. To model this ligand’s binding pose, we removed residues 72–86 from transmembrane helix 3 and extracellular loop 1 in GPR40-compound 1, and modeled those residues back using the corresponding residues in 4PHU.
ligand assay, 50 μM ligand was used. In the double-ligand assay, 25 μM of compound I and 25 μM of TAK-875 were used. Samples were prepared as triplicates and incubated at 4 °C for 1.5 h in the dark before thermal denaturation. The excitation wavelength was set at 365 nm, and the emission wavelength was at 460 nm. Melts were performed over a temperature range of 25–90 °C, ramping 1 °C every 5 s. Melting curves were processed with GraphPad Prism program (GraphPad-Prism v.6.04). The inflection point of the melting curves was used as the Tm, and was determined using the first derivative values in the Rotor-Gene Q real-time software (v. 2.3.1 (Build 49)).

Crystalization. compound I bound GPR40-T4L complex was mixed with monolein containing 10% cholesterol in 1:1.5 parts v/v protein:lipid ratio using the twin-syringe mixing method55. Using a mosquito LCP crystalization robot (TTP labtech), 50 nL size LCP boles were dispensed onto 96-well glass sandwich plates (Hampton) and overlaid with 0.8 μL of precipitant solution. Crystals reached full size in 10–12 days at 20 °C in the optimized conditions containing 0.1 M Tris-HCl, pH 8.5, 30% PEG 400, 0.2 M ammonium formate. Crystals in LCP were harvested and flash frozen in liquid nitrogen without additional cryoprotectant.

Data collection, processing and structure determination. A complete X-ray data set was collected on a single rod-shaped crystal (~10 x 70 μm) with negligible radiation damage (Supplementary Fig. 7) at 100K in a single sweep of 180 x 1° oscillations and 1.2 s exposure with an unattenuated beam at beamline LRL-CAT (31-ID-D) at the Advanced Photon Source in Argonne National Laboratory, Lemont, IL. The beam size was approximately 80 x 70 μm² FWHM (full width half maximum). The whole sample (i.e. the LCP sample/blob in the MiTeGen loop that contained the crystal) was exposed to the beam during data collection, and loop rastering was not performed. The diffraction data were integrated using autoPROC38,39 and merged and scaled in SCALA18 from the CCP4 suite10. Crystal structure of the GPR40-compound I complex was determined by molecular replacement using the previously solved human GPR40-TAK-875 structure: 4PHU41. GPR40 (intracellular and extracellular loops removed) and lysozyme were used as separate search models. Clear density for the ligand was observed immediately. After numerous cycles of refinement with REFMAC541 and model building with COOT42, the models were refined to reasonable R factors. The structure was validated using MolProbity43. The Ramachandran plot reports 95.7% in the most favored region, 4.3% in the allowed region, and none in the disallowed region. For details, see Table 1.

Synthesis of compound 1. Described in detail in Supplementary Method section.

Data availability. Coordinates and structure factors of the GPR40-compound I X-ray structure have been deposited in the Protein Data Bank under accession number 5KW2. The PDB access codes 4PHU and 3SSN were used in this study. All relevant data are available from the corresponding authors upon request.

Received: 16 May 2017 Accepted: 30 August 2017
Published online: 25 April 2018

References

1. World Health Organization. Global report on diabetes, http://apps.who.int/iris/bitstream/10665/204871/1/9789241565237_eng.pdf?ua=1 (2016).
2. Centers for Disease Control and Prevention. Diabetes Report Card, http://www.cdc.gov/diabetes/pdfslibrary/diabetesreportcard2014.pdf (2014).
3. World Health Organization. World Health Organization: diabetes mellitus fact sheet, http://www.who.int/mediacentre/factsheets/fs138/en/ (2016).
4. Poitout, V. & Lin, D. C. Modulating GPR40: therapeutic promise and potential in diabetes. Drug Discov. Today 16, 1301–1308 (2013).
5. Ballesteros, J. A. & Weinstein, H. Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G protein-coupled receptors. Methods Neurosci. 25, 366–428 (2001).
6. Moro, O., Lameh, J., Hogger, P. & Sadee, W. Hydrophobic amino acid in the i2 loop plays a key role in receptor-G protein coupling. J. Biol. Chem. 268, 22273–22276 (1993).
7. Sugimoto, Y. et al. A cluster of aromatic amino acids in the i2 loop plays a key role for Gs coupling in prostaglandin EP2 and EP3 receptors. J. Biol. Chem. 279, 11016–11026 (2004).
8. Zheng, C. et al. The second intracellular loop of the human cannabinoid CB2 receptor governs G protein coupling in coordination with the carboxyl terminal domain. PLoS ONE 8, e63262 (2013).
9. Eastwood, B. J. et al. The minimum significant ratio: a statistical parameter to characterize the reproducibility of potency estimates from concentration-response assays and estimation by replicate-experiment studies. J. Biomol. Screen 11, 253–261 (2006).
10. SYBYL 1.8.0.2: OpenEye Software, Santa Fe, NM. http://www.eyesopen.com.
11. Case, D. A. AMBER 10 (University of California, 2008).
12. Gerber, P. R. & Muller, K. MAB, a generally applicable molecular force field for structure modelling in medicinal chemistry. J. Comput. Aided. Mol. Des. 9, 251–268 (1995).
13. Omsager, R. Liquid moments of molecules in liquids. J. Am. Chem. Soc. 58, 1486–1493 (1936).
14. Yabuki, C. et al. A novel antidiabetic drug, fasiglifam/TAK-875, acts as an ago-allosteric modulator of FFAR1. J. Biol. Chem. 289, 302 (2014).
15. Liu, J. et al. Structural basis for the cooperative allosteric activation of the free fatty acid receptor GPR40. Nat. Struct. Mol. Biol. 24, 570–577 (2017).
16. Lin, D. C. et al. Identification and pharmacological characterization of multiple allosteric binding sites on the free fatty acid 1 receptor. Mol. Pharmacol. 84, 822–839 (2012).
17. Srivastava, A. et al. High-resolution structure of the human GPR40 receptor bound to allosteric agonist TAK-875. Nature 513, 124–127 (2014).
18. Evans, P. Scaling and assessment of data quality. Acta Crystallogr. D Biol. Crystallogr. 62, 72–82 (2006).
19. Zhang, D. et al. Two disparate ligand-binding sites in the human PZ1 receptor. Nature 520, 317–321 (2015).
20. Jazayeri, A. et al. Extra-helical binding site of a glucagon receptor antagonist. Nature 533, 274–277 (2016).
21. Ballesteros, J. A. & Weinstein, H. Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G protein-coupled receptors. Methods Neurosci. 25, 366–428 (2001).
22. Bartosz, G. & Slomiany, B. A. Melting points of TAK-875 were used. Samples were prepared as triplicates and incubated at 4 °C for 1.5 h in the dark before thermal denaturation. The excitation wavelength was set at 365 nm, and the emission wavelength was at 460 nm. Melts were performed over a temperature range of 25–90 °C, ramping 1 °C every 5 s. Melting curves were processed with GraphPad Prism program (GraphPad-Prism v.6.04). The inflection point of the melting curves was used as the Tm, and was determined using the first derivative values in the Rotor-Gene Q real-time software (v. 2.3.1 (Build 49)).
42. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. Acta Crystallogr. D Biol. Crystallogr. 66, 486–501 (2010).
43. Chen, V. B. et al. MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr. D Biol. Crystallogr. 66, 12–21 (2010).
44. Connolly, M. L. Solvent-accessible surfaces of proteins and nucleic acids. Science 221, 709–713 (1983).
45. Wootten, D. et al. Allosteric modulation of endogenous metabolites as an avenue for drug discovery. Mol. Pharmacol. 82, 281–290 (2012).

Acknowledgements
We thank Han Wu for statistical assistance, and Justine Cudal, Erin Krause, Blake Gagner, Shelby Hedrick, and Derek Mori for technical support. We thank David B. Wainscott and Francis S. Willard for analyzing the data on ligand cooperativity and allostery. We thank Stephen Wasserman and the staff at the Lilly Research Laboratories Collaborative Access Team (LRL-CAT) beamline Sector 31 of the Advanced Photon Source. This research used resources of the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357.

Author contributions
J.D.H. devised the experimental strategy, purified the protein, and solved the structure. B.C. carried out the thermal stability assays. B.C. and L.R. purified the protein. F.L. optimized the crystallization conditions and harvested crystals. C.G., R.E.M., and K.C. carried out the molecular biology and baculovirus expression of GPR40. J.P.L. synthesized compound 1. J.B. collected and processed X-ray diffraction data. S.D.K. and A.P.L. performed the radioligand binding assays. K.L.W. and K.A.O. performed the functional assays. Y.C., M.S., J.P.R., H.-C.Y., and N.A.R. performed the OGTT, IPGTT and incretin secretion assays. T.G. performed the analytical fluorescence size-exclusion chromatography. T.G. and M.D.S. developed the mass spectrometry protocol for analyzing the GPR40 crystallization samples. M.R.L. performed and analyzed the computational modeling of AM-1638, TAK-875, stearic acid and γ-LA. Supervision of the work was carried out by J.D.H, C.M.-R., A.R.M., and C.H. The manuscript was prepared by J.D.H., B.C., L.R., S.D.K., K.L.W., K.A.O., M.R.L., J.P.L., Y.C., J.B., K.A.R., A.R.M., and C.H. All authors contributed to the final editing and approval of the manuscript.

Additional information
Supplementary Information accompanies this paper at https://doi.org/10.1038/s41467-017-01240-w.

Competing interests: All authors are current or past employees of Eli Lilly and Company. The authors declare no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2018